Tensor networks and the numerical renormalization group

Andreas Weichselbaum

Habilitationsschrift

Fakultät für Physik,
Center for NanoScience (CeNS), und
Arnold Sommerfeld Zentrum für theoretische Physik

Ludwig Maximilians Universität

München, Mai 2012

Preface

This *habilitation* project retrospectively started in fall 2004 when after finishing my doctoral studies at Ohio University, I accepted an offer for a postdoctoral position in the group of Prof. Jan von Delft at the Ludwig Maximilians University in Munich. The project was initially for two years with the option for extension. The goal of the project was to bring together two widely successful numerically elaborate methods to analyze strongly correlated quantum systems through renormalization group means: the long-established numerical renormalization group by Kenneth Wilson (NRG, 1975) and its later spin-off, the density matrix renormalization group by Steve White (DMRG, 1992). It had just become apparent then that both methods can be combined in an elegant transparent way on the same algebraic basis in terms of so-called matrix product states. The project therefore required to develop the necessary numerical tools, together with the subsequent question what one method can learn from the other. For me, this was an excellent new field to explore, paired up with my expertise and enthusiasm for numerical algorithms.

In the years since then the field of matrix product states has greatly flourished. The advent of real-time evolution both within the DMRG as well as in the NRG, the conceptual simplification and streamlining of dynamical quantities in the NRG through complete basis sets, as well as the extreme diversification of matrix product states towards tensor networks, all represented exciting new developments in the field. With flexible extensive numerical codes readily developed by myself and several excellent student projects on the way, the question whether to stay in Munich for some time longer was eventually decided in favor of *habilitation*. I am greatly indebted to the many students (Arne Alex, Cheng Guo, Markus Hanl, Andreas Holzner, Theresa Hecht, Wolfgang Münder, Hamed Saberi; Francesco Alaimo, Wael Chibani, Katharina Stadler) and my office mates (Vitaly Golovach and Ireneusz Weymann) for bearing with me, Jan von Delft for his continued support, and to all of them for the many discussions we had, and for giving me the opportunity to be part of this vibrant place for doing research.

Munich, May 2012   Andreas Weichselbaum

In dedication to my parents and my wife Sonali.

Contents

CONTENTS

Abstract

Tensor networks have emerged within the past two decades as a powerful framework to simulate strongly correlated quantum-many-body systems. In particular, it was realized in 2004 that the two powerful and widely successful methods of the numerical renormalization group (NRG) and the density matrix renormalization group (DMRG) shared the same algebraic basis in terms of matrix product states (MPS) which thus served as the starting point for this work. While the NRG is truly based on a renormalization group (RG) ansatz, this is not exactly the case for the DMRG, which by now is considered rather a variational ansatz. Even more so, bringing together these two methods in terms of their shared common algebraic basis has proven a very fruitful and instructive approach. It allowed for a better understanding of the NRG through the quantum information concepts borrowed from DMRG. Strict RG related constraints could be loosened by sidestepping to a variational ansatz. The advent of complete basis sets within the NRG, elegantly formulated in MPS, led us to a clear conceptual simplification and streamlining of dynamical quantities within the NRG. With quantum impurity models the standard realm of NRG, this also has seen an increased interest in DMRG simulations with adapted coarse-graining of macroscopic leads in energy space motivated by the NRG. In view of the extreme recent diversification of one-dimensional MPS into more general tensor networks, finally, we developed a powerful tensor library for arbitrary-rank tensors that can deal with any abelian as well as generic non-abelian symmetries beyond SU(2). Powerful applications within the realm of MPS to cutting edge research in physics are demonstrated, with the application to real two-dimensional physical systems kept as an outlook.

Chapter 1

Introduction

Quantum impurity models represent the systems of central interest to this *habilitation* project. They consist of a small interacting quantum system (the *impurity*) with only a few degrees of freedom in contact with an (effective) macroscopic non-interacting fermionic or bosonic environment. As such they represent a wide class of problems. Both, the impurity as well as the bath, are simple to solve on their own. In the presence of interaction, however, this introduces quantum-many-body correlations, which requires the full treatment of the exponentially large Hilbert space of quantum-many-body states. Besides actual impurities in a host material, quantum impurity models include quantum dot setups, qubit systems coupled to an environment, as well as effective impurity models, such as they emerge from dynamic mean field theory (DMFT).[25,41] Transport experiments through quantum dots and quantum point contacts have become highly controlled tunable settings,[27] which allows for time-dependent manipulation of the *quantum impurity* and real out-of-equilibrium transport experiments far away from linear response.[46] The theoretical understanding of these systems, however, can quickly become highly non-trivial at low temperatures when strong correlations start to play an important role. A similar situation is faced in the rapidly growing field of ultra-cold atoms with temperatures in the range of nano-Kelvins, which also emerges as a highly controllable arena for quantum many-body physics.[19,26,28,73]

The understanding of strongly correlated quantum systems requires a well-controlled systematic way to deal with exponentially large Hilbert spaces. Exact solutions are rare, hence reliable numerical tools are extremely important. With Hilbert spaces quickly beyond what exact diagonalization can handle, examples for quasi-exact numerical methods are quantum Monte Carlo or renormalization group based methods such as the density matrix renormalization group (DMRG)[70] or the numerical renormalization group (NRG).[72] Each of these methods has its specific realm where it works particularly well. In this work, the emphasis is on quantum impurity models for which the method of choice is typically the NRG. The NRG is constructed to deal with a small interacting quantum system (the *impurity*) that interacts with a non-interacting (effective) bath, allowing for dynamically generated arbitrarily-small low-energy scales. Nevertheless, it becomes quickly expensive numerically, if the complexity of the impurity is increased or if several bath channels are added. Hence formulating NRG in a way that reveals its common algebraic and thus numerical basis with the variational approach intrinsic to the DMRG, this can lead to powerful mixed settings in which the tools and concepts from both methods are combined. The latter defines the overall theme of this *habilitation* project.
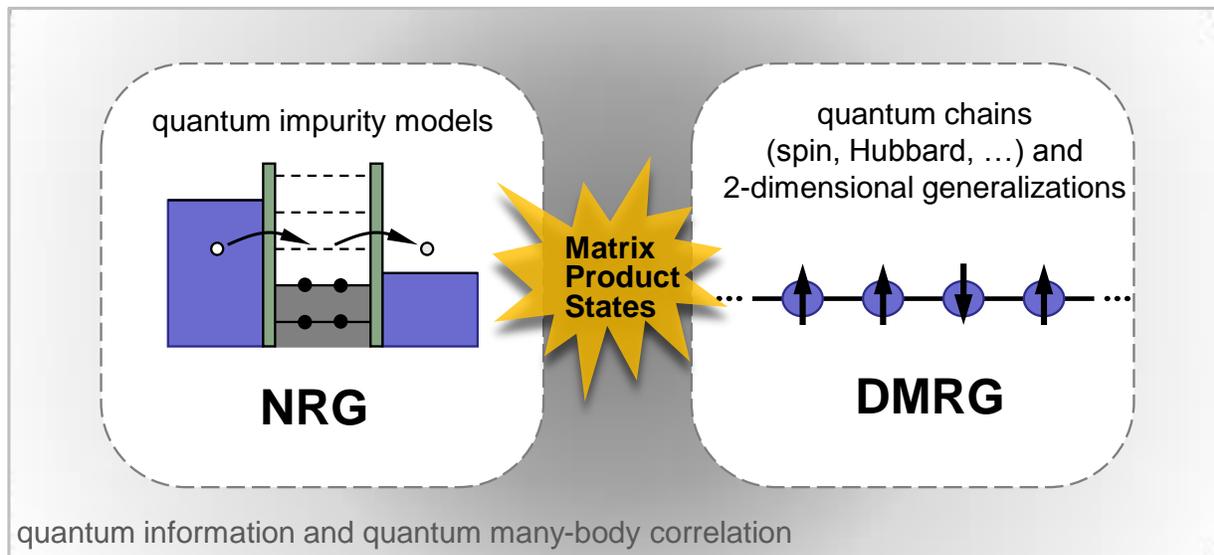
Figure 1.1: Common underlying algebraic structure of NRG and DMRG in terms of MPS. While NRG is exclusively constructed for quantum impurity models, DMRG is usually associated with interacting real-space lattices such as spin or Hubbard chains, with intensely studied generalizations to two- or higher-dimensional systems. Both methods live in the algebraic space of MPS or their generalizations to tensor-network states. The latter represent a vehicle of central interest in the analysis of quantum information and quantum many-body correlations.

1.1 Research highlights

At the start of this work in 2004, the numerical renormalization group (NRG) and density matrix renormalization group (DMRG) were two separate disciplines. Nevertheless, it had just been realized then that these two widely successful methods shared a common algebraic basis in terms of so-called matrix product states. Therefore the goal of this work was to bring both methods together and investigate their mutual benefits. The following lists the major highlights that were part of this research. Publications that are reprinted in this thesis together with a more extended introduction, are marked by their respective section **Pn** (for paper Pn) in the appendix.

▶ **Complete basis sets and correlation functions in the NRG**

In 2005, Anders and Schiller introduced a very important technical innovation within NRG: they showed[2] that based on energy scale separation it is possible to construct a complete many-body basis of approximate eigenstates of the full Hamiltonian, and used this to study real-time evolution after a quantum quench. Nevertheless, it turned out that the concept is significantly more generic. In particular, these complete basis sets allow to calculate dynamical quantities such as correlation functions for arbitrary

temperatures in a black-box like algorithm which is in stark contrast to prior patching schemes. The MPS framework proved enormously useful in its understanding and setup. The MPS diagrammatics described in Chap. 4 will be published in another independent publication.

The first generic application to spectral functions was published in **P1**
Sum-Rule Conserving Spectral Functions from the Numerical Renormalization Group
Andreas Weichselbaum, and Jan von Delft, Phys. Rev. Lett. **99**, 076402 (2007)

▶ **DMRG-enhanced NRG**

While the NRG per se is not applicable to typical uniform lattice models that are analyzed by the DMRG, the reverse, of course, holds: DMRG *can* be used to analyze Hamiltonians in the Wilson chain setup.

Due to the NRG's inherent constraint of logarithmic discretization for the sake of energy scale separation, NRG has finite spectral resolution at finite frequencies. This is problematic if sharp features at finite frequency occur. Nevertheless, this can be strongly improved upon by using a tailored discretization. However, this rules out the iterative NRG prescription which thus must be replaced by DMRG specific quasi-variational methods. This was demonstrated in detail for the Kondo model in the presence of a strong magnetic field, $B \gg T_K$, *i.e.* much larger than the Kondo temperature T_K, in the paper **P2** *Variational matrix-product-state approach to quantum impurity models*, A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and Jan von Delft Phys. Rev. B 80, 165117 (2009) [cond-mat/0504305v2, (2005)].

Having a common algebraic basis, one of our very first questions we were interested in was: can one improve upon the one-way prescription of the NRG by (subsequent) variational sweeping? This was analyzed in much detail by the Ph.D. student Hamed Saberi in the publication *Matrix-product-state comparison of the numerical renormalization group and the variational formulation of the density-matrix renormalization group*, Hamed Saberi, Andreas Weichselbaum, and Jan von Delft Phys. Rev. B **78**, 035124 (2008). The conclusion was that in the regime where NRG works efficiently, the NRG is *essentially equivalent* to the variational setup of the DMRG.

Importantly, the common algebraic basis allows to carry over to the NRG much of the standard analysis of MPS within the DMRG. These include entanglement entropy measures, as well as concepts such as entanglement spectra or discarded weights. The latter were introduced for the first time to the NRG arena in the paper **P3**
Discarded weight and entanglement spectra in the numerical renormalization group
Andreas Weichselbaum Phys. Rev. B **84**, 125130 (2011).

▶ **NRG-enhanced DMRG**

A simple obvious feedback from the NRG to the DMRG was the (logarithmic) discretization of the bath in energy for quantum impurity setups. Rather than using the prevalent plain tight-binding chain as the model for non-interacting leads, which has

the largest density of states at the band edges, tailored logarithmic discretization allows to coarse-grain high energies that are less relevant for the dynamics. This can be used both in in-[8] and out-of-equilibrium in the low-energy sector.[9]

Strict logarithmic discretization allows to accurately describe strongly-correlated low-energy physics and resolve phase boundaries. This can also be of interest for the DMRG in certain contexts. In particular, this applies to the spin-boson model. Despite its conceptual simplicity, the spin-boson model with a sub-ohmic spectral distribution of the bath is highly non-trivial and still controversial at finite temperature. While bosonic NRG exists, it naturally runs into problems if the bosonic state space per Wilson site is occupied up to very high excitations. The variational MPS approach provides a strong alternative then. A detailed analysis of the spin-boson model using variational MPS was performed in an independent project with the Ph.D. student Cheng Guo. *Critical and Strong-Coupling Phases in One- and Two-Bath Spin-Boson Models*, Cheng Guo, Andreas Weichselbaum, Jan von Delft, and Matthias Vojta Phys. Rev. Lett. **108**, 160401 (2012).

▶ **FDM-NRG applications**

The complete basis sets by Anders and Schiller (2005) turned out a fruitful ground for many applications. Specifically, from a quantum impurity point of view, any thermodynamical quantity that can be written down in Lehmann representation is amenable to the FDM-NRG approach for arbitrary temperature.

A consistent generic approach to arbitrary temperatures is important when theoretical computations are related to actual experiments. The power of this was demonstrated in larger collaboration that targeted the until then still controversial question regarding the microscopic model for the very original system where the Kondo effect was observed, namely iron impurities in gold or silver.[20]

P4 *Kondo Decoherence: Finding the Right Spin Model for Iron Impurities in Gold and Silver*; T. A. Costi, L. Bergqvist, A. Weichselbaum, J. von Delft, T. Micklitz, A. Rosch, P. Mavropoulos, P. H. Dederichs, F. Mallet, L. Saminadayar, and C. Bäuerle Phys. Rev. Lett. **102**, 056802 (2009).

Absorption and emission spectra, which combine initial and final Hamiltonian, have been studied early on within the NRG.[30,45,53] Nevertheless, it had been significant patch work still. By realizing that the standard Fermi-Golden rule expression corresponds to a Lehmann representation, however, absorption or emission spectra are ideally suited for the FDM-NRG framework. This resulted in an excellent collaboration with Hakan E. Türeci and the experimental group of Atac Imamoglu, with the two publications,

Many-Body Dynamics of Exciton Creation in a Quantum Dot by Optical Absorption: A Quantum Quench towards Kondo Correlations, Hakan E. Türeci, M. Hanl, M. Claassen, A. Weichselbaum, T. Hecht, B. Braunecker, A. Govorov, L. Glazman, A. Imamoglu, and J. von Delft, Phys. Rev. Lett. **106**, 107402 (2011).

P5 *Quantum quench of Kondo correlations in optical absorption*
Christian Latta, Florian Haupt, Markus Hanl, Andreas Weichselbaum, Martin Claassen, Wolf Wuester, Parisa Fallahi, Stefan Faelt, Leonid Glazman, Jan von Delft, Hakan E. Türeci, and Atac Imamoglu, Nature **474**, 627 (2011).

Absorption spectra are intrinsically related to the phenomenon of Anderson orthogonality (AO). It turns out that the MPS framework is ideally suited for a fully self-contained, accurate, and transparent study of AO.

P6 *Anderson orthogonality and the numerical renormalization group*, Andreas Weichselbaum, Wolfgang Münder, and Jan von Delft, Phys. Rev. B **84**, 075137 (2011).

Anderson Orthogonality in the Dynamics After a Local Quantum Quench, W. Münder, A. Weichselbaum, M. Goldstein, Y. Gefen, and J. von Delft, cond-mat/1108.5539v2 (accepted by PRB).

▶ **Non-abelian symmetries in tensor-networks (QSpaces)**

This *habilitation* project required to develop a proper flexible and suitable numerical framework to matrix product states. Initially implemented in terms of a general tensor library for abelian symmetries only, dubbed QSpaces, this has been recently extended to include non-abelian symmetries in a completely generic fashion. By construction of a tensor library for arbitrary rank tensors then, QSpaces can equally well deal with NRG as well with DMRG settings. Moreover, the framework should also be straightforwardly applicable to more recent tensor network setups such as MERA[64] or PEPS[43], as long as *all* (effective) state spaces can be represented in terms of well-defined symmetry labels. The code includes about 70,000 lines of objected oriented C++ code, which makes use of the highly optimized Lapack and BLAS routines for linear algebra wherever possible. The code is mostly deployed from within MatLab, including model setup, pre- and postprocessing of data and wrapper routines through mex-files (about 20,000 lines of MatLab code). All of this code was fully developed by myself for this project.

The general conceptual framework for treating non-abelian symmetries has been written up in much detail with an extended pedagogical appendix in the paper
P7 *General framework for non-abelian symmetries in tensor networks*
A. Weichselbaum, cond-mat/1202.5664v1 (submitted to Annals of Physics).

This includes a successful demonstration of the QSpace tensor library for non-abelian symmetries on the symmetric three-channel model introduced in P4. While initially analyzed using abelian symmetries only, this turned out insufficient for general model parameters. Nevertheless, the model has strong non-abelian symmetries which thus served as major motivation for the general implementation of the non-abelian symmetries. For comparison, various non-abelian symmetry settings are analyzed for exactly the same model Hamiltonian. These range from the more traditional symmetry setting based on plain SU(2) symmetries, $\mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)_{charge}^{\otimes 3}}$, to the explicit treatment of the SU(3) channel symmetry in $\mathrm{SU(2)_{spin}} \otimes \mathrm{U(1)_{charge}} \otimes \mathrm{SU(3)_{channel}}$, to the actual full symmetry of the model, $\mathrm{SU(2)_{spin}} \otimes \mathrm{Sp(6)}$. By going from state space representation to

multiplet spaces, a tremendous reduction in the reduced coefficient spaces is achieved, which renders these calculations feasible to start with. A follow-up paper to P4 and P7 analyzing the experimentally measured magnetic field dependence of conductivity is in preparation (Markus Hanl et al.).

1.2 Overview

The remainder of this cumulative *habilitation* thesis splits up in three parts.

(i) A brief introduction to matrix product states and their related algebra specific to NRG and DMRG (Chap. 2), followed by a short introduction of the essentials of NRG (Chap. 3). A detailed discussion of general non-abelian symmetries in tensor networks has been compiled into the independent paper **P7**. Hence the discussion of (non-)abelian symmetries is kept at a minimum here.

(ii) A more extended introduction and discussion of the actual tensor networks that emerge when using complete basis sets (Chap. 4, *MPS diagrammatics for the numerical renormalization group*): matrix product states can be understood simply as a specific language to describe the NRG. While this does not change the basic underlying principles of the NRG per se, the MPS framework, nevertheless, proves extremely useful for the discussion and implementation of complete basis sets.

(iii) Reprints of selected papers that have been published in peer reviewed journals as part of this *habilitation*, together with a short introduction and accompanying remarks for each.

While this thesis is mainly concerned about matrix product states and their application, its title *tensor networks and the numerical renormalization group* has been chosen more general nevertheless for two reasons: (1) the application of complete basis in part (ii) does generate more complex networks [see Chap. 4], and (2) by construction, the QSpace library in **P7** which introduces non-abelian symmetries to tensors of arbitrary rank, clearly also can deal with higher-rank tensors, assuming that all involved (effective) state spaces allow a well-defined categorization in terms of symmetry sectors. Within the NRG, tensors up to rank equal to 4 were used frequently.

Chapter 2

Matrix Product States

In the following a brief basic introduction to matrix product states (MPS) is given, while also introducing the corresponding notation and conventions. The discussion is based within the framework of the density matrix renormalization group (DMRG)[70] where historically the term MPS was coined[50,55,56] [see Schollwöck (2011) for an extensive review on this topic]. Much of the language is further borrowed from the quantum information community which also put much effort into the analysis of MPS and their entanglement properties.[58,63,74] In particular, I am indebted to Frank Verstraete who kindly introduced me to MPS by providing me with an initial basic MPS code many years back. The typical essential MPS steps as well as the corresponding semantics I learned from there, which therefore is also reflected in what follows. While MPS concepts have been significantly extended to a large variety of tensor networks in recent years,[22,43,64] this chapter will only deal with the essential concepts already also present within MPS. The connection to other numerical quantum-many-body methods such as the Numerical Renormalization Group (NRG)[72] is straightforward[51,62,65,66], and will be discussed in much more detail later. In this introduction, however, the DMRG context will be emphasized for simplicity. While MPS is based on an underlying one-dimensional structure, more complex MPS network structures quickly emerge in practice (*cf.* Chap. 4 on MPS diagrammatics).

Consider a one-dimensional physical chain of length N with local and nearest-neighbor interaction only. The dimension of the local state space σ_n on site n is denoted by d and considered the same for all sites, for simplicity. Typically, d is small, *e.g.* $d = 2$ for a spinless fermionic degree of freedom. The total Hilbert space then of dimension d^N is spanned by
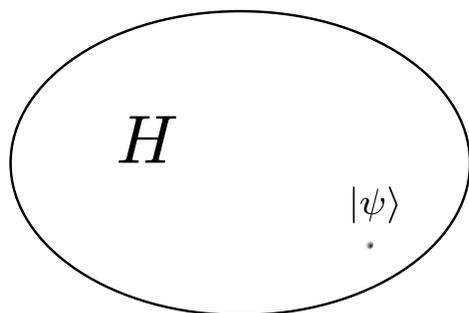

Figure 2.1: "Hilbert space is huge" – schematic depiction of a possibly strongly correlated quantum many-body state (dot), *e.g.* a matrix product state, within the full Hilbert space H.

the Fock space (product states)

$$|\psi\rangle \;=\; \sum_{\sigma_1,\ldots,\sigma_N}^{d} c_{\sigma_1\ldots\sigma_N} \underbrace{|\sigma_1\rangle \otimes \ldots \otimes |\sigma_N\rangle}_{\equiv |\sigma_1,\ldots,\sigma_N\rangle \equiv |\sigma\rangle} \equiv \sum_{\sigma} c_{\sigma}^{(N)} |\sigma\rangle. \tag{2.1}$$

The order in the product state $\sigma \equiv \{\sigma_1,\ldots,\sigma_N\}$ is important when dealing with fermionic systems. The quantum many-body state $|\psi\rangle$ represents a tiny point of the entire Hilbert space, as schematically depicted in Fig. 2.1. This state may carry any amount of bipartite or multipartite entanglement which typically scales with system size. Specific states such as ground states, however, obey an area law.[7,22,74] It states that the entropy of a contiguous q-dimensional subsystem of volume L^q scales only like the surface L^{q-1} separating the chosen region from the rest of the system. For one-dimensional systems, the surface that separates a contiguous block of length L of the rest of the system are two points, *i.e.* the left and right boundary of the block. This is a surface that is independent of the block size L, hence the entropy saturates with increasing L. Strictly speaking, this is the case for gapped systems only, while for critical systems, *i.e.* non-gapped systems, the entropy acquires logarithmic corrections of the type $S \propto \ln L$. Therefore the entropy grows *much* slower than an extensive quantity which would scale with the system size. The entropy S then can be related to an effective state space dimension $D^{\mathrm{eff}} \propto e^S$ required in a numerical description of such a state. This suggests that the numerical resources required in the description of a one-dimensional state obeying an area law, worst case, scales with the length of the system, $e^{\ln L} \sim L$, which thus is polynomial and not exponential with system size.

This underpins the efficiency of matrix product states (MPS)[50] in the simulation of strongly correlated quantum many-body physics. An MPS is in a sense a convenient and compact way of writing a general a priori unknown quantum state in a one dimensional system, with the coefficients of Eq. (2.1) condensed into a set of matrices,

$$|\psi\rangle_Q^{\mathrm{MPS}} \equiv \sum_{\sigma_1,\ldots,\sigma_N} \mathrm{tr}\Big(Q \prod_{n=1}^{N} A^{[\sigma_n]} \Big) |\sigma_1,\ldots,\sigma_N\rangle. \tag{2.2}$$

In particular, every one of the d local states σ_n has a matrix $A^{[\sigma_n]}$ associated with it, resulting in a total of $d \cdot N$ so-called A-tensors. The maximum matrix dimension in constrained by some upper cutoff dimension D, typically ranging from a few tens to a several thousands.

The matrix Q in Eq. (2.2) describes the boundary conditions[50] of the two ends of the chain. In the case of periodic boundary conditions, one may simply take a D-dimensional identity matrix. For an open chain, on the other hand, as considered throughout this work, Q can be considered $Q \equiv |a\rangle\langle b|$ with two arbitrary vectors a and b in the D-dimensional matrix space, which thus disconnects the chain at the boundary. In general, for open boundary conditions, the vectors a and b may be merged with the A-tensors of the first and the last site. Therefore the matrices $A^{[\sigma_1]}$ ($A^{[\sigma_N]}$) become column (row) vectors, respectively. With open boundary conditions the Hilbert space of the first few sites at each end can be
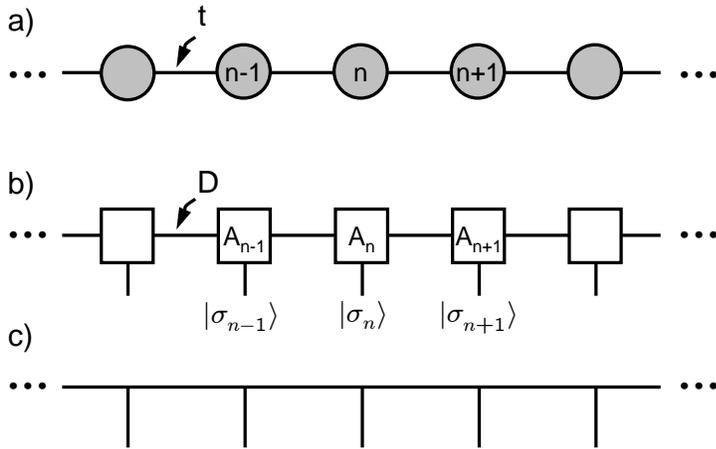
Figure 2.2: Panel (a) Schematic depiction of a quantum chain of a Hamiltonian with nearest neighbor interaction strength t. Panel (b) shows the numerical simulation of the same in terms of matrix product state. In a diagram as shown in panel (b), in general, blocks refer to coefficient spaces, and lines to indices, which themselves correspond to well-defined state spaces. Open lines are open indices, while connecting lines are indices summed over. Thus the horizontal lines refer to the effective D-dimensional state space, with their summation summarized in the product in Eq. (2.3). The vertical lines connect to the local state spaces $|\sigma_n\rangle$, and thus represent the summation over $|\sigma_n\rangle$ in Eq. (2.3). For simplicity, the local state space labels $|\sigma_n\rangle$ are considered implicit and will be dropped in most cases, except for the ones required for the discussion. Furthermore, also the size of the coefficient blocks will be shrunk to zero in later discussions where convenient. In this sense, the resulting diagram in panel (c) is equivalent to the diagram in panel (b).

included exactly, leading to $\dim\big(A^{[\sigma_1]}\big) = 1 \times d$, $\dim\big(A^{[\sigma_2]}\big) = d \times d^2$, and so on until the dimension D is reached [here $\dim(\cdot)$ stands for the dimension of the specified algebraic object]. The situation is completely analogous for the right boundary starting at site N, except that the order is reversed, with the effect that row- and column-dimensions are flipped. The trace over the product of A-tensors then results in a plain number, *i.e.* the coefficient for a given state, and hence the trace in Eq. (2.2) becomes irrelevant, leading to

$$|\psi\rangle^{\mathrm{MPS}} \equiv \sum_{\sigma_1,\ldots,\sigma_N} \Big(\prod_{n=1}^{N} A^{[\sigma_n]}\Big)|\sigma_1,\ldots,\sigma_N\rangle. \tag{2.3}$$

This description of quantum-many body states emerges completely naturally from the iterative prescription of numerical methods such as the DMRG or the NRG.

A pictorial representation of an MPS together with the background topology of the underlying one-dimensional physical Hamiltonian is depicted in Fig. 2.2. In particular, Fig. 2.2(c) represents an essential building block for more complex tensor networks discussed later. Panel (a) depicts the physical quantum chain to be studied. The links between the sites indicate nearest-neighbor interactions, *e.g.* described by a hopping amplitude t. Panel (b) shows the numerical description of this quantum chain in terms of an MPS, with the sets of A-tensors $\{A^{[\sigma_n]}\}$ for a specific site referred to as A_n, in short. Note that there is a

clear one-to-one correspondence of the physical system in panel (a) with the A-tensors of the underlying MPS structure in panel (b). This directly derives from the iterative growth of entire blocks of the chain one site at a time in the numerical simulation. In that the MPS follows the directed structure of the Hamiltonian, this also keeps correlations as localized as possible within the MPS.

The set of A-tensors for each site n in Fig. 2.2(b) is described by the rank-3 tensor A_n with the three indices i, j, σ_n, *i.e.* $(A^{[\sigma_n]})_{ij} \equiv A_{ij}^{[\sigma_n]}$. Pictorially these are represented by *legs* attached to each coefficient block A. The horizontal connected legs are summed over (*contracted*) by standard matrix multiplication to the tensors to the left and to the right, while the third leg downwards links to the local state space σ_n. In principle, the vertical lines are also summed over through Eq. (2.3). From a numerical point of view, however, they appear open since the label $|\sigma_n\rangle$ itself has no numerical meaning other than representing a finite dimensional "open" state space. Only through the actual evaluation of matrix elements or scalar products with other states will state spaces become "closed", *i.e.* connected to other actual coefficient spaces.

2.1 MPS and tensor network diagrammatics

Given an MPS description, tensors of different rank are connected to each other by contracting mutual state spaces, *i.e.* summation over common indices. The explicit notation can become quickly cumbersome with an excessive number of indices when spelled out in detail. In this sense a graphical representation is usually preferred when describing operational procedures. Typical examples were already encountered in the previous basic introduction to MPS, *cf.* Fig. 2.2. The simple basic rules are as follows.[56,63]

- Coefficient spaces are represented by blocks of any shape, possibly reduced to nodes of lines for better readability [*e.g.* see Fig. 2.2(b+c)].

- Indices represent actual state spaces, which are indicated by lines connected to coefficient spaces. The number of lines connected to a coefficient space describes its rank. Closed lines which connect two coefficient spaces, are considered contracted, *i.e.* summed over. Open lines, finally, describe the overall rank and dimensionality of the entire object depicted in the diagram.

A further major advantage of these diagrams is that they *exactly* describe the numerical operations that need to be performed in practice. Given a diagram with a multitude of contractions, these are always performed sequentially, resulting in intermediate temporary objects of altering ranks. A specific elementary contraction then is performed in between two objects that refer to a common set of state spaces. Typically, both objects are reshaped into temporary matrices which then allows to use public highly-optimized linear algebra libraries to perform the tensor-contraction in terms of a standard matrix multiplication.
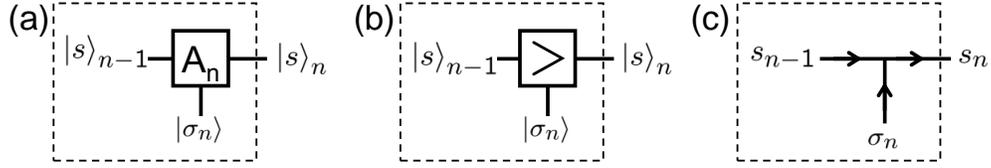
Figure 2.3: Basic MPS iteration step in terms of A-tensor. All panels show essentially the same, yet emphasize different aspects. Panel (a) generic A-tensor which links fully orthonormal state spaces. Panel (b) shows the special case where the state space to the left and and at the bottom (input spaces) are combined, resulting in the output space to the right. This provides a *directed* structure: the arrow within the coefficient block points towards the output space. Panel (c) shows exactly the same as panel (b) yet with simplified notation, in that the A-tensor is shrunk to a node, and input and output spaces indicated by arrows. Mixed notations may be used throughout the discussions that follow. In all panels, the outer dashed box indicates an effective state space with only one open index, *i.e.* the effective state space $|s_n\rangle$ of all sites up to and including site n.

2.2   Orthonormalization and effective state space

An MPS derives from the generic iterative procedure that, given an effective description of part of the system say up to site $n-1$ in terms of the (many-body) basis $|s_{n-1}\rangle$, a new site with local state space $|\sigma_n\rangle$ is added. This results in a typically truncated effective new state space $|s_n\rangle$,

$$|s_n\rangle = \sum_{s_{n-1},\sigma_n} A^{[\sigma_n]}_{s_{n-1},s_n} |\sigma_n\rangle \otimes |s_{n-1}\rangle. \tag{2.4}$$

The coefficients $A^{[\sigma_n]}_{s_{n-1},s_n}$ link to the underlying full product space $|s_{n-1}\rangle \otimes |\sigma_n\rangle$. The corresponding rank-3 tensor will be referred to as A-tensor, with its elementary MPS diagram depicted in Fig. 2.3.

By construction, the input states $|s_{n-1}\rangle$ and $|\sigma_n\rangle$, as well as the output states $|s_n\rangle$ form orthonormal basis sets. Therefore $A^{[\sigma_n]}_{s_{n-1},s_n}$ is part of a unitary transformation, with the property

$$\langle s_n|s'_n\rangle = \sum_{s_{n-1},\sigma_n} A^{[\sigma_n]*}_{s_{n-1},s_n} A^{[\sigma_n]}_{s_{n-1},s_n} \overset{!}{=} \delta_{ss'}. \tag{2.5}$$

Its MPS diagram is shown in Fig. 2.4(a). With s_n drawn to the left and s_{n+1} drawn to the right, this is referred to as left-to-right (LR) orthonormalization. This *directedness* is also indicated by the arrow within the A-tensor (*cf.* Fig. 2.3(b)), or more generally by the arrows on input and output spaces in Fig. 2.3(c).

Conversely, the tensor A_n may have been interpreted as an iterative step coming from

(a) $\sum_{\sigma_n} A^{[\sigma_n]\dagger} A^{[\sigma_n]} = 1$ (b) $\sum_{\sigma_n} A^{[\sigma_n]} A^{[\sigma_n]\dagger} = 1$

Figure 2.4: Orthonormalization property of A-tensors. Panel (a) shows a left-to-right (LR) orthonormalized A-tensor. The mathematical property is shown at the top, while the equivalent MPS diagram is shown at the bottom. The symbol '>' inside the box indicates the direction of orthonormalization, *cf.* Fig. 2.3(b). The asterisk indicates complex conjugation as in Eq. (2.5), which is irrelevant for time-reversal symmetric problems. The resulting bare bracket for the identity tensor to the lower right of panel (a) shows, that the entire contraction of A_n and A_n^* to the left can be eliminated (*short circuited*) by directly connecting from s to s' (for an explicit example, see also Fig. 2.8 below). Panel (b) shows the reversed situation of a right-to-left (RL) orthonormalized A-tensor.

the right end of the system, starting from the last site N. With

$$|\tilde{s}_n\rangle = \sum_{\tilde{s}_{n+1}, \sigma_n} A^{[\sigma_n]}_{\tilde{s}_n, \tilde{s}_{n+1}} |\sigma_n\rangle \otimes |\tilde{s}_{n+1}\rangle. \tag{2.6}$$

the state space $|\tilde{s}_{n+1}\rangle$ then provides an effective representation of the system $n' > n$ to the right of site n. Thus adding site n with its state space σ_n to the effective previous basis $\tilde{s}_{n+1}$, the resulting basis $\tilde{s}_n$ is described by a different A-tensor which is right-to-left (RL) orthonormalized, instead, as depicted in Fig. 2.4(b).

2.2.1   State space orthonormalization

Orthonormal basis sets are desirable quite generally in numerical calculations and specifically with MPS, as they make algorithms efficient and stable in the presence of state-space truncation. For this, consider the generic MPS setup in Fig. 2.5(a), where the state $|\psi\rangle$ is defined on a linear chain with open boundary conditions. Here the A-tensors for sites $n' < n$ ($n' > n$) are considered LR (RL)-orthonormalized, respectively. The tensor A_n for the so-called *current site* n plays a special role, in that it links the effective state spaces $|s_{n-1}\rangle$ and $|\tilde{s}_{n+1}\rangle$ for sites $n' < n$ and $n' > n$, respectively, with the full local state space $|\sigma_n\rangle$ for site n,

$$|\psi\rangle = \sum_{\substack{s_{n-1}, \sigma_n \\ \tilde{s}_{n+1}}} \underbrace{A^{[\sigma_n]}_{s_{n-1}, \tilde{s}_n}}_{\equiv \Psi_{(\sigma_n, s_{n-1}), \tilde{s}_{n+1}}} |\sigma_n, s_{n-1}\rangle \otimes |\tilde{s}_{n+1}\rangle. \tag{2.7}$$
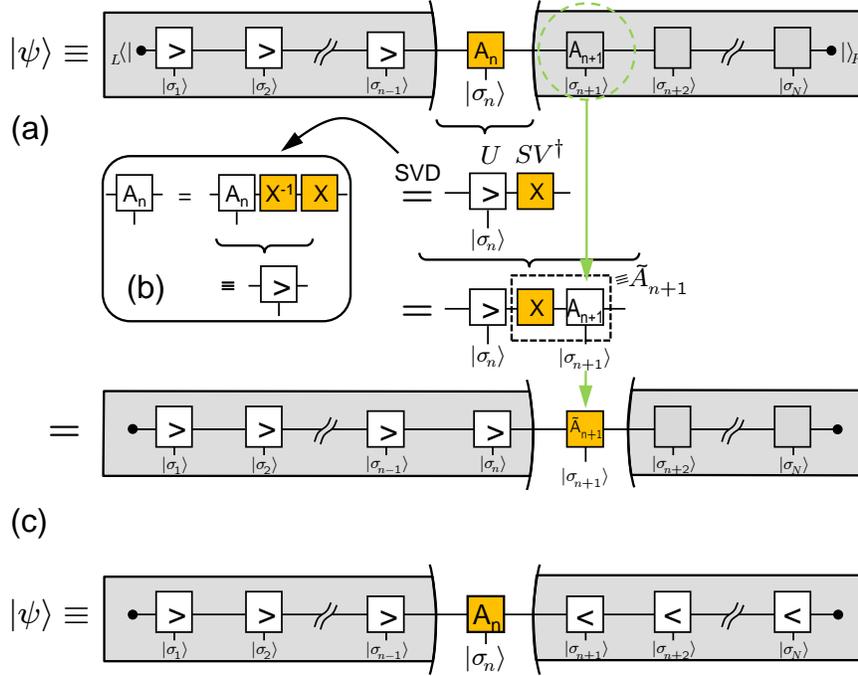
Figure 2.5: Panel (a) Local orthonormalization step of a given MPS at current site n. This shifts the focus on to the next current site $n + 1$ using singular value decomposition (SVD). Having open boundary conditions, the left (right) end of the MPS is terminated by the left (right) vacuum state, respectively, as indicated in the upper MPS diagram. For simplicity, this will be indicated by a terminating bullet for the rest of the paper as shown in the lower MPS diagrams. Panel (b) indicates the equivalence of the orthonormalization step to inserting $X^{-1}X$ at the link between A_n and A_{n+1}. Panel (c) shows current site n now with fully orthonormalized state spaces, *i.e.* including an RL-orthonormalization for sites $n' > n$. The coefficient space A_n links the state spaces together into the actual wave function $|\psi\rangle$. Representing the current site, it is shaded in color in order to distinguish it from the orthonormalized A-tensors which are kept in white and whose direction of orthonormalization is indicated (*cf.* Fig. 2.3).

The tilde for the states $|\tilde{s}_{n+1}\rangle$ for the sites $n' > n$ serves to differentiate from the effective description $|s_{n-1}\rangle$ for the sites $n' < n$. In this sense, the state space $|s_n\rangle$ refers to the sites $n' \in \{1, \dots, n\}$, whereas $|\tilde{s}_n\rangle$ refers to the sites $n' \in \{n, \dots, N\}$. By combining the effective state space $|s_{n-1}\rangle$ with the local state space $|\sigma_n\rangle$, their product-space index range can also be *fused* into a single hyperindex, leading to the coefficient matrix $\Psi_{(\sigma_n, s_{n-1}), \tilde{s}_{n+1}}$, as defined in Eq. (2.7).

The A-tensor, as defined in Eq. (2.7) for the current site, does not fulfill the generic orthonormalization as in Fig. 2.4, yet it does encode the global normalization condition of

the wave function,

$$\langle\psi|\psi\rangle = \sum_{\sigma_n} \mathrm{tr}\,(A^{[\sigma_n]\dagger}A^{[\sigma_n]}) \equiv \mathrm{tr}\,(\Psi^\dagger\Psi) \stackrel{!}{=} 1. \tag{2.8}$$

Assuming orthonormal state spaces for all legs of the tensor A_n, the elementary orthonormalization can be performed now in either direction of the MPS. In the following, the focus will be on LR-orthonormalization, while RL-orthonormalization works in complete analogy, by consistently reversing the role of left and right, *i.e.* essentially working on L $\leftrightarrow$ R reflected version of the MPS.

Even if all A-tensors in Fig. 2.5(a) had been initialized randomly, thus compromising the interpretation of orthonormal effective state spaces for the horizontal lines, the state $|\psi\rangle$ can nevertheless be iteratively LR-orthonormalized by *internal* reorganization of the MPS without physically changing the state.[63] The starting point for this is the very left end of the system. As indicated in Fig. 2.5(a), the effective state space from the previous iteration for the very first A-tensor is the left vacuum state $|\rangle_\mathrm{L}$. By definition, this single state is a proper orthonormal state space. This therefore sets the starting point for the iterative prescription for LR-orthonormalization starting with $n = 1$. Given a proper orthonormal *effective* basis of the system up to and including site $n-1$, the orthonormal local state space of site n can be added. The coefficient space A_n at this point can be arbitrary. It has three legs, $A^{[\sigma_n]}_{s_{n-1},\tilde{s}_n}$. The two state spaces to be combined, $|s_{n-1}\rangle$ and $|\sigma_n\rangle$ are already orthonormal, by construction. The third index $\tilde{s}_n$ does not necessarily have to describe a orthonormal state space $|\tilde{s}_n\rangle$ yet, which is useful for initialization. In general, the (not necessarily normalized) state $|\psi\rangle$ can be written as in Eq. (2.7).

Tracing out all sites $n' > n$, it is straightforward to obtain a reduced density matrix for the system up to and including site n, $\hat{\rho}_n^\mathrm{red} \equiv \mathrm{tr}_{\sigma_{n+1},\ldots,\sigma_N}\big(|\psi\rangle\langle\psi|\big)$. By diagonalizing this density matrix, the dominant states within the combined space $|\sigma_n, s_{n-1}\rangle$ can be obtained. Nevertheless, a more compact superior approach is provided through the *Schmidt decomposition*, generally also referred to as singular value decomposition (SVD).[55] With the indices (σ_n, s_{n-1}) fused to obtain a plain rank-2 coefficient matrix Ψ, *cf.* Eq. (2.7), SVD decomposition leads to

$$\Psi = U\underbrace{SV^\dagger}_{\equiv X}. \tag{2.9}$$

The matrices U and V are column-orthonormal, whereas S is a diagonal matrix with real and positive diagonal elements, *i.e.* the so-called singular values λ_α. If Ψ is an $m \times n$ matrix, then with $k \equiv \min(m, n)$ the matrices U, S, and V are of dimension $m \times k$, $k \times k$,

and $n \times k$, respectively. Substituting Eq. (2.9) into Eq. (2.7) leads to

$$
\begin{aligned}
|\psi\rangle &= \sum_{\{\sigma_n s_{n-1}\}, \tilde{s}_{n+1}} [USV^\dagger]_{(\sigma_n s_{n-1}), \tilde{s}_{n+1}} |\sigma_n, s_{n-1}\rangle \otimes |\tilde{s}_{n+1}\rangle \\
&= \sum_{\alpha=1}^{k} \sum_{\tilde{s}_{n+1}} \sum_{\sigma_n s_{n-1}} \underbrace{\underbrace{u_{(\sigma_n s_{n-1}), \alpha}}_{\equiv U^{[\sigma_n]}_{s_{n-1}, \alpha}} |\sigma_n, s_{n-1}\rangle}_{\equiv |s_n^{(\alpha)}\rangle} \otimes \underbrace{\lambda_\alpha^{(n)} v^*_{\tilde{s}_{n+1}, \alpha}}_{\equiv X_{\alpha, \tilde{s}_{n+1}}} |\tilde{s}_{n+1}\rangle,
\end{aligned}
\tag{2.10}
$$

which is depicted along the MPS diagrams in then center of Fig. 2.5(a). The reshaped $U_{s_{n-1},\alpha} =: A_n^{\mathrm{new}}$ takes the position of the original A_n. By construction, the LR-orthogonal tensor A_n^{new} generates the new orthonormal effective basis set $|s_n^{(\alpha)}\rangle$ which now also includes site n. The remaining information of the original state $|\psi\rangle$ is split off into the further newly generated matrix $X_{\alpha, \tilde{s}_{n+1}}$ which connects A_n^{new} with the original $|\tilde{s}_{n+1}\rangle$ space to the right. By contracting $X_{\alpha, \tilde{s}_{n+1}}$ onto A_{n+1}, thus generating $\tilde{A}_{n+1}$, the *current site* has been shifted from site n to site $n+1$ with all A-tensors up to and now including site n LR-orthonormalized. All transformations in Eq. (2.10) have been exact, hence overall the state $|\psi\rangle$ remains the same.

Assuming that the X arising from SVD in Eq. (2.10) has an inverse, then the effect of LR-decomposition is equivalent to inserting $X^{-1}X = \mathbf{1}$ in between A_n and A_{n+1} as indicated in the box Fig. 2.5(b). Definitely this leaves the state $|\psi\rangle$ intact. Nevertheless, by associating X^{-1} with A_n and X with A_{n+1}, the current gets shifted from site n to $n+1$.[63] This underlines the internal degree of freedom that comes with MPS, in that every contracted index can be decorated with $X^{-1}X = \mathbf{1}$ with an arbitrary invertible matrix, while contracting X^{-1} and X with the different ends of the bond, respectively.

The LR-orthonormalization step can be repeated iteratively starting from site $n = 1$ all the way to last site $n = N$. For given MPS with open boundary also the very right end is terminated by a vacuum state, say $|\rangle_{\mathrm{R}}$, which again represents a well-defined state space of dimension 1. Therefore the X-matrix generated at the very last iteration N is a 1×1 matrix, *i.e.* a simple number, which reflects the norm of the state. For a normalized state, it is exactly equal to 1. For a non-normalized state, the last X can simply be discarded to obtain a fully normalized state $|\psi\rangle$.

Similar to the LR-orthonormalization prescription, the RL-orthonormalize starts from site N and proceeds in complete analogy to the LR-orthonormalization, yet with reversed role of L and R. This way, with site n the current site, all sites $n' < n$ ($n' > n$) can be written in LR- (RL-) orthonormalized form, respectively, through an iterative prescription as sketched above. Consequently, site n has access to fully orthonormal state spaces for each part of the system, including the local state space σ_n, as depicted in Fig. 2.5(c). The full LR/RL-orthonormalization over the full chain, however, may only be used, for example, during setup of a DMRG calculation. When already fully orthonormal basis sets exist *w.r.t.* site n, the current site can be shifted to sites $n \pm 1$ simply using a single basic LR- or RL-orthonormalization step. As such, this is fully compliant with an iterative scheme.

2.3 Block Entropy

Given an MPS in the local picture of current site n, *cf.* Fig. 2.5(c), the wave function is given by Eq. (2.7),

$$|\psi\rangle = \sum_{\sigma_n, s_{n-1}, \tilde{s}_{n+1}} A^{[\sigma_n]}_{s_{n-1}, \tilde{s}_{n+1}} |\sigma_n\rangle |s_{n-1}\rangle |\tilde{s}_{n+1}\rangle, \tag{2.11}$$

where $|s_{n-1}\rangle$, $|\tilde{s}_{n+1}\rangle$, and $|\sigma_n\rangle$ describe an orthonormal basis for the left ($n' < n$), right ($n' > n$), and the local state space at site n, respectively. Partially tracing out the right side of the system ($\tilde{s}_{n+1}$), this leads to a reduced density matrix whose spectrum is equivalent to the square of the SVD spectrum[55] as in Eq. (2.9). Thus with $\rho_\alpha^{(n)} \equiv [\lambda_\alpha^{(n)}]^2$ the eigenspectrum of $\hat{\rho}^{\text{red}}_{1,\dots,n} \equiv \text{tr}_{\sigma_{n+1}\dots,\sigma_N}\big(|\psi\rangle\langle\psi|\big) = \text{tr}_{\tilde{s}_{n+1}}\big(|\psi\rangle\langle\psi|\big)$, the *block entropy* of the block of sites $n' \in \{1, \dots n\}$ is given by the von-Neumann entropy

$$S_n \equiv -\sum_\alpha \rho_\alpha^{(n)} \ln \rho_\alpha^{(n)} \geq 0. \tag{2.12}$$

For systems with open boundary conditions, the block entropy represents physical information on the MPS state under consideration, in that it does not depend on internal degrees of freedom of the MPS.

Using Schmid decomposition, the wave function in Eq. (2.11) can be rewritten as

$$|\psi\rangle = \sum_\alpha \lambda_\alpha^{(n)} |s_n^{(\alpha)}\rangle |\tilde{s}_{n+1}^{(\alpha)}\rangle. \tag{2.13}$$

Here $|s_n^{(\alpha)}\rangle$ is an orthonormal state space within the fused space $|\sigma_n, s_{n-1}\rangle$, whereas $|\tilde{s}_{n+1}^{(\alpha)}\rangle$ forms some orthonormal superposition of the states $|\tilde{s}_{n+1}\rangle$. Given Eq. (2.13), the SVD spectrum $\lambda_\alpha^{(n)}$ forms the basis of the DMRG's truncation criteria: discarding the smallest eigenvalues of the reduced density matrix, $[\lambda_\alpha^{(n)}]^2 \leq \varepsilon \ll 1$, together with their associated state space, this allows to determine the best possible approximation for the wave function in a controlled fashion.[70] From an MPS point of view, finally, the number D_n of states to be kept at iteration n is qualitatively proportional to an *effective dimension*

$$D_n^* \equiv e^{S_n}, \tag{2.14}$$

with the von-Neumann entropy S_n as in Eq. (2.12), and the proportionality constant D_n/D_n^*, typically of the order $10 \dots 100$, roughly independent of the truncation threshold ε, yet dependant on the specific model Hamiltonian analyzed and the accuracy required.

2.4 Role of symmetries

The MPS framework allows to incorporate preserved symmetries in a transparent and efficient way. In general, symmetries imply that many matrix elements or coefficients

are exactly zero due to selection rules. As a consequence, tensors become sparse. The essential role of abelian symmetries then is, that the non-zero coefficients can be condensed into dense blocks. Therefore while abelian symmetries must deal with the actual full state space dimension still, the efficiency derives from reducing the original problem to a sequential treatment of typically significantly smaller blocks, thus exploiting the sparsity of the problem. The numerically negligible overhead lies in an efficient book keeping of the non-zero blocks.

In the presence of non-abelian symmetries, one further realizes that many of the non-zero matrix elements are actually *dependent* on each other (*e.g.* consider the Wigner-Echart theorem for irreducible operators). In general, this translates into splitting off the corresponding Clebsch Gordan coefficient spaces in terms of a tensor product.[59] As a result, tensors can be strongly *compactified*. Rather than talking about the original full state spaces, the language changes to the significantly reduced multiplet spaces.

A basic introduction to abelian and non-abelian symmetries alike together with a detailed description of a transparent framework for their treatment in general tensor networks in terms of so-called QSpaces is given in Weichselbaum (2012). Essentially, QSpaces can be seen as powerful arbitrary-rank tensor-library that can also deal with compact non-abelian symmetries. It is based on the explicit evaluation of Clebsch-Gordan coefficient spaces from the actual generators of their Lie algebra. As such, it provides a flexible framework in particular *w.r.t.* to the implementation of symmetries, where not all quantities of interest are easily available analytically.

2.4.1 State symmetrization

Consider an arbitrary MPS $|\psi\rangle$ that originally has full A-tensors, *i.e.* makes no reference to any symmetry spaces whatsoever. However, assuming that $|\psi\rangle$ is close to a symmetry eigenstate, this state can be cast into an exact symmetry eigenstate. In practice, this procedure can be useful for testing purposes. The corresponding *state symmetrization* requires to reorganize all indices into state spaces with proper symmetry labels. The actual procedure is straightforward for abelian symmetries. For non-abelian symmetries, however, the procedure requires to partially recombine multiplet spaces with their Clebsch Gordan coefficient spaces into an explicit tensor product. Hence the latter is not as easy to implement while possible in principle.

The actual procedure then is as follows. Firstly, the symmetrization of a given state requires a proper mapping of the local state spaces $|\sigma_n\rangle$ for site n into proper symmetry eigenstates. By contracting this mapping onto every local state space, the local state space can therefore be written in proper symmetry labels. The remaining strategy then is completely analogous to the LR-orthonormalization already discussed with Fig. 2.5. As such, it is an iterative prescription. The starting point may again be taken as the very left end of the MPS, *i.e.* site $n = 1$. The local state space has already been cast into a symmetry basis. The effective basis to its left is the left-vacuums state, which transforms like a scalar for all symmetries. Hence its symmetry labels are trivially also known, *i.e.*
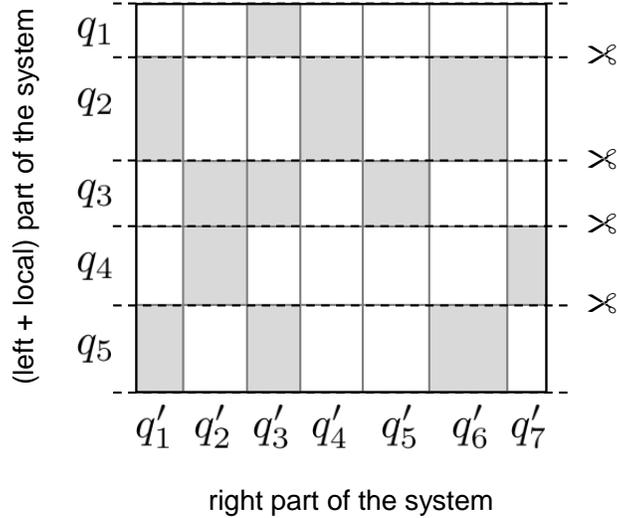
Figure 2.6: Schematic depiction of orthonormalization and possibly truncation of wave function $|\psi\rangle = \sum_{(l,\sigma),r} \psi_{(l,\sigma),r} |l\sigma\rangle |r\rangle$ in the one-site local DMRG picture of left block (l), right block (r), and local state space (σ). Here left (l) and local (σ) state space are assumed to be orthonormal and already written in proper combined symmetry labels q_i. The state space for the right (r) part of the system, however, in principle at this point can be an arbitrary coefficient space. It may already exist in terms of block labels q'_i (with non-zero blocks shaded in gray). his is not required, however, since only full rows are considered anyway. The latter is indicated by the slicing along the lines that separate blocks with different symmetry labels q_i. Therefore in the absence of symmetry labels, for example, the entire state space r may be described by one single block with (irrelevant) block label $q'_1 = 0$.

are all equal to zero. Therefore making the iterative assumption that *w.r.t.* to the current site n the effective basis for the left block of sites $n' < n$ is properly written in terms of their symmetry labels, together with the local Hilbert space of site n, the tensor A_n can be *sliced* into well-defined symmetry spaces, as schematically indicated in Fig. 2.6. There left ($n' < n$) and local (σ) state space are already combined into proper combined symmetry spaces q_i. Since the symmetry labels for the right block are not yet known, the entire state space r can be described by one single block with irrelevant block label $q'_1 = 0$. Subsequent SVD within the combined symmetry sectors q_i allows to extract an orthonormal Schmidt basis for all sites up-to and now including site n, *i.e.* $A_n \to U_n X_n$, as indicated in Fig. 2.5. Here U_n already refers to proper symmetry sectors, such that also the new index connecting U_n and X_n refers to proper symmetry spaces. By contracting X_n onto A_{n+1}, and setting $n \to n + 1$, the iterative prescription can be repeated for site $n + 1$.

Once the end of the chain is reached at site $n = N$, the left index of X_N connecting to A_N will list all symmetry spaces that are contained in the original state $|\psi\rangle$. Having assumed

that $|\psi\rangle$ is close to a symmetry eigenstate, one symmetry sector in X_N will dominate in amplitude. This allows to project the state $|\psi\rangle$ onto the dominant symmetry, by skipping the contributions to X_N from all other symmetries, *i.e.* setting them to zero while also readjusting the global normalization of the state. Overall, now the state $|\psi\rangle$ itself has a well-defined exact symmetry, while also all of its internal state spaces had been reorganized in terms of proper symmetry labels.

2.5 Simple MPS examples

Simple examples for matrix product states with the smallest non-trivial dimension $D = 2$ are[47] the AKLT state[1,56], GHZ state, cluster states, or W-states. These are prototypical states in the analysis of entanglement in the field of quantum information. The simple MPS examples of GHZ and W-state will be motivated in the following. There the local state space refers to qubits, hence this state space is described by $\sigma \in \{0, 1\}$.

2.5.1 GHZ state

The GHZ state for N qubits is given by the state

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|00\ldots0\rangle_N + |11\ldots1\rangle_N). \tag{2.15}$$

With the convention, $|0\rangle \equiv (1, 0)$ and $|1\rangle \equiv (0, 1)$, its MPS representation consists of the uniform, *i.e.* site-independent A-tensors,

$$A_n^{[0]} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_n^{[1]} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad (n = 2, \ldots, N - 1),$$

except for the boundary. For the first qubit (site) one has the A-tensor

$$A_1^{[0]} = \tfrac{1}{2^{1/4}} \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad A_1^{[1]} = \tfrac{1}{2^{1/4}} \begin{pmatrix} 0 & 1 \end{pmatrix},$$

and for the last qubit its transpose, $A_N^{[\sigma]} = (A_1^{[\sigma]})^\dagger$. For a finite contribution for given MPS, clearly *all* of the A-tensors must either be $A^{[0]}$ or $A^{[1]}$, hence this generates the state in Eq. (2.15), indeed.

2.5.2 W-state

The W-state for N qubits is given by

$$|\text{W}\rangle = \tfrac{1}{\sqrt{N}}\Big(|100\ldots0\rangle_N + |010\ldots0\rangle_N + \ldots + |000\ldots1\rangle_N\Big). \tag{2.16}$$

In contrast to the GHZ-state, it is not possible to specify a uniform set of A-tensors which are identical for all sites except for the boundaries.[47] Nevertheless, the *site-dependent* A-tensors for the MPS of the W-state with $D = 2$ can be easily determined.

For any site n within the chain, the W-state can be split into left, local and right part, respectively. All that is needed for either part is its respective contributions to the W-state,

$$|\mathrm{W}\rangle = \tfrac{1}{\sqrt{N}}\left[\sqrt{n_L}|\mathrm{W}\rangle_n^L|0\rangle_n|0\rangle_n^R + |0\rangle_n^L|1\rangle_n|0\rangle_n^R + \sqrt{n_R}|0\rangle_n^L|1\rangle_n|\mathrm{W}\rangle_n^R\right], \qquad (2.17)$$

with $N \equiv n_L + 1 + n_R$. Here $|\mathrm{W}\rangle_n^L$ $[|\mathrm{W}\rangle_n^R]$ refers to the W-state for the qubits to the left $(n' < n)$ [to the right $(n' < n)$] of qubit n, respectively, whereas the states $|0\rangle_n^{[LR]}$ represents the all-zero state for left or right. All states referenced in Eq. (2.17) clearly are already orthonormal, hence a two-dimensional effective state space for both, left and right of site n has been identified, confirming the dimension of the MPS, $D = 2$. Within the effective state spaces for left and right, the A-tensors in the local picture of a specific site n can be readily determined.

For a left-orthonormal state space throughout, the A-tensors can be determined from the iterative character of the W-state,

$$|\mathrm{W}\rangle_n^L = \sqrt{\tfrac{n_L}{n_L+1}}|\mathrm{W}\rangle_{n-1}^L|0\rangle_n + \sqrt{\tfrac{1}{n_L+1}}|0\rangle_{n-1}^L|1\rangle_n. \qquad (2.18)$$

which translates to the LR-orthonormalized A-tensors,

$$A_n^{[0]} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\tfrac{n_L}{n_L+1}} \end{pmatrix}$$
$$A_n^{[1]} = \begin{pmatrix} 0 & \sqrt{\tfrac{1}{n_L+1}} \\ 0 & 0 \end{pmatrix}. \qquad (2.19)$$

The corresponding order of the effective basis for the left block including qubit n is given by $|l_n\rangle \in \{|0\rangle, |\mathrm{W}\rangle\}_n^L$ in exactly this order. For the first iteration, there is no $|\mathrm{W}\rangle^L$ state, hence only the first rows of Eq. (2.19) apply. Together with $n_L = 0$, its A-tensors are

$$A_1^{[0]} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$
$$A_1^{[1]} = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

For the last site, on the other hand, only the $|\mathrm{W}\rangle_N^L$ needs to be selected, hence only the last columns in Eq. (2.19) apply. With $n_L = N - 1$, its A-tensors are

$$A_N^{[0]} = \begin{pmatrix} 0 \\ \sqrt{\tfrac{N-1}{N}} \end{pmatrix}$$
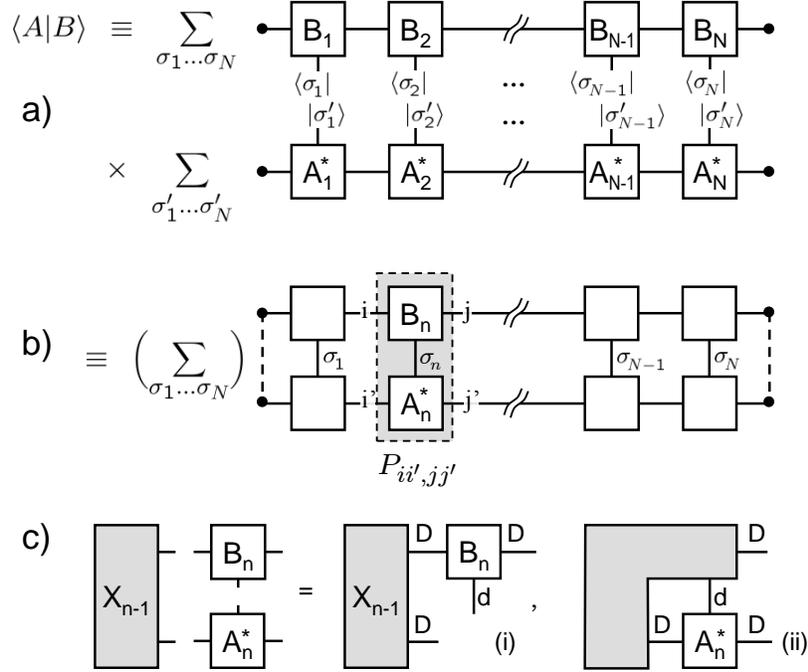$$A_N^{[1]} = \begin{pmatrix} \sqrt{\tfrac{1}{N}} \\ 0 \end{pmatrix}.$$

Figure 2.7: MPS diagram for simple scalar product between states $|A\rangle$ and $|B\rangle$ encoded as MPS (panel a), leading to the plain Fock space contraction in panel (b) (vertical lines). The left- (right-) most horizontal refer to singleton spaces, hence may also be simply contracted for convenience, as indicated by the vertical dashed lines. By convention, connected lines are contracted, *i.e.* summed over, hence the leading sum in panel (b) is implicit, as emphasized by putting it in brackets. Explicit usage of the transfer matrices P_k in evaluating the scalar product would scale like $\mathcal{O}(N \cdot dD^4)$ in cost, but this is not yet optimal.[55] Panel (c) indicates the actual way of calculating a scalar product of MPS states by sequential contraction of (i) B and (ii) A. The cost for each scales like $\mathcal{O}(dD^3)$.

2.6 MPS Algebra

An MPS describes a potentially strongly-correlated quantum many-body-state in an exponentially large Hilbert space, as introduced in Eq. (2.3). While, in principle, the coefficient $\prod_n A^{[\sigma_n]}$ can be calculated for arbitrary but fixed $(\sigma_1, \ldots, \sigma_N)$, this would quickly become exponentially prohibitive for the entire Hilbert space. In practice, however, this is never required. An MPS is stored by its constituting A-tensors, while physical quantities such as expectation values can be calculated efficiently by iterative means. This is based on a generalized matrix-product structure of the underlying (quasi-) one-dimensional physical Hamiltonian.

2.6.1 Simple example: Scalar Product

Consider the very basic example of the scalar product of two states encoded as MPS,

$$
\begin{aligned}
\langle \psi_A | \psi_B \rangle &= \sum_{\sigma_1,\ldots,\sigma_N} \sum_{\sigma'_1,\ldots,\sigma'_N} \Big(\prod_{n=1}^{N} A_k^{[\sigma_n]*} \Big) \Big(\prod_{n=1}^{N} B_k^{[\sigma'_n]} \Big) \langle \sigma_1,\ldots,\sigma_N | \sigma'_1,\ldots,\sigma'_N \rangle \\
&= \prod_{n=1}^{N} \underbrace{\Big(\underbrace{\sum_{\sigma_n}} A_n^{[\sigma_n]*} \otimes B_n^{[\sigma_n]} \Big)}_{\equiv P_n},
\end{aligned}
\tag{2.20}
$$

with its MPS diagram shown in Fig. 2.7. Due to $\langle \sigma_n | \sigma_{n'} \rangle = \delta_{nn'} \delta_{\sigma\sigma'}$, the two MPS in panel (a) directly link vertically, generating the MPS diagram in panel (b), which exactly reflects the last line in Eq. (2.20). Given an open boundary, the left- (right-) most horizontal line connects to the one-dimensional vacuum state. Therefore this singleton index space may also be simply contracted, as indicated by the vertical dashed line. This way, there are no open lines left in the diagram, which allows full contraction to a number, *i.e.* the overlap of the two input states.

Consequently, the scalar product of two MPS has been reduced to roughly N multiplications of the $D^2 \times D^2$ dimensional transfer matrices P_n. While the multiplication of two P_n would scale like $\mathcal{O}(D^6)$, for open boundary conditions and iterative prescription, however, this is essentially reduced to a matrix-vector multiplication in the space of transfer matrices, with overall cost $\mathcal{O}(ND^4)$. The latter can still be further improved upon by *not* combining A_n^* and B_n into a single object P_n, but rather dealing with the original block structure while calculating the scalar product.[56] This is indicated in Fig. 2.7(c). Starting from the left, A_1^* and B_1 can be contracted, generating a matrix X_1. This again sets the stage for an iterative construction. Let X_{n-1} represent the scalar product contracted up to site $n-1$. At the maximum, it is an $D \times D$ matrix (or, equivalently, a D^2 dimensional vector in the space of the transfer matrices P_n). Then including site n requires (i) to contract B_n with cost $\mathcal{O}(dD^3)$, followed by the contraction of A_n^* again with cost $\mathcal{O}(dD^3)$. Hence by working sequentially through the A and B blocks, the cost for calculating the scalar product of two MPS is $\mathcal{O}(dND^3)$.[55]

2.6.2 Operator expectation values

The scalar-product above already showed all the essential ingredients that are also required for the calculation of expectation values. The scalar product calculates the matrix element $\langle A | \mathbf{1} | B \rangle$ with respect to the identity operator,

$$
\mathbf{1} \equiv \mathbf{1}_1 \otimes \mathbf{1}_2 \otimes \ldots \otimes \mathbf{1}_N.
$$

It may be replaced by an arbitrary other tensor product of operators that act locally simultaneously,

$$
\hat{\mathcal{C}} \equiv \hat{c}_1 \otimes \hat{c}_2 \otimes \ldots \otimes \hat{c}_N.
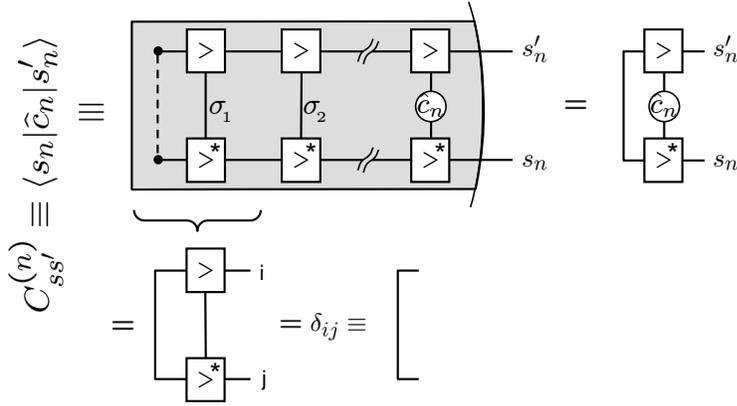\tag{2.21}
$$

Figure 2.8: MPS diagram for obtaining the matrix elements of an operator $\hat{c}_n$ acting locally at the last site n of an effective state space $|s_n\rangle$. The iterations $n' < n$ drop out due to the LR-orthonormalization of the A-tensors and hence can be short-circuited, as explained with Fig. 2.4(a) This leads to the simple actual contraction of the A-tensor of site n with the local operator $\hat{c}_n$ as indicated to the right.

The only extra cost is to also contract the respective operator $\hat{c}_n$ for a given iteration n. This can be done, for example, as a prior step to Fig. 2.7(c-i): $\hat{c}_n$ may be contracted onto the local state space for B_n first, resulting in $\tilde{B}_n$, which then is contracted with X_{n-1}.

Globally, assuming $A = B$, the expectation value then can be simply computed as $\langle\hat{\mathcal{C}}\rangle_A \equiv \langle A|\hat{\mathcal{C}}|A\rangle/\langle A|A\rangle$. Here no specific internal specific orthonormalization of the A-tensors is required. If the operator $\hat{\mathcal{C}}$ is short range or local to some site n, however, then the global contraction to obtain the expectation value can be short-circuited by using LR/RL-orthonormalized state spaces. This is explained next.

2.6.3 Operator representation in effective state space

Matrix elements of local operators can be efficiently obtained given an effective orthonormal MPS basis set $|s\rangle_n$ of the part of the system they act upon. Consider, for example, an operator $\hat{C} = \hat{c}_n$ that acts locally on site n only, as shown in Fig. 2.8. The effective state space $|s\rangle_n$ is assumed to be written as MPS with LR-orthonormal A-tensors. The state space left of site $n' = 1$ describe the empty state, indicated by the bullet. Being a singleton space, it can be simply contracted, for convenience, as indicated by the vertical dashed line. Using Fig. 2.4 then, the contraction of the $A_1^{(*)}$ tensors can be eliminated, *i.e.* shortcut. This implies that now the left legs of the $A_2^{(*)}$ tensors are directly connected, such that again Fig. 2.4 applies. The argument can be repeated up to site $n - 1$, which is equivalent to saying, that by construction, of course, also $|s_{n-1}\rangle$ describes an orthonormal state space. The resulting object on the *r.h.s.* of Fig. 2.4 only involves the last A-tensor from site n together with the operator matrix elements of $\hat{c}_n$ expressed in the local basis $|\sigma_n\rangle$. This simplification is a direct benefit of using orthonormalized state spaces throughout.

Finally, given a one-dimensional physical system with short-range interaction, the matrix elements of the Hamiltonian can be constructed efficiently in an iterative fashion, as demonstrated in Fig. 2.9. Since the effective state space $|s\rangle_n$ is given in terms of an MPS up to iteration n, also only the terms of the Hamiltonian are included that are fully contained
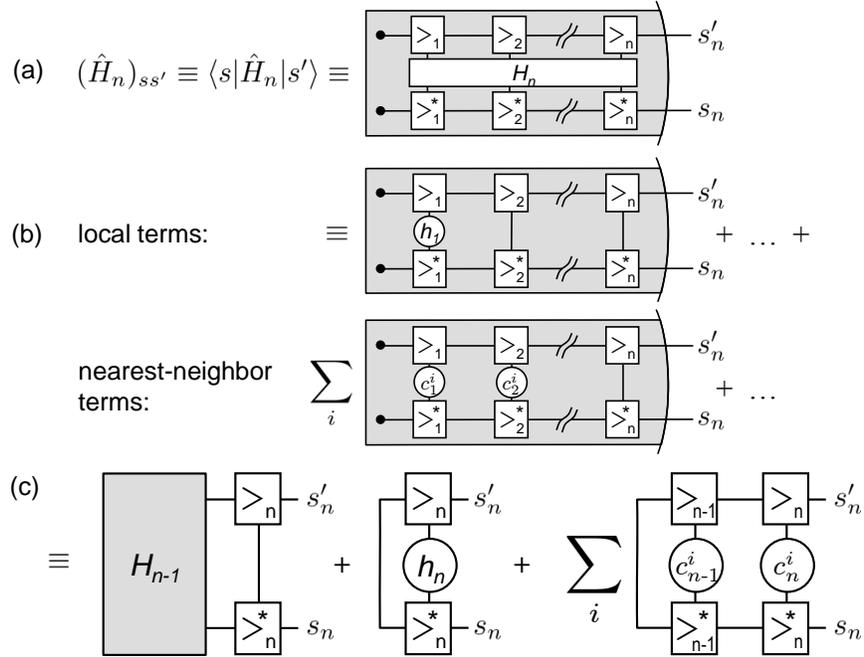
Figure 2.9: MPS diagrams on obtaining the matrix elements of a one-dimensional Hamiltonian with short range interactions only. Panel (a) shows the overall object to be calculated, indicating that H_n acts on all sites involved. Panel (b) depicts the individual local ($\sum_i \hat{h}_i$) and nearest-neighbor terms ($\sum_{n,i} \hat{c}_{n+1}^i \hat{c}_n^i$) constituting the Hamiltonian. Panel (c) shows an efficient iterative scheme, that uses the matrix elements of H_{n-1} of all terms up to and including site $n-1$ obtained from previous iterations.

within the block of sites $n' = 1, \ldots, n$, which is denoted by H_n. By construction, $\hat{H}_n$ can be built iteratively, having $\hat{H}_n = \hat{H}_{n-1} + \hat{h}_{n-1,n}$ where $\hat{h}_{n-1,n} \equiv \hat{h}_n + \sum_i \hat{c}_{n-1}^i \otimes \hat{c}_n^i$ describes the new terms to be added to the Hamiltonian when enlarging the block by one site (local term to site n and nearest-neighbor interaction between sites $n-1$ and n, respectively).[1] As indicated in Fig. 2.9(c), the iterative scheme uses H_{n-1} obtained from previous iterations. Thus to obtain H_n, H_{n-1} is propagated to site n (first term), while the new local term h_n (center term), as well as the nearest-neighbor terms explicitly involving site n interacting with site $n-1$ (right term) need to be added. By convention, hats are reserved for operators acting in the full Hilbert space, while no hats are used for explicit matrix representations of operators.

---

[1]While operators such as the Hamiltonian may also be represented as matrix product operators.[40] Within the NRG, however, this was neither required nor useful, since H_n must only represent those terms of the Hamiltonian that are fully contained *within* the sites $n' \leq n$.

2.7 Connection to DMRG

The density matrix renormalization group (DMRG)[70] is understood as a variational prescription within the space of MPS. Thus given a (quasi-) one-dimensional Hamiltonian $\hat{H}$, DMRG may start from a random and subsequently orthonormalized MPS. It then proceeds with iteratively updating A_n with the goal to reduce the energy,

$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \to \text{MIN} \;\Rightarrow\; \frac{\partial}{\partial A_n^*} \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = 0. \tag{2.22}$$

The latter expression results in an eigenvalue problem within a local Hilbert space of typically huge dimension nevertheless. Being interested in low-energy physics such as the ground state, well-established methods for sparse systems such as Davidson algorithm can be employed. Note that even though DMRG focuses on the ground state from a global perspective, on a local perspective it always has to deal with *many* states that provide an effective description for parts of the system with the focus on low energy physics.

More generally, DMRG can be seen as a powerful extension to exact diagonalization (ED). If the system is small enough, DMRG will give the exact ground state, while for larger systems, it produces the best possible ground state given the numerical resources in a controlled fashion. Consequently, many other concepts know from ED can also be transferred to the DMRG, existing examples being quantum trajectory approach[18] to stochastically sample non-equilibrium process, or Chebyshev expansion[32] for the calculation of spectral properties or even real-time evolution. Note, however, that within MPS even the addition of two states typically already represents an approximation, albeit a well-controlled one. With the resulting algorithms also dealing with states far away from their ground state yet still maintaining some variational character that is not necessarily *w.r.t.* to energy, the acronym DMRG is frequently also replaced in favor of *variational matrix product state* (VMPS).[63]

2.8 Connection to NRG

The numerical renormalization group (NRG,[72] to be introduced and discussed in more detail in Chap. 3) produces an MPS in a single sweep.[51,62,65,66,68] The NRG's crucial underlying assumption derived from logarithmic discretization is energy scale separation (ESS). At a given iteration n then, the environment ($n' > n$) can be safely ignored. The Hamiltonian $\hat{H}_n$ is *exactly* diagonalized in the basis $|s\rangle_{n-1} \otimes |\sigma_n\rangle$. Being interested in the low-energy physics, the high-energy states can be discarded (based on ESS). Hence similar to the DMRG, Wilson's NRG also strongly focuses on the low energy physics. As the iteration in n proceeds, ever smaller excitations from the *overall ground state* are explored. Given that the ground state is part of a continuum, it needs to be interpreted in a renormalization group sense: once the stable low-energy fixed point is reached, the NRG iteration cycle can be stopped. In this sense, the Wilson chain acquires a finite length, the ground state of which is referred to (overall) ground state of the Wilson chain. Note also that

the semi-infinite Wilson chain does have a well-defined ground state energy that converges exponentially with $n \to \infty$.

The cost function of the NRG can thus be written as the ground state of the semi-infinite Wilson chain,

$$\lim_{N \to \infty} \langle s_N | \hat{H} | s_N \rangle \to \text{MIN}. \tag{2.23}$$

Despite that from a global perspective this represents a single state, namely the ground state for the semi-infinite system, at any *intermediate* finite iteration n always *many* states are available. Note that this situation is not that much unlike DMRG at all. Nevertheless, through the extra ingredient of energy scale separation there is much more to the NRG still. Most importantly, the NRG does represent a true renormalization group analysis in contrast to the DMRG. Therefore by construction, NRG provides a powerful setup for fixed point analysis at all energy scales.[12,36,72] As a consequence, this finally also allows an efficient systematic treatment of dynamical quantities at arbitrary finite temperature.[12,68]

2.9 MPS and sequential generation of correlated qubit states

This section was born out of the project Saberi et al. (2009) dealing with the numerical optimization of sequential generation of correlated qubit states using an auxiliary physical state space.[37] While this direction will be less important for the rest of this thesis, it nevertheless provides an interesting alternative point of view on the effective state spaces generated within an MPS: akin to the well-known AKLT construction,[1,56] it gives the effective state spaces a clear physical interpretation. Arguments along these lines can be generalized to two-dimensional networks (PEPS),[6,63] and thus be used to argue in favor of symmetries in these tensor networks. Here in the one-dimensional setting, the basic idea is that an initially decoupled auxiliary state space acts sequentially with a linear sequence of local state space, such that at the end of the protocol the auxiliary state is again guaranteed to decouple, yet served to prepare a set of qubits in an arbitrary yet fixed MPS of dimension D.[37] In particular, this section provides an explicit construction of the unitaries required to generate given MPS. This construction is always possible, yet is not unique, as will be seen below.

Consider the setup shown in Fig. 2.10(a). The system starts out in an initial product state (MPS with $D_0 = 1$), indicated by the bullets lined up at the top of the panel. The auxiliary physical system with state space dimension D is initialized in state $|a_0\rangle$, shown at the very left. Without restricting the case, the dimension of the local state space is assumed to be $d = 2$ with $\sigma \in \{0, 1\}$, as for qubits or spin-half systems. The initial states are chosen as $|\sigma_i^!\rangle := |0\rangle \equiv (1, 0)$, and also $|a_0\rangle := (1, 0, \ldots, 0)_D$ a D-dimensional vector. This setting is general in that one is always free to choose a different state by inserting a $V^\dagger V = \mathbf{1}$ with arbitrary unitary V in between any of the initial states and its connecting U_n. Applying V then onto the input state, allows to generate an arbitrary input state,
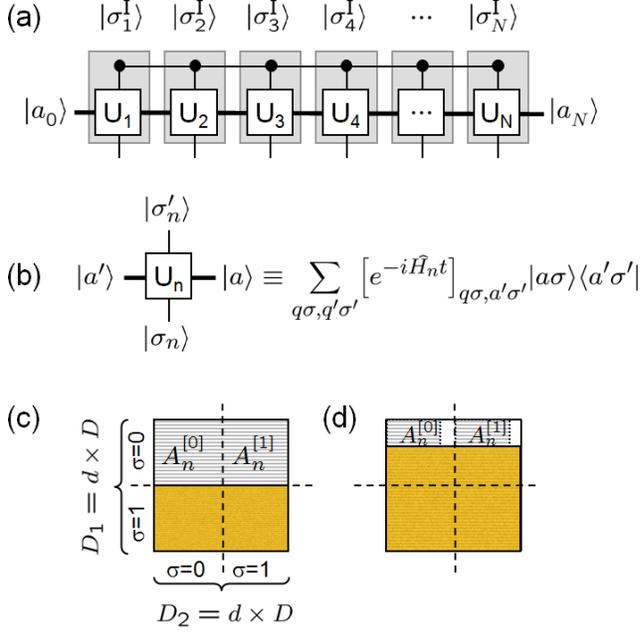
Figure 2.10: Sequential generation of MPS through interaction with auxiliary state space (*e.g.* to creating entangled qubit systems via *sequential quantum factory*[37]). Panel (a) The initial state of the system is a product state $|\psi\rangle^{\mathrm{I}} \equiv \bigotimes_{n=1}^{N} |\sigma_n^{\mathrm{I}}\rangle$ [points at the top, indicating initial state in terms of an MPS of dimension $D_0 = 1$]. The auxiliary system starts in the left state $|a_0\rangle$. It interacts sequentially with each site n sequentially from $n = 1, \ldots, N$ as described by unitaries U_n. After the last site the protocol is such that the auxiliary system again decouples, leaving the actual system in the desired MPS state. Panel (b) shows the interaction of the auxiliary system with site n through a tailored Hamiltonian H_n. Panel (c) explicit construction of the unitaries U_n from the target MPS described the A_n tensors assuming the dimension d of a site to be $d = 2$, for simplicity.

while the unitary $V^\dagger$ can be contracted onto U_n, simply yielding an altered unitary $\tilde{U}_n$. The sequential operation of the auxiliary state space is described by unitary matrices U_n. These are imagined to be the result of letting the auxiliary system interact with site n through some specific Hamiltonian $\hat{H}_n$ over some time t, leading to $U_n = e^{-i\hat{H}_n t}$ as depicted in Fig. 2.10(b). Depending on the specific structure of the MPS, the Hamiltonian may be constrained to a certain set of operations available to the experiment.[52]

Now the aim is to use the physical auxiliary system of dimension D to prepare the actual system in some desired MPS given in terms of the site-specific A_n-tensors of the MPS dimension equal to the dimension D of the auxiliary state space. The explicit construction of the unitaries U_n from the A_n is not entirely unique, as will be seen, thus leaves some room in constraining the Hamiltonian to physically accessible ones.

The construction of the unitaries U_n from the A_n proceeds as follows. Without changing the physical state, the target MPS can be brought into a RL(!)-orthonormal form. The A_n for arbitrary n can thus be rewritten as a $D_1^{(n)} \times D_2^{(n)}$ matrix $A_{s_{n-1}, \{s_n, \sigma_n\}}$, which has orthonormal rows by definition of RL-orthonormalization. Therefore it also must hold $D_1^{(n)} \leq D_2^{(n)}$ to guarantee linearly independent orthonormal vectors. The space of a unitary U_n, on the other hand, is given by $dD \times dD$, *i.e.* it can be represented as a $d \times d$ array of $D \times D$ blocks, as indicated in Fig. 2.10(c). In particular, it holds $dD \geq \{D_1^{(n)}, D_2^{(n)}\}$ for all n. Therefore considering a $dD \times dD$ unitary space U, the d $A^{[\sigma_n]}$-matrices can be inserted in the first row of $D \times D$ blocks. If the dimension of A-tensors is given by their maximal dimension $D \times D$, as is the case in the center of the MPS, then these matrices completely

fill the first row of $D \times D$ blocks in U [Fig. 2.10(c)]. If the $A^{[\sigma_n]}$ matrices have smaller dimensions, they leave some space as shown in Fig. 2.10(d). These rows of U are padded with plain zeros which preserves orthonormality of the rows (indicated by the white space to the right of the $A^{[\sigma]}$), while the remaining space to the bottom needs to be completed into a full unitary U (this is always possible, even though not unique).

The unitaries U_n thus constructed from the A_n do generate the underlying MPS. In order to see this, remember that the local state space was assumed to be initialized in the states $|\sigma_i\rangle = (1,0)$. Its action onto the unitary as rank-4 object as shown in Fig. 2.10(b), effectively projects out *all* except for the first row of blocks in the block-matrix representation of Fig. 2.10(c) or (d). Hence it selects the first row of blocks, *i.e.* the one containing the A-tensors. Furthermore, the auxiliary system was assumed to be in the initial state $|a_0\rangle := (1, 0, \ldots, 0)_D$. The auxiliary system acts *within* the $D \times D$ blocks in Fig. 2.10(c) or (d). With the $A^{[\sigma_1]}$ of the first site being row vectors, $|a_0\rangle$ thus exactly selects the first row in U_1 containing the $A^{[\sigma_1]}$, thus discards the (arbitrarily) completed space for U_1. Having padded the rows of A_n in Fig. 2.10(d) with zeros, the same argument also applies for the next site. That is, the auxiliary space emerging from U_1 again also selects the rows in U_2 that contain $A^{[\sigma_2]}$. Repeating the argument all the way the last site N, the MPS has, in fact, been fully encoded in the unitaries U_n. ∎

All completed spaces in Fig. 2.10(c) or (d), indicated by the orange (dark) shaded areas, are fully projected out, thus they are never referenced. The physical state generated this way can be written as usual MPS, by contracting the input states $|\sigma_0^I\rangle_n$ onto the U_n tensors, thus generating a standard rank-3 tensor, as indicated by the gray boxes in Fig. 2.10(a). By encoding the underlying target MPS, the Schmidt rank after the last iteration $n = N$ is 1, indicating that, by construction, the auxiliary system is again fully decoupled, indeed, as desired.

Chapter 3

Numerical renormalization group and quantum impurity systems

The numerical renormalization group deals with quantum impurity models by construction.[72] In the following, therefore a brief introduction is given to the underlying concepts together with the definition of the single impurity Anderson model for later reference.

The generic quantum impurity system (QIS) consists of a small quantum system (the *quantum impurity*) that is coupled to a non-interacting macroscopic reservoir, *e.g.* a Fermi sea, at the location $\vec{r} = 0$ of the impurity. Let the impurity be described by the general Hamiltonian $\hat{H}_{\mathrm{imp}}$ and the reservoir by the Hamiltonian

$$\hat{H}_{\mathrm{bath}} = \sum_{k\mu} \varepsilon_{k\mu} \hat{c}^{\dagger}_{k\mu} \hat{c}_{k\mu}. \tag{3.1}$$

Here k may be seen as momentum in s-wave approximation or simply as energy label, while μ describes different flavors of particles, such as spin or channel. Therefore $\hat{c}^{\dagger}_{k\mu}$ creates a particle with momentum or energy k and flavor μ at energy $\varepsilon_{k\mu}$. More generally, $\hat{c}^{\dagger}_{k\mu}$ describes the single-particle eigenstates of the non-interacting bath Hamiltonian. Typically, the bath for different flavors look identical from the point of view of the impurity, hence $\varepsilon_{k\mu} \equiv \varepsilon_k$ can be taken independent of μ. The state of the bath at the location of the impurity is given by $\hat{f}_{0\mu} \equiv \frac{1}{\sqrt{\mathcal{N}}} \sum_k \hat{c}_{k\mu}$, with proper normalization $\mathcal{N}$, *e.g.* determined via canonical anticommutator relations $\{\hat{f}_{0\mu}, \hat{f}^{\dagger}_{0\mu}\} = 1$. With this, the combined quantum impurity system is described by the generic Hamiltonian

$$\hat{H}_{\mathrm{QIS}} = \underbrace{\hat{H}_{\mathrm{imp}} + \hat{H}_{\mathrm{cpl}}(\{\hat{f}_{0\mu}\})}_{\equiv H_0} + \hat{H}_{\mathrm{bath}}, \tag{3.2}$$

where the coupling term $\hat{H}_{\mathrm{cpl}}(\{\hat{f}_{0\mu}\})$ can act arbitrarily within the impurity system, while it interacts with bath μ only through $\hat{f}^{(\dagger)}_{0\mu}$, *i.e.* its degree of freedom at the location of the impurity. The Hilbert space of the typically interacting *local Hamiltonian* $\hat{H}_0$ in Eq. (3.1) is considered small enough so it can be easily solved exactly numerically. Thus while $\hat{H}_0$ and $\hat{H}_{\mathrm{bath}}$ are easy to solve individually, their combination can give rise to highly nontrivial strongly correlated quantum-many-body effects such as Kondo physics.

The description of an interacting many-body Hamiltonian requires the treatment of the full exponentially large Hilbert space. This is typically dealt with using some kind of
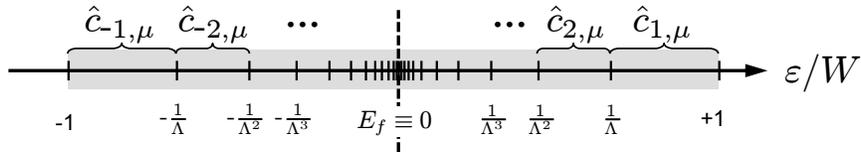
Figure 3.1: Logarithmic discretization of a conduction band of half-bandwidth W resulting in the new coarse-grained states $\hat{c}_{m\mu}$ with m an integer.

renormalization group ansatz. In the case of quantum impurity model, this is the numerical renormalization group (NRG).[12,36,72] The NRG is a two-step process. (i) The continuum of states in the bath is coarse grained relative to the Fermi energy using the discretization parameter $\Lambda > 1$, such that with W the half-bandwidth of a Fermi sea, this defines a set of intervals $\pm W[\Lambda^{-(m-z+1)/2}, \Lambda^{-(m-z)/2}]$, each of which will be described by a *single* fermionic degree of freedom, as depicted schematically in Fig. 3.1. Here m is a positive integer, with the additional constant $z \in [0, 1[$ introducing an arbitrary shift.[44,76] The great advantage of using this kind of discretization is that it allows to zoom in to arbitrarily small energy scales close to the Fermi energy. By construction, this allows to resolve dynamically generated possibly exponentially small energy scales such as the Kondo temperature. (ii) With $\hat{f}_{0\mu}$ being a linear superposition of states in the original diagonal bath states $\hat{c}_{k\mu}$, for each flavor μ the bath can be unitarily rotated into a new orthonormal basis with $\hat{f}_{0\mu}$ representing the seed state. In particular, given $\hat{f}_{0\mu}$, the non-interacting bath can always be tridiagonalized, which thus allows to map the coarse grained bath exactly onto a semi-infinite chain with the first site described by $\hat{f}_{0\mu}$. This one-dimensional linear setup is called the Wilson chain,[72]

$$\hat{H}_{\text{QIS}} \simeq \lim_{n \to \infty} \hat{H}_n \tag{3.3a}$$

with

$$\hat{H}_n \equiv \underbrace{\hat{H}_{\text{imp}} + \hat{H}_{\text{cpl}}(\{\hat{f}_{0\mu}\})}_{(=H_0)} + \sum_{\mu} \sum_{n'=0}^{n-1} t_{n'} \left(\hat{f}_{n'\mu}^{\dagger} \hat{f}_{n'+1,\mu} + \text{H.c.} \right), \tag{3.3b}$$

where the upper limit in the last sum was chosen such, that $\left(\hat{H}_n \right)_{n=0}$, indeed, corresponds to the local Hamiltonian $\hat{H}_0$, *cf.* Eq. (3.2). Within the setting of the Wilson chain, the impurity (first site in the chain) couples to the first site of the bath $f_{0\mu}$ (second site in the chain) only, while the remainder of the chain consists of a plain tight-binding chain between bath sites with hopping amplitudes t_n. For larger n, it quickly holds

$$\omega_n \equiv \lim_{n \gg 1} t_n = \frac{\Lambda^{z-1}(\Lambda-1)}{\log \Lambda} W \Lambda^{-\frac{n}{2}}, \tag{3.3c}$$

where ω_n describes the (smallest) energy scale of a Wilson chain including all sites up to and including site $n+1$ for arbitrary Λ and z-shift. For simplicity, a flat hybridization function

is assumed throughout, *i.e.* $\Gamma(E) = \Gamma\vartheta(W - |E|)$, with the discretization following the prescription of Žitko and Pruschke (2009).

Given the exponentially decaying hopping amplitudes t_n in Eq. (3.3c), this justifies the powerful assumption of energy scale separation along the Wilson chain for large enough Λ, typically $\Lambda \gtrsim 1.7$. In particular, this justifies the iterative diagonalization of the Hamiltonian in Eq. (3.3b)

$$\hat{H}_{n+1} = \hat{H}_n + \sum_\mu t_n\big(\hat{f}^\dagger_{n\mu}\hat{f}_{n+1,\mu} + \text{H.c.}\big), \tag{3.4}$$

which also goes along with an extremely systematic analysis of the full QIS.[72] By starting with the local Hamiltonian H_0 and adding one Wilson site after another, for large enough Λ energy scales separate[72] in the sense that energy levels remain in closer vicinity of the initial energy level when another site is added due the nearest-neighbor hopping structure. The energy scale separation thus leads to the usual NRG truncation criteria prescription of discarding states with highest energies, while keeping the low-lying states at every iteration. Even though NRG would be exact in the limit $\Lambda \to 1$, this limit cannot be taken in practice, however, since truncation by a fixed energy would lead to infinitely large state spaces that must be kept from one iteration to the next. Since this cannot be done, energy scale separation would be violated.

Energy scale separation (ESS) is an essential ingredient to the NRG. It can be directly linked to the NRG's only approximation, namely the coarse graining in terms of Λ. As a consequence, ESS directly leads to the following three rigid drawbacks of NRG.[12] (1) With a $\Lambda \gtrsim 1.7$ required for energy scale separation, dynamical quantities such as correlation function also have *finite resolution* at finite frequency. That is, the discrete data from an NRG calculation needs to be broadened substantially to obtain smooth curves[12] [see also supplementary material to Weichselbaum and von Delft (2007)]. Averaging over uniformly distributed z-shifts[44,76] greatly improves resolution, yet clearly cannot provide the equivalent of the continuous limit $\Lambda \to 1$ while using $\Lambda \gtrsim 1.7$. (2) Being interested in thermodynamic quantities at finite temperature T, in general, this implies that the thermal density matrix has its dominant contributions from energy shells corresponding to T.[68] These shells, however, only have the finite energy resolution corresponding to these shells. Consequently, sharp features in dynamical quantities for $|\omega| \ll T$, if existent, typically *cannot* be resolved. Nevertheless, FDM-NRG[68] (see later discussion), gives reliable resolution for $\omega \geq \alpha T$ with $\alpha \lesssim 1$. (3) Due to the separation of energy scales, all terms in the Wilson chain setup of the Hamiltonian that contribute to the same energy scale at a given NRG iteration *must* be included at the same time. This implies that multichannel models becomes exponentially expensive in the number of channels.

These limitations of the NRG can be relaxed, if the rigid requirement of energy scale separation is loosened. While this no longer justifies the one-way iterative diagonalization of the Wilson chain, the ground state properties of the Wilson chain can nevertheless be analyzed from a variational point of view.[21,49,51,62,65,66] Through the common algebraic structure of matrix product states, this naturally leads to a DMRG based analysis of the Wilson chain.

3.1 Single impurity Anderson model

A prototypical quantum impurity model applicable to the NRG is the single impurity Anderson model (SIAM).[3,4,23,24] It consists of a single interacting fermionic level (d-level), *i.e.* the impurity,

$$\hat{H}_{\mathrm{imp}} = \sum_{\sigma} \varepsilon_{d\sigma} \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}. \tag{3.5a}$$

with level-position $\varepsilon_{d\sigma}$ and onsite interaction U. This impurity is coupled through the hybridization

$$\hat{H}_{\mathrm{cpl}} = \sum_{k\sigma} \left(V_{k\sigma} \hat{d}^{\dagger}_{\sigma} \hat{c}_{k\sigma} + \mathrm{H.c.} \right) = \sum_{\sigma} \left(\hat{d}^{\dagger}_{\sigma} \cdot \underbrace{\sum_{k} V_{k\sigma} \hat{c}_{k\sigma}}_{\equiv \tilde{f}_{0\sigma}} + \mathrm{H.c.} \right) \tag{3.5b}$$

to a spinful non-interacting fermi sea,

$$\hat{H}_{\mathrm{bath}} = \sum_{k\sigma} \varepsilon_{k\sigma} \hat{n}_{k\sigma}, \tag{3.5c}$$

thus referred to as a one-channel system. Here $\hat{d}^{\dagger}_{\sigma}$ ($\hat{c}^{\dagger}_{k\sigma}$) creates an electron with spin σ at the d-level (in the bath with energy k), respectively. Moreover, $\hat{n}_{d\sigma} \equiv \hat{d}^{\dagger}_{\sigma} \hat{d}_{\sigma}$, and $\hat{n}_{k\sigma} \equiv \hat{c}^{\dagger}_{k\sigma} \hat{c}_{k\sigma}$. The model has three physical parameter regimes that can be tuned through temperature: (i) the free orbital regime (FO) at large energies allows all states at the impurity from empty to doubly occupied. (ii) the local moment regime (LM) at intermediate energies with a single electron at the impurity and the the empty and double occupied state only accessible through virtual transitions. (iii) the low-energy strong coupling (SC) fixed-point where the the local moment is fully screened into a quantum-many-body singlet with the electrons in the bath. Through Schrieffer-Wolff transformation[57] the FO regime can be projected out, leading to an effective $J\hat{S}_{\mathrm{imp}} \cdot \hat{S}_{\mathrm{bath}}$ spin interaction with positive J between the spin of the impurity and the spin of all electrons in the bath, resulting in the so-called Kondo Hamiltonian.[33]

Given Eqs. (3.5b) and (3.5c), form the point of view of the impurity, the effects of the bath are fully captured by the bath Greens function $\Delta_{(\sigma)}(\omega) \equiv \langle \tilde{f}_{0\sigma} \| \tilde{f}^{\dagger}_{0\sigma} \rangle$,

$$\Delta_{(\sigma)}(\omega) = \sum_{k} \frac{|V_{k\sigma}|^2}{\omega^+ - \varepsilon_{k\sigma}}, \tag{3.6}$$

with $\omega^+ \equiv \omega + i0^+$ and 0^+ a positive infinitesimal, and having used the *non*-normalized fermionic state $\tilde{f}_{0\sigma}$ in Eq. (3.5b). Due to analytical, *i.e.* retarded, structure of $\Delta_{(\sigma)}(\omega)$, it is sufficient to consider its imaginary part only, since the real part can simply be obtained through Kramers-Kronig relations. The imaginary part then defines the hybridization

function $\Gamma_{(\sigma)}(\omega)$,

$$
\begin{aligned}
\Gamma_{(\sigma)}(\omega) &\equiv -\mathrm{Im}\Delta_{(\sigma)}(\omega) = \sum_k \pi |V_{k\sigma}|^2 \delta(\omega - \varepsilon_{k\sigma}) \\
&\equiv \pi \rho_{(\sigma)}(\varepsilon) V_{(\sigma)}^2(\varepsilon),
\end{aligned}
\tag{3.7}
$$

where the last expression represents the continuum limit, with $\rho(\varepsilon)$ the density of states of the bath. The hybridization function therefore corresponds to a strictly positive spectral density.

Since the coupling in Eq. (3.5b) preserves spin, the hybridization function $\Delta_\sigma(\omega)$ can be defined for each spin individually. Moreover, using a spin symmetric setup, $\Delta(\omega)$ becomes independent of spin, indicated by the brackets around σ in Eqs. (3.6) and (3.7). Since typically the bandwidth is much larger as compared to the relevant physical scales of the model Hamiltonian, only the properties of the bath near the the fermi energy ε_f are relevant. Using a particle-hole symmetric setup for the bath, then the typical choice taken is the box distribution $\Gamma_{(\sigma)}(\omega) = \vartheta(|\omega| - 1)\Gamma_{(\sigma)}$ with constant $\Gamma_{(\sigma)}$, chemical potential $\varepsilon_f = 0$, and bandwidth 1. Consequently, all physical parameters are considered in units of bandwidth.

The state $\tilde{f}_{0\sigma}$ of the bath finally can be written as a properly normalized state $\hat{f}_{0\sigma}$,

$$
\hat{f}_{0\sigma} \equiv \frac{\tilde{f}_{0\sigma}}{\sqrt{\int \frac{d\omega}{\pi}\Gamma_{(\sigma)}(\omega)}},
$$

derived from the fermionic anticommutator relation $\{\hat{f}_{0\sigma}, \hat{f}_{0\sigma}^\dagger\} = 1$, where the inverse factor of π originates in the definition of $\Gamma(\omega)$ in Eq. (3.7). Using the simple box distribution above with total bandwidth of 2., it follows, $\tilde{f}_{0\sigma} = \sqrt{\frac{2\Gamma_{(\sigma)}}{\pi}}\hat{f}_{0\sigma}$.

3.2 NRG and area laws

By construction, the NRG generates a matrix product state.[51,62,65,66] Moreover, it can be argued [*cf.* Sec. 2.8] that the NRG prescription based on energy scales is quasi-variational. By now the success of variational MPS, *i.e.* DMRG, to ground state calculations of quasi-one-dimensional systems is understood to be firmly rooted in the area law for the entanglement entropy S_A.[58,63,74] The latter is defined as the entanglement of some contiguous region A with the rest of the system. In this sense, MPS is ideally suited to capture ground state properties efficiently for quasi-one-dimensional systems.

In comparison, the NRG references all energy scales through its iterative diagonalization scheme, but nevertheless zooms in towards the low energy scales (*"ground state properties"*) of the full semi-infinite Wilson chain. Therefore given a Wilson chain of sufficient length N, without restricting the case, one may simply consider the fully mixed density matrix built from the ground state space $|0\rangle_N$. This then allows to analyze the entanglement entropy S_n of the states $|s\rangle_n$, *i.e.* the block of sites $n' < n$, with respect to its environment $|e\rangle_n$.
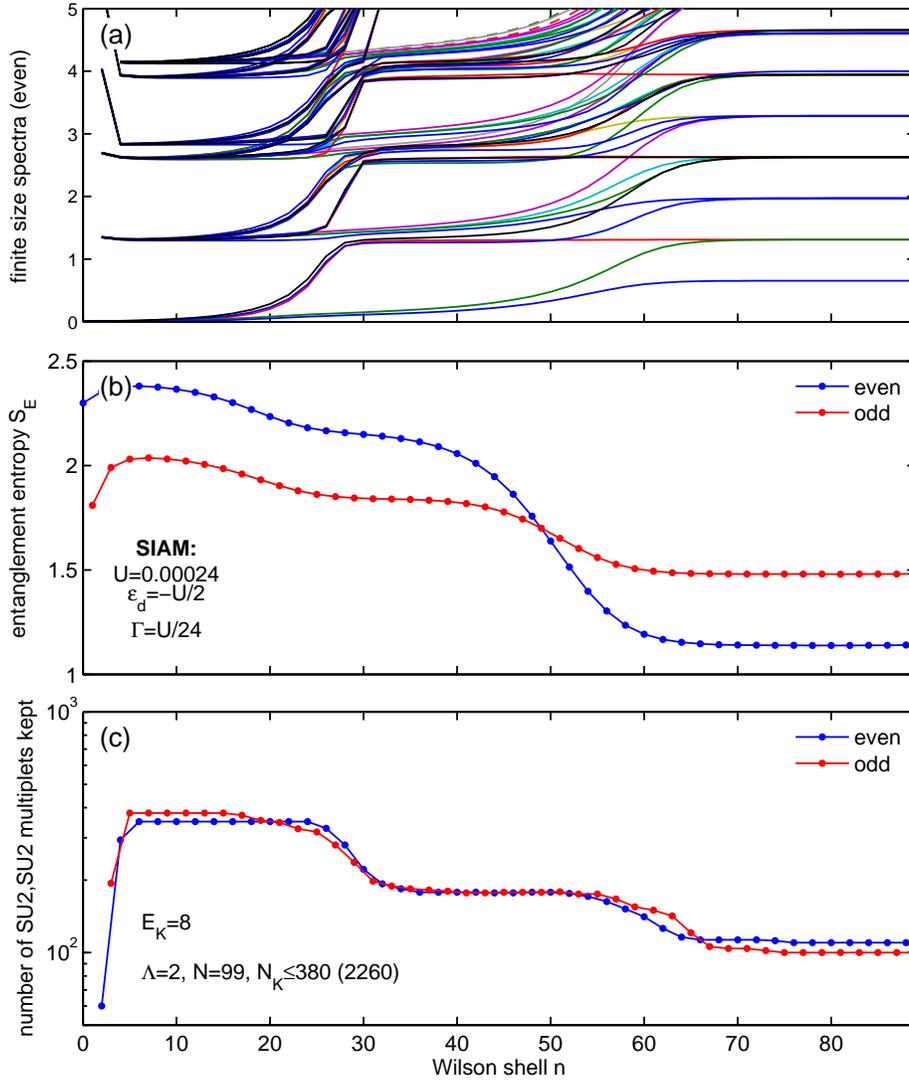
Figure 3.2: NRG and area law – analysis for the symmetric single impurity Anderson model for the parameters as shown in panels (b) and (c) [*cf.* Eq. (3.5); all energies in units of bandwidth]. Panel (a) shows the standard NRG energy flow diagram for even iterations where the different colors indicate different symmetry sectors. Panel (b) shows the entanglement entropy S_n of the Wilson chain up to and including site n with the rest of the chain *w.r.t.* the overall ground state. Due to intrinisic even-odd alternations, even and odd iterations n are plotted separately. Panel (c) shows the actual number of multiplets kept from one iteration to the next, using a dynamical truncation criteria *w.r.t.* a given threshold energy E_K. The calculation used $SU(2)_{\mathrm{spin}} \otimes SU(2)_{\mathrm{charge}}$ symmetry, hence the *actual* number of kept states is by about an order of magnitude higher (*e.g.* as indicated with the maximum number of states kept, N_K in panel (c): the value in brackets gives the corresponding number of states).

The interesting consequence in terms of area law is that one expects the (close to) lowest entanglement entropy S_n for the stable low-energy fixed point, while one expects S_n to increase for higher energies, *i.e.* with decreasing Wilson shell index n. This is nicely confirmed, for example, in a test calculation for the single impurity Anderson model (SIAM), as shown in Fig. 3.2.

Figure. 3.2(a) shows the standard NRG energy flow diagram for even iterations, which clearly outlines the physical regimes of free orbital (FO, $n \lesssim 25$), local moment (LM, $25 \lesssim n \lesssim 60$), and strong coupling (SC, $n \gtrsim 60$) regime. In particular, in order to have a sufficiently wide FO regime, a very small onsite interaction U was chosen. Panel (b) shows the entanglement entropy. Up to the very beginning or the very end of the chain (the latter is not shown), this shows a smooth monotonously decaying behavior *vs.* energy scale. In particular, consistent with the area law for lowest-energy states, the entanglement is smallest once the stable low-energy fixed point is reached. Having chosen a dynamical (quasi-variational)[66] truncation scheme *w.r.t.* to energy, the qualitative behavior of the entanglement entropy is also reflected in the number of states that one has to keep for some fixed overall accuracy, as shown in Fig. 3.2(c). Clearly, up to the very few first shells prior to truncation, the largest number of states must be kept at early iterations. While this is a hand-waving argument, this nevertheless confirms the empirical fact, that the first few Wilson shells with truncation are usually the most important, *i.e.* most expensive ones. Therefore for good overall accuracy all the way down to the low energy sector, one must allow for a sufficiently large number of states to be kept at early iterations.

Chapter 4

MPS diagrammatics for the numerical renormalization group

The complete basis sets within the NRG represent an excellent setting to describe in detail all relevant steps when computing thermodynamic quantities. Using FDM-NRG instead of the more traditional patching schemes, all these calculations become transparent well-defined procedures, in the sense that all steps are *completely fixed* without any requirement for further algorithmic fine-tuning. The underlying MPS algebra nevertheless leads to contractions of larger tensor networks, which quickly can become complex mathematical expression if spelled out explicitly. Alternatively, it has proven much more convenient to represent the relevant tensor networks graphically,[56] which in the following is dubbed MPS diagrammatics. This chapter then describes in detail a range of tensor networks that need to be dealt with when calculating thermodynamical quantities within the NRG. For this, however, a brief introduction to complete basis sets is given, together with their relevant general properties.

Within the NRG, a complete full-many body basis can be constructed from the state space of the iteratively computed NRG eigenstates $\hat{H}_n |s\rangle_n = E_s^n |s\rangle_n$.[2] With the NRG stopped at some final length N of the Wilson chain, the NRG eigenstates *w.r.t.* to site $n < N$ can be complemented by the complete state space of the rest of the chain, $|e\rangle_n$, living on sites $n+1, \ldots, N$. The latter space will be referred to as the *environment*, which due to energy scale separation will only weakly effect the states $|s\rangle_n$. Therefore given the combined states,

$$|se\rangle_n \equiv |s\rangle_n \otimes |e\rangle_n, \tag{4.1}$$

spanning the full Wilson chain, within the validity of energy scale separation,[2]

$$\hat{H}_N |se\rangle_n \simeq E_s^n |se\rangle_n, \tag{4.2a}$$

the NRG eigenstates at iteration n are, to a good approximation, also eigenstates of the full Wilson chain. This holds for a reasonably large discretization parameter $\Lambda \gtrsim 1.7$.[51]

The iteratively *discarded* state space allows to build a complete many-body *eigen*basis of the full Hamiltonian,[2]

$$\mathbf{1}^{(d_0 d^N)} = \sum_{se,n} |se\rangle_n^{\mathrm{DD}} {}_n\langle se|, \tag{4.2b}$$
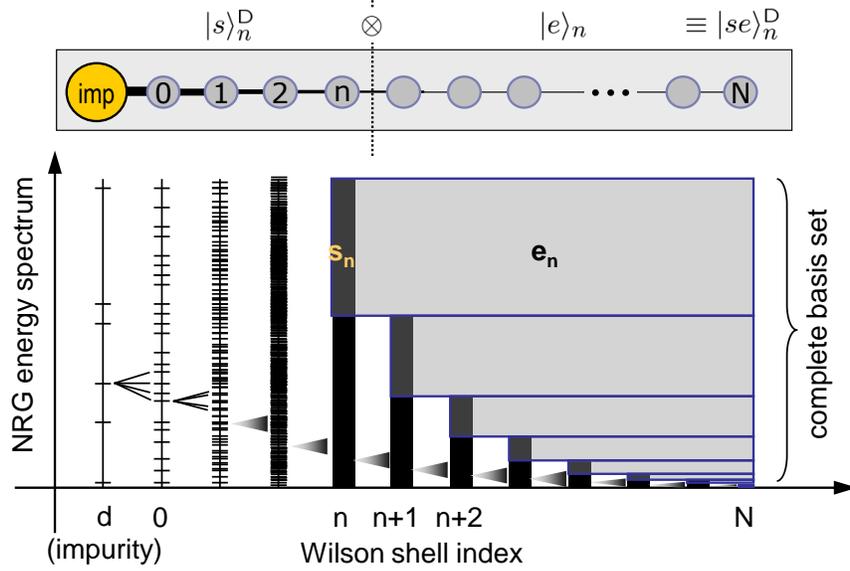
Figure 4.1: Iterative construction of complete basis set within the NRG by collecting the *discarded* state spaces $|s\rangle_n^D$ from all iterations $n \leq N$ (black space a the left of the gray blocks), which is complimented by the environment $|e\rangle_n$ for the rest of the system starting from site $n+1$ up to the overall chain length N considered (gray blocks). In a hand-waving picture, by adding a site to the chain, this site introduces a new lowest energy scale to the system, with the effect that existing levels become split within a narrow energy window (indicated by the spread of levels from one iteration to the next). The impurity, and also the first few sites can be considered exactly with a manageable total number of states in its Hilbert space still. Yet as the state space grows exponentially, truncation quickly sets in. The discarded state spaces then when collected form a complete basis. At the last iteration, where NRG is stopped, all states are considered discarded, for completeness.

where $d_0 d^N$ refers to the full Hilbert space dimension of the Hamiltonian H_N. Here d refers the state space dimension of a single Wilson site, while d_0 refers to the state space dimension of H_0, which also includes the impurity. It is further assumed, that the local Hamiltonian H_0 is never truncated, *i.e.* truncation sets in for some $n = n_0 > 0$. Finally, at the last iteration $n = N$, all states are considered discarded for consistency.

The completeness of the state space in Eq. (4.2b) can be easily motivated by realizing that at every NRG truncation step, by construction, the discarded space (eigenstates at iteration n with largest energies) is orthogonal to the kept space (eigenstates with lowest eigenenergies). The subsequent refinement of the kept space at later iterations will not change the fact, that the discarded states at iteration n remain orthogonal to the state space generated at later iterations. This iterative reduction of Hilbert space while building up a complimentary complete orthogonal state space is schematically depicted in Fig. 4.1.

4.1   Identities

Given the complete basis in Eq. (4.2), the following identities hold,[2]

$$\underbrace{\sum_{se} |se\rangle_{n}^{KK}{}_{n}\langle se|}_{\equiv \hat{P}_n^K} = \sum_{n'>n} \underbrace{\sum_{se} |se\rangle_{n'n'}^{DD}\langle se|}_{\equiv \hat{P}_{n'}^D}. \tag{4.3}$$

Here the state space projectors $\hat{P}_n^X$ are defined to project into the kept (X = K) or discarded (X = D) space of Wilson shell n. This then allows to rewrite Eq. (4.3) more compactly as

$$\hat{P}_n^K = \sum_{n'>n}^{D} \hat{P}_{n'}^D. \tag{4.4}$$

With this, two independent sums over Wilson shells can be reduced into a single sum over shells,

$$
\begin{aligned}
\sum_{n_1,n_2} \hat{P}_{n_1}^D \hat{P}_{n_2}^D &= \sum_{(n_1=n_2)\equiv n} \hat{P}_n^D \hat{P}_n^D + \sum_{n_1>(n_2\equiv n)} \hat{P}_{n_1}^D \hat{P}_n^D + \sum_{(n_1\equiv n)<n_2} \hat{P}_n^D \hat{P}_{n_2}^D \\
&= \sum_n \left(\hat{P}_n^D \hat{P}_n^D + \hat{P}_n^K \hat{P}_n^D + \hat{P}_n^D \hat{P}_n^K \right) \\
&\equiv \underbrace{\sum_n^{XX'\neq KK} \hat{P}_n^X \hat{P}_n^{X'}}_{\equiv \sum_n'}.
\end{aligned} \tag{4.5}
$$

For simplified notation, the prime in the last single sum over Wilson shells ($\sum'$) indicates that also the kept-sectors are included in the sum, yet excluding the all-kept sector XX' $\neq$ KK, since this sector is still refined in later iterations.[2,68]

While Eq. (4.5) holds for the entire Wilson chain, exactly the same line of arguments can be repeated starting from some arbitrary but fixed reference shell n,

$$\sum_{n_1,n_2>n} \hat{P}_{n_1}^D \hat{P}_{n_2}^D = \underbrace{\sum_{n'>n}^{XX'\neq KK} \hat{P}_{n'}^X \hat{P}_{n'}^{X'}}_{\equiv \sum_{n'>n}'}.$$

Together with Eq. (4.4), this becomes

$$\hat{P}_n^K \hat{P}_n^K = \sum_{n_1,n_2>n} \hat{P}_{n_1}^D \hat{P}_{n_2}^D = \sum_{n'>n}' \hat{P}_{n'}^X \hat{P}_{n'}^{X'}. \tag{4.6}$$

4.1.1   Generalization to multiple sums over shells

Consider the evaluation of some physical correlator that requires m insertions of the identity in Eq. (4.2b) in order to obtain a simple Lehmann representation. Examples are thermal expectation values of observables or (higher-order) Greens functions. In all cases, the resulting independent sum over arbitrarily many identities as in Eq. (4.2b) can always be rewritten into a *single* sum over Wilson shells.

Claim: Given m full sums as in Eq. (4.2b), this can be rewritten in terms of a single sum over a Wilson shell n, such that Eq. (4.5) generalizes to

$$\sum_{n_1,\ldots,n_m} \hat{P}_{n_1}^{D_1} \ldots \hat{P}_{n_m}^{D_m} = \underbrace{\sum_n^{X_1 \cdots X_m \neq K_1 \ldots K_m} \hat{P}_n^{X_1} \ldots \hat{P}_n^{X_m}}_{\equiv \sum_n'} \tag{4.7}$$

Proof: Note that via Eq. (4.4), the *l.h.s.* of Eq. (4.7) can be rewritten as

$$\hat{P}_{n_0-1}^{K_1} \ldots \hat{P}_{n_0-1}^{K_m} = \sum_{n_1,\ldots,n_m} \hat{P}_{n_1}^{D} \ldots \hat{P}_{n_m}^{D}$$

where $n_0 > 0$ is the first iteration where truncation occurred. This way, $\hat{P}_{n_0-1}^{K}$ refers to the full Hilbert space still. Proving Eq. (4.7) hence is again equivalent to proving for general n that

$$\hat{P}_n^{K_1} \ldots \hat{P}_n^{K_m} = \sum_{n_1,\ldots,n_m>n} \hat{P}_{n_1}^{D} \ldots \hat{P}_{n_m}^{D} = \sum_{n'>n}' \hat{P}_{n'}^{X_1} \ldots \hat{P}_{n'}^{X_m}, \tag{4.8}$$

where the sum in the middle term indicates an independent sum of all $n_i > n$ with $i = 1, \ldots, m$. The case of two sums ($m = 2$) was already shown in Eq. (4.6). Hence one may proceed via induction. Assume, Eq. (4.8) holds for $m - 1$. Then for the case m one has,

$$\hat{P}_n^{K_1} \ldots \hat{P}_n^{K_{m-1}} \cdot \hat{P}_n^{K_m} =$$
$$= \left(\sum_{n'>n}' \hat{P}_{n'}^{X_1} \ldots \hat{P}_{n'}^{X_{m-1}} \right) \left(\sum_{n_m>n}^{D} \hat{P}_{n_m}^{D_m} \right)$$
$$= \left(\sum_{n'=n_m>n}' + \sum_{n<n'<n_m}' + \sum_{n'>n_m>n}' \right) \hat{P}_{n'}^{X_1} \ldots \hat{P}_{n'}^{X_{m-1}} \cdot \hat{P}_{n_m}^{D_m}$$
$$= \sum_{\bar{n}>n}' \hat{P}_{\bar{n}}^{X_1} \ldots \hat{P}_{\bar{n}}^{X_{m-1}} \left(\hat{P}_{\bar{n}}^{D_m} + \hat{P}_{\bar{n}}^{K_m} \right) + \hat{P}_{\bar{n}}^{K_1} \ldots \hat{P}_{\bar{n}}^{K_{m-1}} \cdot \hat{P}_{\bar{n}}^{D_m}$$
$$\equiv \sum_{\bar{n}>n}' \hat{P}_{\bar{n}}^{X_1} \ldots \hat{P}_{\bar{n}}^{X_m},$$

where the last term in the fourth line followed from the inductive hypothesis. ∎

More intuitively, the m independent sums over $\{n_1, \ldots n_m\}$ in Eq. (4.8) can be rearranged such, that for a specific iteration $\bar{n}$, either one of the indices n_i may carry $\bar{n}$ as minimal value, while all other sums range from $n_{i'} \geq \bar{n}$. This way, by construction, the index n_i stays within the *discarded* state space, while all other sums $n_{i'}$ are unconstrained up to $n_{i'} \geq n_i = \bar{n}$, thus represent either discarded at iteration $\bar{n}$ or discarded at any later iteration which corresponds to the kept space at iteration $\bar{n}$. From this, Eq. (4.8) immediately follows.

4.2 Full density matrix

Given the complete NRG energy eigenbasis $|se\rangle_n^{\mathrm{D}}$, the full density matrix (FDM)[68] at arbitrary temperature $T \equiv 1/\beta$ is given by

$$\hat{\rho}^{\mathrm{FDM}}(T) = \sum_{sen} \frac{e^{-\beta E_s^n}}{Z} |se\rangle_{nn}^{\mathrm{DD}} \langle se|, \tag{4.9}$$

with $Z(\beta) \equiv \sum_{ne,s \in \mathrm{D}} e^{-\beta E_s^n}$. By construction of a thermal density matrix, all energies E_s^n from all shells n appear on an equal footing relative to a single global energy reference. Hence any prior iterative rescaling or shifting of the energies E_s^n, which is a common procedure within the NRG, *must be undone*. From a numerical point of view, typically the ground state energy at the last iteration $n = N$ for a given NRG run is taken as energy reference. This enforces that all Boltzmann weights are smaller or equal 1.

Note that the energies E_s^n are considered *independent* of the environmental index e. As a consequence, this leads to *exponentially large* degeneracies in energy for the states $|se\rangle_n$. The latter must be properly taken care of within FDM, as it contains information from all shells. By already tracing out the environment for each shell (this will be further justified below), this leads to

$$\hat{\rho}^{\mathrm{FDM}}(T) = \sum_n \underbrace{\frac{d^{N-n} Z_n}{Z}}_{\equiv w_n} \underbrace{\sum_s \frac{e^{-\beta E_s^n}}{Z_n} |s\rangle_{nn}^{\mathrm{DD}} \langle s|}_{\equiv \rho_n^{\mathrm{FDM}}(T)}, \tag{4.10}$$

with d the state-space dimension of a single Wilson site, and introducing the normalization $Z_n(\beta) \equiv \sum_{s \in \mathrm{D}_n} e^{-\beta E_s^n}$ of the density matrices $\rho_n^{\mathrm{FDM}}(T)$ built from the discarded space of a specific shell n only. Therefore, $Z = \sum_n Z_n$ and $\mathrm{tr}(\rho_n^{\mathrm{FDM}}(T)) = 1$.

The qualitative behavior of the weights w_n can be understood straightforwardly. With the typical energy scale of shell n given by

$$\omega_n = a\Lambda^{-n/2}, \tag{4.11}$$

with a some constant of order 1. [*cf.* Eq. (3.3c)], this allows to estimate the weights as follows,

$$\ln(w_n) \simeq \ln\!\big(d^{N-n} e^{-\beta \omega_n}/Z\big) = (N-n)\ln(d) - \beta \omega_n + \mathrm{const},$$

the maximum of which as a function of shell n is determined by

$$\frac{d}{dn}\ln(w_n) \;\simeq\; -\ln(d) + \frac{a\beta\ln(\Lambda)}{2}\Lambda^{-n/2} \stackrel{!}{=} 0,$$

with the solution

$$a\Lambda^{-n^*/2} \simeq \frac{2\ln(d)}{\beta\ln(\Lambda)} \sim T. \tag{4.12}$$

The middle term is $1/\beta$ times some constant of order 1., from which the last proportionality follows. This shows that the distribution w_n is strongly peaked around the energy scale of given temperature T. With $T \equiv a\Lambda^{-n_T/2}$ and therefore $n_T \simeq n^*$, the distribution therefore decays super-exponentially fast towards larger energies $n \ll n_T$ (dominated by $e^{-\beta\omega_n}$ with *exponentially* increasing ω_n with decreasing n), while it decays in a plain exponential fashion towards smaller energy scales $n \gg n_T$ (dominated by d^{-n}, since with $\beta\omega_n \ll 1$, $e^{-\beta\omega_n} \to 1$). An actual example in terms of the single impurity Anderson model (SIAM) is shown in Fig. 4.2 which clearly supports all of the above qualitative analysis. It follows for typical discretization parameter Λ and local dimension d, that n_T is slightly smaller than n^*, *i.e.* to the left of the maximum in w_n typically at the onset of the distribution w_n on a linear scale, as is seen in Fig. 4.2 (n_T is indicated by the vertical dashed line). An important consequence of the exponential decay of the weights w_n for $n \gg n_T$ is that by taking a long enough Wilson chain to start with, FDM-NRG *automatically* truncates the length of the Wilson chain around $n \sim n_T$.

The weights w_n are fully determined within an NRG calculation, yet depend sensitively on the specific physical as well as numerical parameters. Most obviously, this includes the state space dimension d of a given Wilson site. However, the weights w_n clearly also depend on the specific number of states kept from one iteration to the next. For example, the weights are clearly zero for iterations where no truncation takes place. This is usually the case for the first few NRG iterations including the impurity. In general, the weights fully adapt to the specific truncation scheme adopted, which includes, for example, adaptive truncation based on an energy threshold. In the case of fixed $N_K = 512$ in Fig. 4.2, note that if $d = 4$ times the number of states had been kept, *i.e.* $N_K = 512 \to 2048$, this essentially would have shifted the entire weight distribution in Fig. 4.2 by one iteration to lower energy scales, *i.e.* to the right, resulting in an improved spectral resolution for frequencies $\omega \lesssim T$.[68]

Interestingly, given a constant number N_K of kept states in Fig. 4.2, the weights w_n show a completely *smooth* behavior, irrespective of even or odd iteration. This is somewhat surprising at first glance, considering that NRG typically does show even-odd behavior, specifically so for the SIAM. There, at even iterations an overall non-degenerate singlet can be formed to be the ground state. Having no unpaired spin in the system, this typically lowers the energy more strongly as compared to odd iterations which do have an unpaired spin. Therefore while even iterations show a stronger energy reduction in its low energy states, its ground state space consists of a single state. In contrast, for odd iterations the energy reduction by adding the new site is weaker, yet the ground state space is degenerate, having no magnetic field (Kramers degeneracy). As can be clearly seen in Fig. 4.2, both
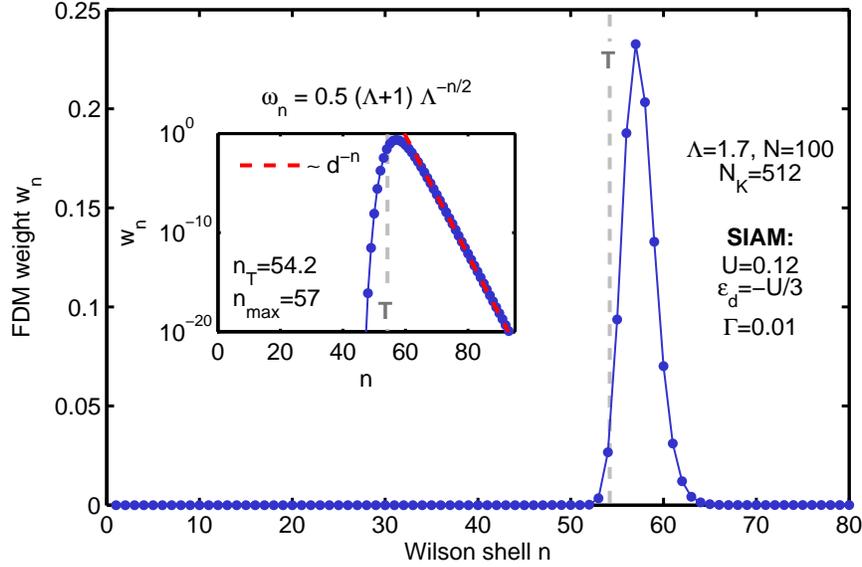
Figure 4.2: Typical FDM weight distribution calculated for the SIAM for the parameters as shown in the panel [*cf.* Eq. (3.5); all energies in units of bandwidth]. The number of states N_K kept at every iteration was taken constant. The distribution is strongly peaked around the energy shell n_T (vertical dashed line) corresponding to the energy scale of temperature. The inset shows the weights w_n on a logarithmic scale, showing plain exponential decay for small energies $n > n_T$, and super-exponentially fast decay towards large energies ($n < n_T$).

effects combined together compensate each other, resulting in a smooth distribution of the FDM weights w_n.

Above analysis shows that the density matrix generated by FDM is dominated by several shells around the energy scale of temperature. The physical information encoded in these shells can critically affect physical observables at much larger energies. This construction therefore *shall not* be shortcut in terms of the density matrix in the kept space at much earlier iterations, say using $\hat{H}|s\rangle_n^K \simeq E_s^n |s\rangle_n^K$ with the Boltzmann weights in Eq. (4.10). The latter can fail for exactly the reasons already discussed in detail with the introduction of the DM-NRG by Hofstetter (2000): the low-energy physics can have important feedback to larger energy scales. To be specific, the low-energy scales on the order of temperature can have a decisive role on the decay channels of high-energy excitations. As a result, for example, the low-energy physics can lead to a significant redistribution of spectral weight in the local density of states at large energies.

4.3 Applications

4.3.1 Spectral functions

The calculation of spectral properties within the NRG started with Oliveira and Wilkins (1981) in the context of X-ray absorption spectra, which was extended to spectral functions at zero temperature by Sakai et al. (1989). Finite temperature together with transport properties was finally introduced by Costi and Hewson (1992). An occasionally crucial feedback from small to large energy scales finally was taken care of by the explicit incorporation of the reduced density matrix for the remainder of the Wilson chain (DM-NRG) by Hofstetter (2000). While these methods necessarily combined data from all NRG iterations to cover the full spectral range, they did so through heuristic patching schemes. Moreover, in the case of finite temperature, these methods had been formulated in a single-shell setup that associates a well-chosen temperature corresponding to the energy scale of this shell. For a more complete listing of references see Bulla et al. (2008).

The possible importance of a true multi-shell framework for out-of-equilibrium situations had already been pointed out by Costi (1997). As it turns out, this can be implemented in a fully transparent well-defined way using the complete basis sets by Anders and Schiller (2005). While nevertheless more traditional single-shell formulations of the NRG exist using complete basis sets,[48] a clean multi-shell formulation can be obtained using the full density matrix in Eq. (4.9) (FDM-NRG, Weichselbaum and von Delft (2007)). The clean treatment of reduced density matrices to describe the low-energy sector essentially generalizes the DM-NRG[31] to a black-box algorithm, which in addition also allows to treat arbitrary finite temperatures on a fully generic footing.

Consider the retarded Greens function

$$G^R_{BC}(t) = -i\vartheta(t)\underbrace{\langle \hat{B}(t)\hat{C}^\dagger \rangle_T}_{\equiv G_{BC}(t)} \tag{4.13}$$

with $\hat{B}(t) \equiv e^{i\hat{H}t}\hat{B}e^{-i\hat{H}t}$. Here the Hamiltonian $\hat{H}$ of the system is assumed to be time-independent. In Eq. (4.13), the operator $\hat{C}^\dagger$ acts at time $t = 0$ on a system in thermal equilibrium at temperature T, described by the thermal density matrix $\hat{\rho}(T) = \frac{e^{-\beta\hat{H}}}{Z}$, i.e. $\langle \cdot \rangle_T \equiv \text{tr}(\hat{\rho}(T)\cdot)$. The system then evolves to some time $t > 0$, where a possibly different operator $\hat{B}$ acts on the system. The overlap with the time evolved wave function then defines the causal correlation of the two events. Fourier-transformed into frequency space, $G^R(\omega) \equiv \int \frac{dt}{2\pi}e^{i\omega t}G^R(t)$, its spectral function is defined by

$$\begin{aligned} A_{BC}(\omega) &\equiv -\frac{1}{\pi}\text{Im}G^R_{BC}(\omega) = \int \frac{dt}{2\pi}e^{i\omega t}G(t) \\ &= \int \frac{dt}{2\pi}e^{i\omega t}\,\text{tr}\left(\hat{\rho}(T)e^{i\hat{H}t}\hat{B}e^{-i\hat{H}t}\hat{C}^\dagger\right). \end{aligned} \tag{4.14}$$

When evaluated in the full many-body eigenbasis, this requires the insertion of two identities as in Eq. (4.2b), (i) to evaluate the trace, and (ii) in between the operators $\hat{B}$ and $\hat{C}^\dagger$. Using

for simplified notation, the eigenbasis sets $\mathbf{1} = \sum_a |a\rangle\langle a| = \sum_b |b\rangle\langle b|$, the spectral function becomes,

$$
\begin{aligned}
A_{BC}(\omega) &= \sum_{ab} \int \frac{dt}{2\pi} e^{i(\omega - E_{ab})t} \rho_a \langle a|\hat{B}|b\rangle\langle b|\hat{C}^\dagger|a\rangle \\
&\equiv \sum_{ab} \rho_a B_{ab} C_{ab}^* \cdot \delta\left(\omega - E_{ab}\right),
\end{aligned}
\tag{4.15}
$$

with $E_{ab} \equiv E_b - E_a$ and $\rho_a \equiv \frac{e^{-\beta E_a}}{Z}$. Equation (4.15) is referred to as the Lehmann representation of given correlation function. In the case of equal operators, $\hat{B} = \hat{C}$, the spectral function is a strictly positive function, *i.e.* a spectral density. In either case, the integrated spectral function results in the plain thermodynamic expectation values,

$$
\int d\omega A_{BC}(\omega) = \sum_{ab} \rho_a B_{ab} C_{ab}^* = \left\langle \hat{B}\hat{C}^\dagger \right\rangle_T.
\tag{4.16}
$$

Using the complete NRG eigenbasis, $|a\rangle \to |se\rangle_n$ and $|b\rangle \to |s'e'\rangle_{n'}$, one may be tempted of directly reducing the double sum in Eq. (4.15) to a single sum over Wilson shells, as in Eq. (4.5). This implies that the thermal weight would be constructed as $\rho_a(T) \sim e^{-\beta E_a} \to e^{-\beta E_s^{n,X}}$ from simply the discarded (X = K) or kept (X = K) space at iteration n. This ignores, however, a possible feedback from small to large energy scales which has been shown to be crucial in certain NRG calculations by the DM-NRG.[31] The solution is to keep track of the FDM as it stands in Eq. (4.9). This, however, introduces another independent sum c over Wilson shells in addition to a and b which, nevertheless, can be dealt with in complete analogy to Eq. (4.7). With $\{a, b, c\} \in \{\text{XX}'\text{X}_\rho \neq \text{KKK}\}$, respectively, it turns out that X = X_ρ are locked to each other due to the orthogonality of the discarded and the kept states at iteration n. Therefore only the contributions $\text{XX}' \neq \text{KK}$ as known from a double sum remain. Nevertheless, the FDM in Eq. (4.9) must be projected into the respective state spaces.

Let me give a more detailed explanation of the previous argument. Using the cyclic property of the trace in Eq. (4.14), the Lehmann representation in Eq. (4.15) can be rewritten as

$$
A_{BC}(\omega) \equiv \sideset{}{'}\sum_{n,ss'} (\hat{C}^\dagger \hat{\rho}_T^{\text{FDM}})_{s's} B_{ss'} \delta\left(\omega - E_{ss'}^n\right).
\tag{4.17}
$$

The prime with the sum in Eq. (4.17) again indicates that only states $ss' \notin \text{KK}$ are to be considered. The MPS diagram of underlying tensor structure is shown in Fig. 4.3. Every *leg* of the "ladders" in Fig. 4.3 corresponds to an NRG eigenstate (MPS) $|s\rangle_n$ for some intermediate iteration n. The blocks for the MPS coefficient spaces (A-tensors) are no longer drawn, for simplicity. The outer sum over the states s' in Eq. (4.17) corresponds to the overall trace. Hence the upper- and lower-most leg in Fig. 4.3 at iteration n carry the same state label s' which are connected by a line (contraction). Furthermore, the inserted identity in the index s initially also would have been identified with two legs [similar to
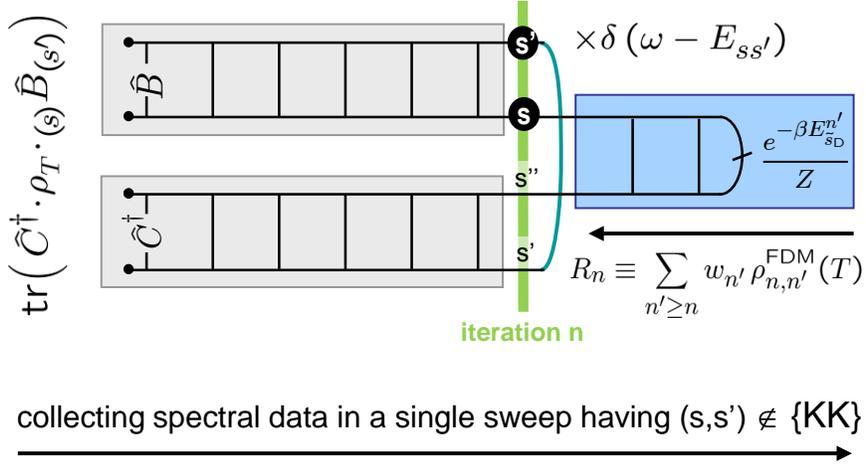
Figure 4.3: MPS diagram for calculating spectral functions within the FDM-NRG. Spectral functions are calculated after a prior NRG *forward sweep*, which generates the NRG eigenbasis decomposition (horizontal ladders, with the boxes for the A-tensors skipped for simplicity; *cf.* Fig. 2.3). The calculation of a spectral function as in Eq. (4.15) then includes the evaluation of the matrix elements $\mathrm{tr}\big(\hat{A} \cdot \rho_T \cdot {}_{(s)}\hat{B}_{(s')}\big)$, as indicated to the left of above figure. The energies of the indices (states) s and s' are "probed" such that their difference determines the energy $\omega = E_{ss'}^n \equiv E_{s'}^n - E_s^n$ of an individual contribution to the spectral function, as indicated by the $\times\delta(\omega - E_{ss'})$ next to the indices s and s' in the upper right of the figure. The sum $\sum_{n'>n}$ in the discarded state space of $\hat{\rho}^{\mathrm{FDM}}(T)$, indicated to the lower right, results in the object R_n. The individual contributions $\rho_{n,n'}^{\mathrm{FDM}}(T)$ are generated by the Boltzmann weights in the discarded space at iteration n', as indicated to the right. The contribution at $n' = n$, *i.e.* R_n^{D}, can simply be determined when needed. On the other hand, the cumulative contributions $n' > n$ are obtained in a simple *backward sweep*, starting from the last Wilson shell N included, as indicated by the small arrow pointing to the left. Having $n' > n$, this calculation always maps to the kept space, thus resulting in R_n^{K}. As such, this operation is typically much faster as compared to the actual collection of the spectral data in a final forward sweep, as indicated at the bottom of the figure.

what is seen in Fig. 4.4 later]. At iteration n, however, the state space s directly hits the FDM, leading to the overlap matrix ${}_n^{\mathrm{X}}\langle s|s'\rangle_n^{\mathrm{X'}} = \delta_{ss'}\delta_{\mathrm{XX'}}$ [hence this eliminates the second block from the top in Fig. 4.4]. Therefore only the single index s from the second complete sum remains in Fig. 4.3.

The two legs in the center of Fig. 4.3 finally stem from the insertion of the FDM which can extend to all iterations $n' \geq n$. Note that the case $n' < n$ does not appear, since there the discarded state space used for the construction of the FDM is orthogonal to the state space s at iteration n. This justifies to trace out the environmental states $|e\rangle_n$ for the density matrices $\rho_n^{\mathrm{FDM}}(T)$, as already pointed out with Eq. (4.10). The FDM thus reduces

to the following structure at iteration n,

$$R_n \equiv \sum_{n' \geq n} w_{n'} \rho_{n,n'}^{\mathrm{FDM}}(T) \tag{4.18}$$

with the (*reduced*) density matrices

$$\rho_{n,n'}^{\mathrm{FDM}}(T) \equiv \operatorname*{tr}_{\{\sigma_{n+1},\dots,\sigma_{n'}\}}\big(\hat{\rho}_{n'}^{\mathrm{FDM}}(T)\big), \tag{4.19}$$

defined for $n' \geq n$, with $\rho_{n,n}^{\mathrm{FDM}}(T) = \rho_n^{\mathrm{FDM}}(T)$. Note that in the definition of the $\hat{\rho}_{n'}^{\mathrm{FDM}}(T)$ in Eq. (4.10) the *environment* consisting of all sites $\bar{n} > n'$ had already been traced out, hence only the sites $\bar{n} = n+1, \dots, n'$ remain to be considered in Eq. (4.19). For $n' > n$ then, $\rho_{n,n'}^{\mathrm{FDM}}(T)$ is built from the effective basis $|s\rangle_n^{\mathrm{D}}$ at iteration n, where subsequently the local state spaces $\sigma_{\bar{n}}$ of sites $\bar{n} = n', n'-1, \dots, n+1$ are traced out. In Eq. (4.19), $\rho_{n,n'}^{\mathrm{FDM}}(T)$ and thus also R_n are understood as the matrix elements *w.r.t.* to the states $|s\rangle_n^{\mathrm{X}}$, hence are simple coefficient spaces written without hats (the latter are reserved for explicit operators).

The object R_n can be written in terms of two contributions, (i) the contribution from iteration $n' = n$ itself (encoded in D-space, *i.e.* the discarded space X=D), and (ii) the contributions of all later iterations $n' > n$ (encoded in K-space, *i.e.* the kept space X = K, at iteration n),

$$R_n = \underbrace{w_n \rho_n^{\mathrm{FDM}}(T)}_{\equiv R_n^{\mathrm{D}}} + \underbrace{\sum_{n' > n} w_{n'} \rho_{n,n'}^{\mathrm{FDM}}(T)}_{\equiv R_n^{\mathrm{K}}}, \tag{4.20}$$

Consequently, R_n only has matrix elements in K and D, referred to as R_n^{X} with X $\in \{$K, D$\}$, similar to the structure of the projectors in Eq. (4.3). By construction, there are no mixed matrix elements between K- and D-sectors (due to the orthogonality of kept and discarded space at iteration n). Therefore R_n has a simpler structure as compared to a generic operator with matrix elements $B^{\mathrm{XX}'}$. The fundamental reason for this is the fact that reduced density matrices transform as scalars with respect to the symmetries of the Hamiltonian, and thus are block-diagonal.

The double sum of Eq. (4.17) includes all sectors XX$'$ for a given iteration n except for the KK sector, *i.e.* $ss' \notin$ KK. In summary, by insisting on using the FDM in Eq. (4.9) this only leads to the minor complication that R_n^{K} needs to be constructed and included in the calculation. The construction of R_n^{K}, on the other hand, can be done in a simple prior backward sweep, which allows to generate R_n^{K} iteratively and thus efficiently. All of the R_n^{K} need to be stored for the actual calculation of the correlation function. Living in kept space, however, their respective space requirement is typically negligible.

Implications for complex Hamiltonians

Typically, the Hamiltonians analyzed by NRG are time-reversal invariant, and therefore can be computed using non-complex numbers. In case the Hamiltonian is not time-reversal

invariant, *i.e.* the calculation is intrinsically complex, the A-tensors on the lower leg of the ladders for the operators $\hat{B}$ and $\hat{C}^\dagger$ in Fig. 4.3 must be complex conjugated. For consistency, this implies for the FDM contributions R_n, that in Fig. 4.3 its corresponding A-tensors in the *upper* leg need to be complex conjugated.

Sum rules

By construction, FDM allows to exactly obey sum-rules for spectral functions as a direct consequence of Eq. (4.16) and fundamental quantum mechanical commutator relations. For example, after completing the Greens function in Eq. (4.13) to a proper many-body correlation function for fermions, $G_d(t) \equiv -i\vartheta(t)\langle\{\hat{d}(t), \hat{d}^\dagger\}\rangle_T$, with $\hat{d}^\dagger$ creating an electron in level d at the impurity and $\{\cdot, \cdot\}$ the anticommutator, the integrated spectral function results in

$$\int d\omega A(\omega) = \langle\{\hat{d}, \hat{d}^\dagger\}\rangle_T = 1, \tag{4.21}$$

due to the fundamental fermionic anticommutator relation, $\{\hat{d}, \hat{d}^\dagger\} = 1$. In practice, Eq. (4.21) is obeyed exactly within numerical double precision noise (10^{-16}), which underlines the fact that the full exponentially large quantum-many body state space can be dealt with in practice, indeed. The fundamental reason for this is the underlying one-dimensional MPS structure, as shown in Fig. 4.3.

4.3.2 Thermal expectation values

Arbitrary thermodynamic expectation values can be calculated consistently and accurately within the FDM-NRG framework through Eq. (4.16). Given the spectral data on the *l.h.s.* of Eq. (4.16), for example, this can be integrated to obtain the thermodynamic expectation value on the *r.h.s.* of Eq. (4.16). In practice, this corresponds to a simple sum of the non-broadened discrete spectral data obtained from FDM-NRG. Using the plain discrete data has the advantage that it does not depend on any further details of smoothening procedures which usually would somewhat larger error bars for the expectation values otherwise.

Typically only *local operators* are of interest within the NRG, in the sense that the operators $\hat{B}$ or $\hat{C}$ in Eq. (4.16), for example, only act at the impurity or in the bath at the location of the impurity (Wilson shell $n = 0$). With temperature usually much smaller than the bandwidth of the system, the weight distribution w_n has absolutely negligible contribution at very early iterations [see for example Fig. 4.2 and its discussion]. Moreover, with the state space truncation starting with iteration $n_0 > 0$, the reduced thermal density matrix for iteration n is fully described for $n < n_0$ by $R_n^{\rm K}(T)$. For a given temperature, the aforementioned simple backward sweep to calculate $R_n^{\rm X}$ then already provides all necessary information for the simple evaluation of the thermal expectation value of any local operator $\hat{C}$ [*e.g.* $\hat{C} := \hat{B}\hat{C}^\dagger$ in Eq. (4.16)],

$$\langle\hat{C}\rangle_T = \text{tr}\big[R_n^{(\rm K)}(T)C_n^{(\rm KK)}\big], \qquad (n < n_0) \tag{4.22}$$

with $C_n^{(\mathrm{KK})}$ the matrix elements of the operator $\hat{C}$ in the kept space of iteration n. With no truncation yet at iteration n, the kept space is the only space available, *i.e.* represents the full state space up to iteration n (hence the brackets around the K's). In Eq. (4.22), it was assumed that the operator $\hat{C}$ acts on sites $\leq n$ only. For local operators acting within the state space of H_0, only the matrix elements $C_0 \equiv C_0^{(\mathrm{KK})}$ are required, leading to $\langle \hat{C} \rangle_T = \mathrm{tr}\big[R_0(T)C_0\big]$ with $R_0(T) \equiv R_0^{(\mathrm{K})}(T)$.

In the case that the operator $\hat{C}$ is not a local operator, in the sense that it does not act on the very first sites of the Wilson chain, but nevertheless acts locally on some specific Wilson site n already in the presence of truncation, *i.e.* $n \geq n_0$, then using Eqs. (4.10) and (4.20) it follows,

$$\langle \hat{C} \rangle_T \;\; = \;\; \mathrm{tr}\big[R_n^{\mathrm{K}}(T)C_n^{\mathrm{KK}}\big] + \mathrm{tr}\big[R_n^{\mathrm{D}}(T)C_n^{\mathrm{DD}}\big] + c \sum_{n'<n} w_{n'}. \tag{4.23}$$

The last term in Eq. (4.23) derives from the discarded state spaces for the Wilson shells $n' < n$, and corresponds to the trace over the environmental states $|e\rangle_n$ in Eq. (4.10). Therefore the constant $c \equiv \frac{1}{d}\mathrm{tr}_{\sigma_n}\big(\hat{C}\big)$ is the plain average of the operator $\hat{C}$ in the local basis $|\sigma_n\rangle$ that it acts upon. Equation (4.22) follows from Eq. (4.23), in that for $n < n_0$, due to the absence of truncation, the second and third term in Eq. (4.23) are zero by construction.

4.3.3   Time-dependent NRG

Non-equilibrium dynamics for quantum quench settings were pioneered within the NRG by Costi (1997). Nevertheless, this original setting was still significant patch work. Only through the introduction of the complete basis sets by Anders and Schiller (2005) [*cf.* Eq. (4.2b)] finally, this could be based on a firm transparent framework. This milestone development allowed for the first time to use the quasi-exact method of NRG to perform real-time evolution to exponentially long time-scales. It emerged together with other approaches to real-time evolution of quantum many-body systems such as the DMRG.[17,71] While the complete basis sets were originally introduced within the NRG for the sole purpose of real-time evolution,[2] however, they turned out significantly more versatile.[48,66,68,69]

Starting from the thermal equilibrium of some initial Hamiltonian H^{I}, at time $t = 0$ a quench at the location of the quantum impurity occurs with the effect that for $t > 0$ the time-evolution is governed by a different final Hamiltonian H^{F}. While initially introduced within the single-shell framework for finite temperature,[2] the same analysis can also be rigorously generalized to the multi-shell approach of FDM-NRG. Here the description will focus on the FDM approach.

In the presence of a quantum quench, the time-dependent expectation value of interest is

$$C(t) \;\; \equiv \;\; \langle \hat{C}(t) \rangle_T \equiv \mathrm{tr}\big[\rho^{\mathrm{I}}(T) \cdot e^{iH^{\mathrm{F}}t}\hat{C}e^{-iH^{\mathrm{F}}t}\big], \tag{4.24}$$

with $\hat{C}$ some observable. While one is interested in the dynamics after the quench only, *i.e.* for $t > 0$, one is nevertheless free to extend the definition of Eq. (4.24) also to negative times. The advantage of doing so is, that the Fourier transform into frequency space of the $C(t)$ in Eq. (4.24) defined for arbitrary times becomes purely real, as will be shown shortly. With this the actual time-dependent calculation can be performed in frequency space first in a simple and for the NRG natural way,

$$
C(\omega) \;\; = \;\; \int \tfrac{dt}{2\pi} e^{i\omega t} \operatorname{tr}\!\big(\hat{\rho}^{\mathrm{I}}(T) \cdot e^{i\hat{H}^{\mathrm{F}}t}\hat{C}e^{-i\hat{H}^{\mathrm{F}}t}\big) \; , \tag{4.25}
$$

A Fourier transform back into the time-domain at the end of the calculation finally provides the desired time-dependent expectation value $C(t) = \int C(\omega)e^{-i\omega t}\,d\omega$ for $t \geq 0$. In order to obtain smooth data closer to the thermodynamic limit, a weak log-Gaussian broadening in frequency space quickly eliminates artificial oscillations that come from the logarithmic discretization. Note that for the sole purpose of damping these artificial oscillations in the time domain, typically a *significantly* smaller broadening parameter $\alpha \lesssim 0.1$ suffices as compared to what is typically used to obtain fully smoothened correlation functions in the frequency domain, *e.g.* $\alpha \gtrsim 0.5$ for $\Lambda = 2$ [see EPAPS in Weichselbaum and von Delft (2007)].

Lehmann representation

In order to evaluate Eq. (4.25), three complete basis sets are required: one completed basis set c derived from an NRG run in H^{I} to construct $\rho^{\mathrm{I}}(T)$, and two complete basis sets a and b from an NRG run in H^{F} to be inserted right before and after the $\hat{C}$ operator, respectively, to describe the dynamical behavior. Clearly, two NRG runs in H^{I} and H^{F} are required to describe the quantum quench. With this, the spectral data in Eq. (4.25) becomes

$$
C(\omega) \;\; = \;\; \sum_{abc} \underbrace{\langle a|c\rangle}_{\equiv S_{ac}} \cdot \rho_c^{\mathrm{I}}(T) \cdot \underbrace{\langle c|a\rangle}_{\equiv S_{ac}^*} \cdot C_{ab} \cdot \delta\left(\omega - E_{ab}^{\mathrm{F}}\right) , \tag{4.26}
$$

which generates the overlap matrix S. Using the complete NRG eigenbasis sets, the corresponding MPS diagram that needs to be evaluated, is shown in Fig. 4.4.

The MPS diagram in Fig. 4.4 is similar to Fig. 4.3, yet with a few essential differences: the block describing the matrix elements of the original operator $\hat{B}$ has now become the block containing $\hat{C}$. The original operator $\hat{C}^\dagger$ is absent, *i.e.* has become the identity. Yet since its "matrix elements" are calculated with respect to two different basis sets (initial and final Hamiltonian), an overlap matrix remains (lowest block in Fig. 4.4). In context of the correlation functions in Fig. 4.3, the complete basis set in the index s could be reduced to the single index s. Here, however, two different complete basis sets hit upon each other, which inserts another overlap matrix (second block from the top in Fig. 4.4, which corresponds to the Hermitian conjugate of the lowest block). The reduced density matrices R_n^{X}, finally, are built from the initial Hamiltonian, yet are completely identical in structure otherwise to the ones already discussed with Eq. (4.20) for correlation functions.
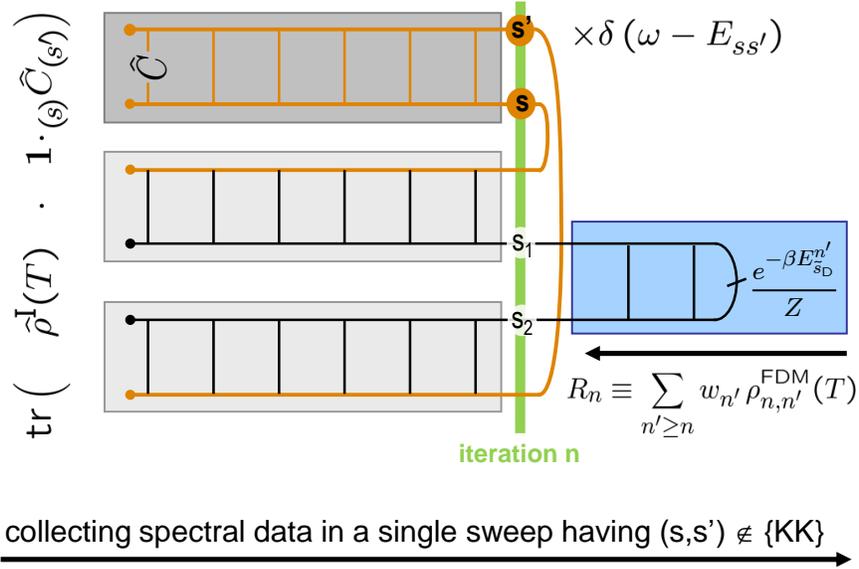
Figure 4.4: MPS diagram for the simulation of quantum quenches in the NRG (TDM-NRG with extension to FDM). The calculation is performed in frequency space as depicted, which only at the very end is Fourier transformed into the time domain to obtain the desired time-dependent expectation value $\hat{C}(t)$. The calculation requires initial (black horizontal lines) and final Hamiltonian (dark orange horizontal lines). An overlap matrix is required between the complete basis sets of initial and final Hamiltonian (light gray boxes at the lower left), which are computed in two prior NRG runs. The partial contributions R_n^{I} to the FDM (box at the lower right) are evaluated with respect to the initial Hamiltonian, but have exactly the same structure otherwise as already described with Fig. 4.4. The plain contraction $SRS^\dagger$ *w.r.t.* the indices s_1 and s_2 at the bottom of the MPS diagram can simply be evaluated through efficient matrix multiplication, while nevertheless respecting the block structure of the symmetries included in the calculation.

The basis of the initial Hamiltonian enters through the two legs connecting to the density matrix in Fig. 4.4 (horizontal black lines), while all other legs refer to the NRG basis generated by the final Hamiltonian [horizontal dark orange (gray) lines]. It is emphasized here, that the reduction of multiple sums in Wilson shells as in Eqs. (4.5) and (4.7) is *not* constrained to having the complete basis sets being identical to each other. It is easy to see that it equally applies to the current context of different basis sets from initial and final Hamiltonian.

4.3.4 Fermi-Golden-Rule calculations

The NRG is perfectly suited to deal with local quantum events such as absorption or emission of a generalized quantum impurity system in contact with non-interacting reser-

voirs.[30,38,45,61,69] If the rate of absorption is weak, such that the system has sufficient time to equilibrate on average, then the resulting absorption spectra are described by the Fermi-Golden rule,[54]

$$
A(\omega) \;\; = \;\; 2\pi \sum_{i,f} \rho_i^{\mathrm{I}}(T) \cdot |\langle f|\hat{C}|i\rangle|^2 \cdot \delta(\omega - E_{if}), \tag{4.27}
$$

where i and f describe complete basis sets for initial and final system, respectively. The system starts in the thermal equilibrium of the initial system. The operator $\hat{C}$ describes the absorption event at the impurity system, *i.e.* corresponds to the term in the Hamiltonian that couples to the light field. The transition amplitudes between initial and final Hamiltonian are fully described by the matrix elements $C_{fi} \equiv \langle f|\hat{C}|i\rangle$. In Eq. (4.27), the frequency ω shows threshold behavior, with the frequency threshold given by the difference in the ground state energies of initial and final Hamiltonian, $\omega_{\mathrm{thr}} \equiv \Delta E_g \equiv E_g^{\mathrm{F}} - E_g^{\mathrm{I}}$.

The only difference between emission and absorption spectra is the reversed role of initial and final system, while also having $\hat{C} \to \hat{C}^\dagger$. Specifically, the emission process starts in the thermal equilibrium of the final Hamiltonian, with subsequent transition matrix elements to the initial system. This also implies that emission spectra have their contributions at negative frequencies, *i.e.* frequencies smaller than the threshold frequency yet blurred by temperature, indicating the emission of a photon.

While absorption or emission spectra are already defined in frequency domain, they can nevertheless be translated into the time domain through Fourier transform,

$$
\begin{aligned}
A(t) \;\; &\equiv \;\; \int \frac{d\omega}{2\pi} e^{-i\omega t} A(\omega) = \sum_{i,f} \rho_i^{\mathrm{I}}(T) \langle i|\hat{C}^\dagger|f\rangle e^{-iE_f t} \langle f|\hat{C}|i\rangle e^{iE_i t} \\
&= \;\; \big\langle \underbrace{e^{i\hat{H}^{\mathrm{I}}t} \hat{C}^\dagger e^{-i\hat{H}^{\mathrm{F}}t}}_{\equiv \hat{C}(t)} \cdot \hat{C} \big\rangle_T^{\mathrm{I}}.
\end{aligned} \tag{4.28}
$$

Thus absorption spectra can also be interpreted similar to correlation functions: at time $t = 0$ an absorption event occurs (application of $\hat{C}$, which for example rises an electron from a low lying level into some higher level that participates in the dynamics). This alters the Hamiltonian, such that the subsequent time evolution is governed by the final Hamiltonian. At some time $t > 0$ then, the absorption event relaxes back to the original configuration (application of $\hat{C}^\dagger$), such that $A(t)$ finally describes the overlap amplitude of the resulting state with the original state with no absorption. While the "mixed" time evolution of $\hat{C}(t)$ in Eq. (4.28) appears somewhat artificial, it can be easily rewritten in terms of a regular time-dependent Heisenberg operator with a *single* Hamiltonian. By explicitly including a further static degree (*e.g.* a low lying hole from which the electron was lifted through the absorption event, or the photon itself), this switches $\hat{H}^{\mathrm{I}}$ to $\hat{H}^{\mathrm{F}}$, *i.e.* between two dynamically disconnected sectors in Hilbert space [compare discussion of type-1 and type-2 quenches in Münder et al. (2011)].

Within the complete NRG eigenbasis, Eq. (4.27) becomes

$$
A(\omega) \;\; = \;\; 2\pi \sum_{ss' \notin \mathrm{KK}} {}_n^{\mathrm{F}}\langle s'|\hat{A}|s\rangle_n^{\mathrm{I}} \cdot R_n^{\mathrm{X}_s\mathrm{I}}(T) \cdot {}_n^{\mathrm{I}}\langle s|\hat{A}^\dagger|s'\rangle_n^{\mathrm{F}} \times \delta(\omega - E_{ss'}), \tag{4.29}
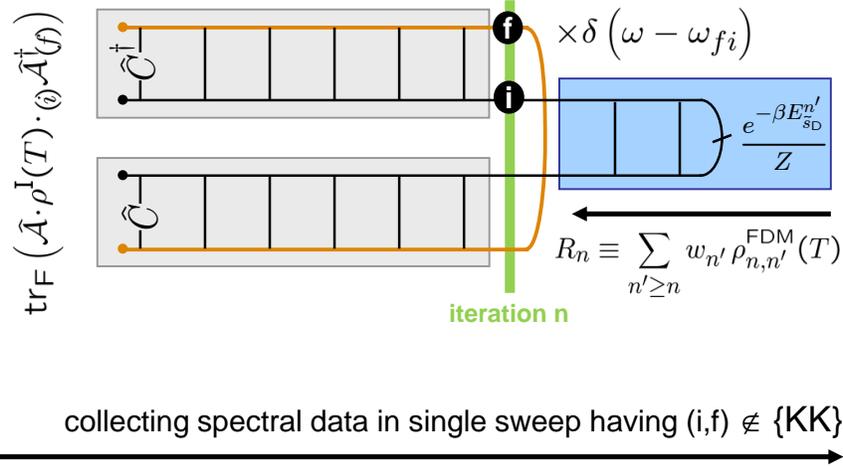$$

Figure 4.5: MPS diagram for the calculation of absorption or emission spectra using Fermi-Golden-rule (FGR-NRG) mediated by the operator $\hat{C}$. The two center legs (horizontal black lines) refer to the state space of the initial Hamiltonian, while the outer legs (horizontal dark gray lines) refer to the state space of the final Hamiltonian. Note that the matrix elements of $\hat{C}$ are *mixed* matrix elements between eigenstates of initial and final Hamiltonian.

with $X_s \in \{K, D\}$ the state space sector of state s. The MPS diagram for Eq. (4.29) to be evaluated is shown in Fig. 4.5. Its structure is completely analogous to the calculation of generic correlation functions in Fig. 4.3, except that similar to the quantum quench earlier, here again the basis sets from two different Hamiltonians come into play.[30] In contrast to the quantum quench situation in Fig. 4.4, however, *no* overlap matrices emerge. Instead, all matrix elements of the local operator $\hat{C}^\dagger$ are calculated in a mixed basis between initial and final eigenstates. The double sum over Wilson shells (one from the outer trace Fig. 4.5 in the complete basis of the final Hamiltonian, and one in the construction of the FDM) is again reduced to a single sum over Wilson shells with the constraint $(i, f) \equiv (s, s') \notin$ KK. The reduced density matrices $R_n^{X,I}$ are constructed *w.r.t.* the initial Hamiltonian, but exactly correspond to the ones introduced in the FDM context in Eq. (4.20) otherwise.

Technical remarks

Absorption or emission spectra in the presence of Anderson orthogonality or strongly correlated low-energy physics typically exhibit sharply peaked features close to the threshold frequency with clear physical interpretation. While in principle, a single Hamiltonian with dynamically disconnected Hilbert space sectors may be used, this is ill-suited for an NRG simulation. Using a single NRG run, this can only resolve the low-energy of the full Hamiltonian, *i.e.* of the initial system as it is assumed to lie lower. Consequently, the sharp features at the threshold frequency will have to be smoothened by an energy window comparable to $\omega_{\text{thr}} = \Delta E_g$ in order to suppress discretization artifacts. This problem is fully
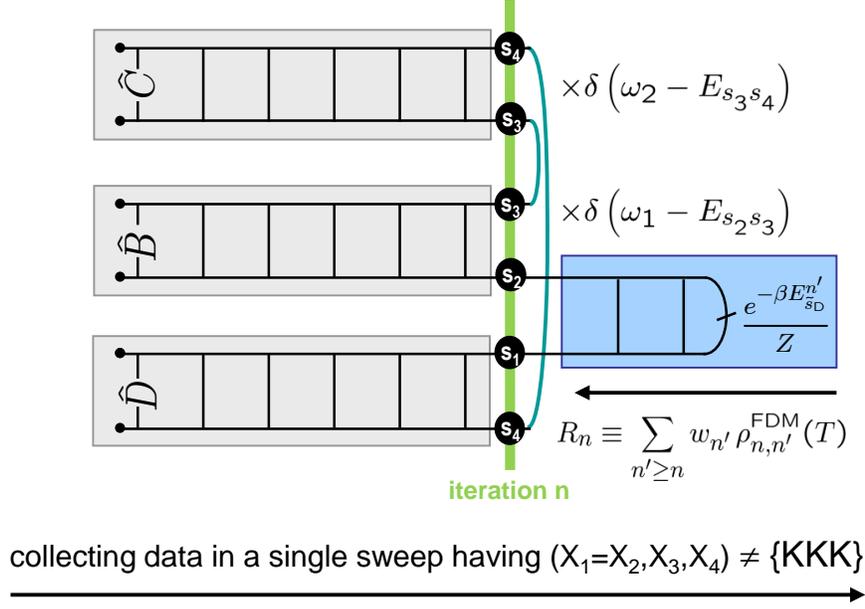
Figure 4.6: MPS diagram for the evaluation of a three-point correlation functions as in Eq. (4.31).

circumvented only by using two separate NRG runs, one for the initial and and one for the final Hamiltonian. With the NRG spectra typically collected in logarithmically spaced bins, having two NRG runs then, it is important that the data is collected in terms of the frequencies $\nu \equiv \omega - \omega_{\text{thr}}$ taken *relative to the threshold frequency* ω_{thr}.

4.3.5 Higher-order correlation functions

Consider the three-point correlation function

$$
\begin{aligned}
G_{BCD}(t_1, t_2) &= \langle \hat{B}(t_1)\hat{C}(t_2)\hat{D} \rangle_T \\
&= \text{tr}\left(\hat{\rho}(T) \cdot e^{i\hat{H}t_1}\hat{B}e^{i\hat{H}(t_2-t_1)}\hat{C}e^{-i\hat{H}t_2}\hat{D} \right)
\end{aligned}
\tag{4.30}
$$

Within the NRG energy eigenbasis, this acquires four independent sums over shells, (1) from the FDM, (2) from performing the overall outer trace, (3) and (4) by inserting an identity after $\hat{B}$ and $\hat{C}$, respectively. All four sums can be combined into a single sum over Wilson shells, by also including the kept state spaces, yet excluding $(X_2, X_3, X_4) \neq$ KKK (note that X_1 and X_2 must be the same, as they directly act onto each other). In frequency space, this implies

$$
G_{BCD}(\omega_1, \omega_2) = \sum_{n}{}' \left[D^{X_4 X_2}_{s_4 s_1} R^{X_2}_{s_1 s_2} \right] B^{X_2 X_3}_{s_2 s_3} C^{X_3 X_4}_{s_3 s_4} \, \delta(\omega_1 - E_{s_2,s_3}) \, \delta(\omega_2 - E_{s_3,s_4}),
\tag{4.31}
$$

where the prime next to the sum again indicates $(X_2, X_3, X_4) \neq \mathrm{KKK}$. The corresponding MPS diagram is shown in Fig. 4.6. The more challenging part with Eq. (4.31) is the dependence on two frequencies. So the corresponding full collection of data into bins (ω_1, ω_2) can become expensive. Certain fixed frequency points together with different kernels corresponding to a different analytic structure of the higher-order correlation function other than the δ-functions in Eq. (4.31), however, appear feasible with reasonable effort.

4.4   Fermionic signs

The NRG is typically applied to fermionic systems (while nevertheless also extensions to bosonic applications exist[11,12,29]). Through its iterative prescription, the resulting MPS has a specific natural fermionic order in Fock space,

$$|s\rangle_n = \sum_{\sigma_d, \sigma_0, \sigma_1, \ldots, \sigma_n} \left(A^{[\sigma_n]} \cdot \ldots \cdot A^{[\sigma_1]} A^{[\sigma_0]} A^{[\sigma_d]} \right)_s \cdot \underbrace{|\sigma_n\rangle \ldots |\sigma_1\rangle |\sigma_0\rangle |\sigma_d\rangle}_{\equiv |\sigma_n, \ldots, \sigma_0, \sigma_d\rangle}, \qquad (4.32)$$

where $|\sigma_d\rangle$ stands for the local state space of the impurity. Site $n' > n$ is added after site n, hence the state space $|\sigma_{n'}\rangle$ appears to the left $|\sigma_n\rangle$. The environmental states $|e\rangle_n$ *w.r.t.* to iteration n which refers to the sites $n' > n$ is irrelevant for the following discussion, and hence will be skipped.

Let $\hat{c}$ be a fermionic operator that acts onto the impurity. Here $\hat{c}$ is assumed an arbitrary operator that nevertheless creates or destroys an odd number of fermionic particles such that fermionic signs apply. A very frequent task then is to represent this operator in the effective many-body-basis at iteration n, *i.e.* to calculate the matrix elements $C_{ss'}^n \equiv {}_n\langle s|\hat{c}|s'\rangle_n$. This involves the basic matrix-element,

$$\langle \sigma_n, \ldots, \sigma_0, \sigma_d|\hat{c}|\sigma_n', \ldots, \sigma_0', \sigma_d'\rangle = \left[\prod_{i=n,\ldots,0} \underbrace{\left(\delta_{\sigma_i, \sigma_i'}(-1)^{n_{\sigma_i'}} \right)}_{\equiv (\hat{z}_i)_{\sigma_i, \sigma_i'}} \right] \cdot \langle \sigma_d|\hat{c}|\sigma_d'\rangle, \qquad (4.33)$$

with $\hat{z} \equiv (-1)^{\hat{n}} = \exp(i\pi\hat{n})$. That is, by pulling the operator $\hat{c}$ acting on the impurity to the right past the second quantization operators that create the states σ_{n_i}, fermionic signs apply, resulting in a *Jordan-Wigner string*

$$\hat{Z} \equiv \bigotimes_{i=0,\ldots,n} \hat{z}_i, \qquad (4.34)$$

to be called *z-string* in short (since $\hat{z}$ is akin to the Pauli z-matrix). Note that through the Jordan-Wigner transformation, which maps fermions onto spins and vice versa, exactly the same string operator as in Eq. (4.33) emerges. For a one-dimensional system with nearest neighbor hopping, the Jordan-Wigner transformation to spins allows to eliminate on the level of the Hamiltonian further complications with fermionic signs. This is fully equivalent, of course, to the explicit treatment of the Jordan-Wigner string in a numerical
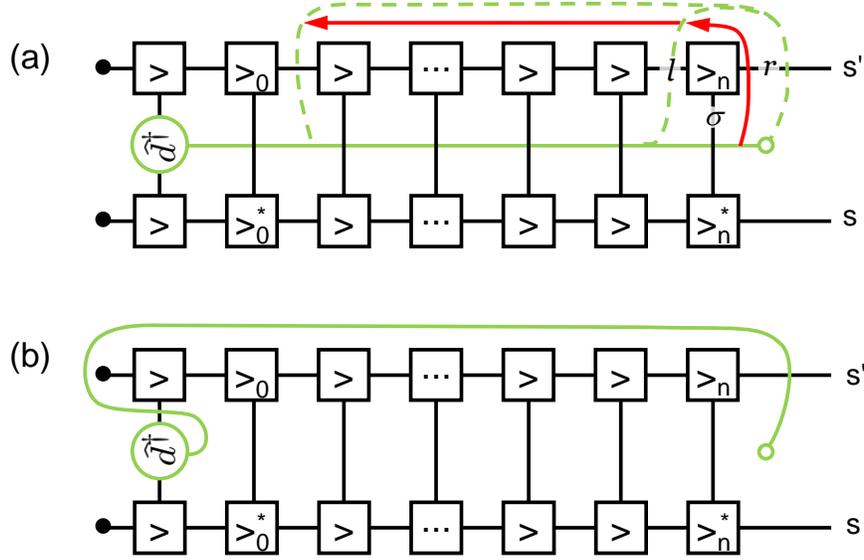
Figure 4.7: MPS diagrams and fermionic signs. Consider the matrix elements of a local operator $\hat{d}^\dagger$ which creates a particle at the impurity. Given the MPS states $|s\rangle_n$ and $|s'\rangle_n$, a z-string (Jordan-Wigner string) $\hat{Z} = \bigotimes_i \hat{z}_i$ arises (light solid horizontal line). The endpoints (open circles) indicate the range of the z-string, *i.e.* starting from and including site 0 to site n. For every crossing of the z-string with a black line, which represent state spaces, fermionic signs apply. Panel (a) shows that a z-string can be rerouted (light dashed lines, pushed in the direction of the red arrow). The final resulting configuration shown in panel (b) shows that by rerouting the z-string significantly fewer crossings with black lines can be achieved. In particular, the z-strings which applied to all sites to the right of $\hat{d}^\dagger$, can be significantly reduced to local fermionic signs at the impurity and another fermionic sign with the state space s'.

setting that keeps a fermionic basis. The operators $\hat{z}_i$ in Eq. (4.33) take care of the book keeping of fermionic signs, by inserting -1 $(+1)$ for all states σ_i at site i with odd (even) number of particles n_{σ_i}. The operators $\hat{z}_i$ are diagonal and hence commute with each other. In the case of additional explicit spin-degrees of freedom, such as the localized spin in the Kondo model, its z-operator is proportional to the identity matrix and hence can be safely ignored.

The following viewpoints highlight three alternative equivalent ways of dealing with fermionic signs in the MPS setup of the NRG. To be specific, the following discussion assumes $\hat{c} = \hat{d}^\dagger$ which creates a particle at the impurity. As such, it generates a Jordan Wigner string for all sites added later to the MPS, *i.e.* sites $i = 0, \ldots, n$ [*cf.* Eq. (4.34)].

Viewpoint 1: rerouting of z-string in tensor network

Figure 4.7 depicts an MPS diagram for the typical evaluation of matrix elements with relevant fermionic signs. The individual boxes (A-tensors) derive from the iterative state space generation of the NRG. This basis generation is obtained in a prior single left-to-right sweep, with all A-tensors orthonormalized left-to-right, as indicated by the '>' in each block (*cf.* Fig. 2.3). All black lines correspond to state space indices. Hence by keeping track of the total number n of particles for all indices, for some specific index a the fermionic sign is given by $(-1)^{n_a}$.

The z-string (depicted by the light green line), that is required for the evaluation of the matrix elements of $d^\dagger$, stretches across all local state spaces σ_i with $0 \leq i \leq n$. Here the interpretation is such, that a *crossing* of the z-string with a state space inserts fermionic signs for this state space.[13] Consider then, for example, the upper right A-tensor in Fig. 4.7. For simplicity, its three legs are labeled l (state space from previous iteration), σ (new local state space), and r (specific combinations of the tensor product of state spaces l and σ, yet with well-defined total particle number). Hence it must hold, $n_l + n_\sigma = n_r$. The index σ is *crossed* by z-string, hence fermionic signs apply,

$$z_\sigma \equiv (-1)^{n_\sigma} = (-1)^{n_r} \underbrace{(-1)^{-n_l}}_{=(-1)^{+n_l}} \equiv z_l z_r. \tag{4.35}$$

Therefore, instead of applying fermionic signs with index σ, it is equally correct to apply fermionic signs with the indices l and r. This allows to *reroute* the z-string[5,13,35] as indicated in Fig. 4.7 (dashed line to the upper right, with the shift in the z-string indicated by short red arrow). Note that for this rerouting to work, the actual left-to-right orthonormalization is not strictly required, and could be relaxed, in general, to the more general condition $n_l \pm n_r \pm n_\sigma =$ even. In particular, this includes $n_l \pm n_r \pm n_\sigma = 0$, which suggests that *any* direction of orthonormalization is acceptable, together with a generic *current site* that combines all (effective) state spaces to an even number of particles, *i.e.* $n_l + n_r + n_\sigma = n_{\text{tot}} =$ even (for $n_{\text{tot}} =$ odd, a global minus sign would apply).

The basic rerouting step as indicated above can be repeated, such that the z-string can be pulled to the top outside the MPS diagram in Fig. 4.7(a), with the final configuration shown in Fig. 4.7(b). The state to the very left (black dot) is the vacuum states with no particles, hence the z-string can also be pulled outside to the left. Thus two crossings of the z-string with state spaces (black lines) remain: one crossing with the local state space at the impurity itself, leading to

$$\hat{d}^\dagger \to \hat{d}^\dagger \hat{z}_d, \tag{4.36}$$

which fully acts within the state space of the impurity, and another crossing with the state space $|s'\rangle_n$ at iteration n.

In typical applications which include thermal expectation values or correlation functions, however, an operator $\hat{d}^\dagger$ never comes alone, as its expectation value with respect to any state with well-defined particle number would be zero. Therefore creation and annihilation
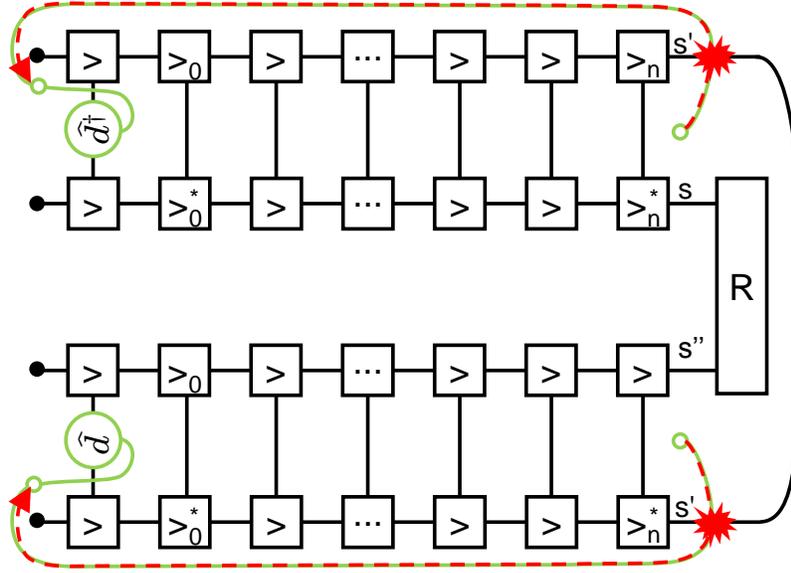
Figure 4.8: Example: fermionic signs in correlation functions. Two MPS diagrams as in Fig. 4.8 for the matrix elements of $\hat{d}$ and $\hat{d}^\dagger$ are combined, as required, for example, for the calculation of correlation functions. The resulting product of matrix elements $_n\langle s'|\hat{d}|s''\rangle_n \cdot R^{[n]}_{s'',s} \cdot {_n}\langle s|\hat{d}^\dagger|s'\rangle_n$ leads to cancelation of the fermionic signs in the index s' in the rerouted z-strings (light green lines), as indicated by the two splashes to the right. Hence the right end-point of the z-string can be retracted to the very left of the system, as indicated by the dashed red arrows. The partial contribution R to the FDM is a scalar operator, such that assuming charge conservation, the particle number of the states s and s'' also must be same. Hence the z-string in Fig. 4.8 could have been equally well also rerouted downwards, instead. The respective fermionic signs with states s and s'' still would have canceled, while the order of application of the z-operator with the impurity would have changed.

operators always appear in pairs. For the local spectral function, for example, $\hat{d}^\dagger$ is paired with its daggered version $\hat{d}$. In their overall combination, the fermionic signs *w.r.t.* the index s' appear *twice* and hence disappear. This situation is sketched in Fig. 4.8. The matrix element discussed previously with Fig. 4.7 is shown at the top of the figure. Given the case of spectral functions (*cf.* Fig. 4.3), its counterpart is shown at the bottom. The contribution R to the reduced density matrix is a scalar operator, such that the particle number of the states s and s'' must match. Similarly, the outer two states are connected through the overall trace (black solid line to the very right), hence even correspond to the same state. Consequently, the same fermionic sign factor applies twice with the rerouted z-strings, which thus cancels, *i.e.* $[(-1)^{n_s}]^2 = 1$ (indicated by the two splashes with s' at the right). Consequently, the right end-point of the z-strings can be retracted along the rerouted z-string all to the way to the left of the impurity (indicated by the red dashed

arrow). Given the A-tensors for the basis transformations from a prior NRG run that only generates the basis, above line of argument therefore allows to ignore fermionic signs for most of the subsequent calculation of thermodynamic quantities or spectral properties. In given example, it is sufficient to calculate the spectral functions for the operator $\hat{d} \rightarrow \hat{z}_d \hat{d}$ and fully ignore fermionic signs for the rest of the chain. This is in contrast to the original setup where the full z-string had to be included, stretching from the impurity across all sites $0 \leq n' \leq n$.

Viewpoint 2: Operator representation

An alternative way to demonstrate the effect of rerouting of the z-string can be given by looking at the equivalent (numerical) representation of operators in the full many-body Hilbert space, *i.e.* without making reference to MPS notation. Given the fermionic order of sites as in Eq. (4.32), a fermionic operator $\hat{f}_k$ that destroys a particle at site $k < n$, has the tensor-product form

$$\hat{F}_k \equiv \hat{1}_d \otimes \hat{1}_0 \otimes \ldots \hat{1}_{k-1} \otimes \hat{f}_k \otimes \hat{z}_{k+1} \otimes \ldots \otimes \hat{z}_n, \tag{4.37}$$

where $\hat{1}_i$ is the identity matrix at site i, $\hat{f}_k$ the the desired operator acting within the state space of site k, and $\hat{z}_i \equiv (-1)^{\hat{n}_i}$ the z-operator taking care of fermionic signs. Now, applying a z-operator to the states s' at the last site n is equivalent to applying a z-operator to each individual site,

$$\begin{aligned}\hat{Z}\hat{F}_k &\equiv \left(\bigotimes_i \hat{z}_i\right)\hat{f}_k \\ &= \hat{z}_d \otimes \hat{z}_0 \otimes \ldots \hat{z}_{k-1} \otimes [\hat{z}\hat{f}]_k \otimes \hat{1}_{k+1} \ldots \hat{1}_n,\end{aligned} \tag{4.38}$$

since $(\hat{z}_i)^2 = \hat{1}_i$. In the application to thermodynamic quantities such as correlations functions, the operator $\hat{F}_k$ would again appear together with its daggered version $\hat{F}_k^\dagger$, hence insertion of $\hat{Z}^2$ has no effect, yet can be split in equal parts, *i.e.* $\hat{F}_k^\dagger \hat{F}_k = (\hat{Z}\hat{F}_k)^\dagger(\hat{Z}\hat{F}_k)$. Therefore, $\hat{Z}\hat{F}_k$ can be equally well used instead of $\hat{F}_k$. As a result, similar to Fig. 4.8, the z-strings have again been flipped from the sites to the right of site k to the left of site k, with the additional transformation $\hat{f}_k \rightarrow [\hat{z}\hat{f}]_k$.

Viewpoint 3: Auxiliary fermionic level

In the case of absorption spectra, the absorption of a photon creates an electron-hole pair, $\hat{h}^\dagger \hat{d}^\dagger$, where the hole $\hat{h}^\dagger$ can be simply treated as a spectator in the dynamics. Nevertheless, by explicitly including the hole in the correlation function, *i.e.* by using the operator $\hat{d}^\dagger \rightarrow \hat{h}^\dagger \hat{d}^\dagger$, this operator itself already forms a pair of fermions that preserves particle number (assuming that $\hat{h}^\dagger$ creates a hole). Therefore, by construction, $\hat{h}^\dagger \hat{d}^\dagger$ simply commutes with all Wilson sites except for the impurity upon which it acts.

The same argument can be repeated for a standard spectral function, by introducing an auxiliary fermionic level $\hat{h}$ that does not participate in the dynamics, *i.e.* does not appear

in the Hamiltonian. In general, *prepending* the states in Eq. (4.32) by the states $|\sigma_h\rangle$ of the "hole", *i.e.*

$$|\sigma_n, \ldots, \sigma_0, \sigma_d\rangle \rightarrow |\sigma_n, \ldots, \sigma_0, \sigma_d\rangle|\sigma_h\rangle, \tag{4.39}$$

immediately results in the same consistent picture as already encountered with Fig. 4.8 or Eq. (4.38).

Bibliography

[1] Ian Affleck, Tom Kennedy, Elliott H. Lieb, and Hal Tasaki. Valence bond ground states in isotropic quantum antiferromagnets. *Commun. Math. Phys.*, 115:477, 1988.

[2] F. B. Anders and A. Schiller. Real-time dynamics in quantum impurity systems: A time-dependent numerical renormalization group approach. *Phys. Rev. Lett.*, 95: 196801, 2005.

[3] P. W. Anderson. Localized magnetic states in metals. *Phys. Rev.*, 124(1):41–53, Oct 1961.

[4] P. W. Anderson. Infrared catastrophe in fermi gases with local scattering potentials. *Phys. Rev. Lett.*, 18(24):1049–1051, Jun 1967.

[5] Thomas Barthel, Carlos Pineda, and Jens Eisert. Contraction of fermionic operator circuits and the simulation of strongly correlated fermions. *Phys. Rev. A*, 80(4):042333, Oct 2009.

[6] B. Bauer, P. Corboz, R. Orús, and M. Troyer. Implementing global abelian symmetries in projected entangled-pair state algorithms. *Phys. Rev. B*, 83:125106, Mar 2011.

[7] Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev. D*, 7(8):2333–2346, Apr 1973.

[8] D. Bohr, P. Schmitteckert, and P. Wölfle. Dmrg evaluation of the kubo formula – conductance of strongly interacting quantum systems. *EPL (Europhysics Letters)*, 73 (2):246, 2006.

[9] Dan Bohr and Peter Schmitteckert. Strong enhancement of transport by interaction on contact links. *Phys. Rev. B*, 75:241103, Jun 2007.

[10] R. Bulla, A. C. Hewson, and T. Pruschke. Numerical renormalization group calculations for the self-energy of the impurity anderson model. *J. Phys.: Condens. Matter*, 10:8365, 1998.

[11] Ralf Bulla, Ning-Hua Tong, and Matthias Vojta. Numerical renormalization group for bosonic systems and application to the sub-ohmic spin-boson model. *Phys. Rev. Lett.*, 91:170601, 2003.

[12] Ralf Bulla, Theo Costi, and Thomas Pruschke. Numerical renormalization group method for quantum impurity systems. *Rev. Mod. Phys.*, 80:395, 2008.

[13] Philippe Corboz, Román Orús, Bela Bauer, and Guifré Vidal. Simulation of strongly correlated fermions in two spatial dimensions with fermionic projected entangled-pair states. *Phys. Rev. B*, 81(16):165104, Apr 2010.

[14] T. A. Costi. Renormalization-group approach to nonequilibrium green functions in correlated impurity systems. *Phys. Rev. B*, 55(5):3003–3009, Feb 1997.

[15] T. A. Costi and A. C. Hewson. Resistivity cross-over for the non-degenerate anderson model. *Philosophical Magazine Part B*, 65(6):1165–1170, 1992.

[16] T. A. Costi, L. Bergqvist, A. Weichselbaum, J. von Delft, T. Micklitz, A. Rosch, P. Mavropoulos, P. H. Dederichs, F. Mallet, L. Saminadayar, and C. Bäuerle. Kondo decoherence: Finding the right spin model for iron impurities in gold and silver. *Phys. Rev. Lett.*, 102(5):056802, 2009.

[17] A. J Daley, C. Kollath, U. Schollwöck, and G. Vidal. Time-dependent density-matrix renormalization-group using adaptive effective hilbert spaces. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(04):P04005, 2004.

[18] A. J. Daley, J. M. Taylor, S. Diehl, M. Baranov, and P. Zoller. Atomic three-body loss as a dynamical three-body interaction. *Phys. Rev. Lett.*, 102(4):040402, 2009.

[19] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose-einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.*, 75(22):3969–3973, Nov 1995.

[20] W. J. de Haas, J. de Boer, and G. J. van dën Berg. The electrical resistance of gold, copper and lead at low temperatures. *Physica*, 1(7-12):1115–1124, 1934.

[21] Luis G. G. V. Dias da Silva, F. Heidrich-Meisner, A. E. Feiguin, C. A. Büsser, G. B. Martins, E. V. Anda, and E. Dagotto. Transport properties and kondo correlations in nanostructures: Time-dependent dmrg method applied to quantum dots coupled to wilson chains. *Phys. Rev. B*, 78(19):195317, Nov 2008.

[22] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82(1):277–306, Feb 2010.

[23] J. Friedel. Electronic structure of primary solid solutions in metals. *Advances in Physics*, 3(12):446–507, 1954.

[24] J. Friedel. Metallic alloys. *Nuovo Cimento Suppl.*, 7(10):287, 1958.

[25] Antoine Georges, Gabriel Kotliar, Werner Krauth, and Marcelo J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68(1):13, Jan 1996.

[26] Stefano Giorgini, Lev P. Pitaevskii, and Sandro Stringari. Theory of ultracold atomic fermi gases. *Reviews of Modern Physics*, 80(4):1215, 2008.

[27] D. Goldhaber-Gordon, Hadas Shtrikman, D. Mahalu, David Abusch-Magder, U. Meirav, and M. A. Kastner. Kondo effect in a single-electron transistor. *Nature*, 391:156–159, 1998.

[28] Markus Greiner, Olaf Mandel, Tilman Esslinger, Theodor W. Hansch, and Immanuel Bloch. Quantum phase transition from a superfluid to a mott insulator in a gas of ultracold atoms. *Nature*, 415:39–44, 2002.

[29] Cheng Guo, Andreas Weichselbaum, Jan von Delft, and Matthias Vojta. Critical and strong-coupling phases in one- and two-bath spin-boson models. *Phys. Rev. Lett.*, 108:160401, Apr 2012.

[30] R. W. Helmes, M. Sindel, L. Borda, and J. von Delft. Absorption and emission in quantum dots: Fermi surface effects of anderson excitons. *Phys. Rev. B*, 72(12):125301, Sep 2005.

[31] Walter Hofstetter. Generalized numerical renormalization group for dynamical quantities. *Phys. Rev. Lett.*, 85(7):1508–1511, Aug 2000.

[32] Andreas Holzner, Andreas Weichselbaum, Ian P. McCulloch, Ulrich Schollwöck, and Jan von Delft. Chebyshev matrix product state approach for spectral functions. *Phys. Rev. B*, 83(19):195115, May 2011.

[33] Jun Kondo. Resistance minimum in dilute magnetic alloys. *Progress of Theoretical Physics*, 32:37–49, 1964.

[34] Robert König, Ben W. Reichardt, and Guifré Vidal. Exact entanglement renormalization for string-net models. *Phys. Rev. B*, 79:195123, May 2009.

[35] Christina V. Kraus, Norbert Schuch, Frank Verstraete, and J. Ignacio Cirac. Fermionic projected entangled pair states. *Phys. Rev. A*, 81(5):052338, May 2010.

[36] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson. Renormalization-group approach to the anderson model of dilute magnetic alloys. i. static properties for the symmetric case. *Phys. Rev. B*, 21(3):1003–1043, Feb 1980.

[37] L. Lamata, J. León, D. Pérez-García, D. Salgado, and E. Solano. Sequential implementation of global quantum operations. *Phys. Rev. Lett.*, 101(18):180506, Oct 2008.

[38] Christian Latta, Florian Haupt, Markus Hanl, Andreas Weichselbaum, Martin Claassen, Wolf Wuester, Parisa Fallahi, Stefan Faelt, Leonid Glazman, Jan von Delft, Hakan E. Türeci, and Atac Imamoglu. Quantum quench of kondo correlations in optical absorption. *Nature*, 474:627–630, 05 2011.

[39] I. P. McCulloch and M. Gulcsi. The non-abelian density matrix renormalization group algorithm. *EPL (Europhysics Letters)*, 57(6):852, 2002.

[40] Ian P. McCulloch. From density-matrix renormalization group to matrix product states. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(10):P10014, 2007.

[41] Walter Metzner and Dieter Vollhardt. Correlated lattice fermions in $d = \infty$ dimensions. *Phys. Rev. Lett.*, 62(3):324–327, Jan 1989.

[42] W. Münder, A. Weichselbaum, M. Goldstein, Y. Gefen, and J. von Delft. Anderson orthogonality in the dynamics after a local quantum quench. *arXiv:1108.5539v2 [cond-mat.str-el]*, August 2011. URL http://xxx.lanl.gov/abs/1108.5539.

[43] V. Murg, F. Verstraete, and J. I. Cirac. Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states. *Phys. Rev. A*, 75(3):033605, Mar 2007.

[44] L. N. Oliveira. The numerical renormalization group and the problem of impurities in metals. *Braz. J. Phys*, 22:155, 1992.

[45] Luiz N. Oliveira and John W. Wilkins. New approach to the x-ray-absorption problem. *Phys. Rev. B*, 24:4863–4866, Oct 1981.

[46] Jiwoong Park, Abhay N. Pasupathy, Jonas I. Goldsmith, Connie Chang, Yuval Yaish, Jason R. Petta, Marie Rinkoski, James P. Sethna, Hector D. Abruna, Paul L. McEuen, and Daniel C. Ralph. Coulomb blockade and the kondo effect in single-atom transistors. *Nature*, 417:722–725, 2002.

[47] David Perez-Garcia, Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac. Matrix product state representations. *Quantum Inf. Comput.*, 7:401, 2007.

[48] Robert Peters, Thomas Pruschke, and Frithjof B. Anders. Numerical renormalization group approach to green's functions for quantum impurity models. *Phys. Rev. B*, 74 (24):245114, 2006.

[49] Carsten Raas, Götz S. Uhrig, and Frithjof B. Anders. High-energy dynamics of the single-impurity anderson model. *Phys. Rev. B*, 69(4):041102, Jan 2004.

[50] Stefan Rommer and Stellan Östlund. Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev. B*, 55:2164, 1997.

[51] Hamed Saberi, Andreas Weichselbaum, and Jan von Delft. Matrix-product-state comparison of the numerical renormalization group and the variational formulation of the density-matrix renormalization group. *Phys. Rev. B*, 78(3):035124, 2008.

[52] Hamed Saberi, Andreas Weichselbaum, Lucas Lamata, David Perez-Garcia, Jan von Delft, and Enrique Solano. Constrained optimization of sequentially generated entangled multiqubit states. *Phys. Rev. A*, 80(2):022334, 2009.

[53] Osamu Sakai, Yukihiro Shimizu, and Tadao Kasuya. Single-particle and magnetic excitation spectra of degenerate anderson model with finite f-f coulomb interaction. *J. Phys. Soc. Japan*, 58:3666, 1989.

[54] Jun John Sakurai. *Modern Quantum Mechanics*. Addison-Wesley Publishing Company, Reading, MA, Revised edition, 1994.

[55] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77(1): 259–315, Apr 2005.

[56] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.*, 326:96–192, 2011.

[57] J. R. Schrieffer and P. A. Wolff. Relation between the anderson and kondo hamiltonians. *Phys. Rev.*, 149(2):491–492, Sep 1966.

[58] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac. Entropy scaling and simulability by matrix product states. *Phys. Rev. Lett.*, 100:030504, Jan 2008.

[59] Sukhwinder Singh, Robert N. C. Pfeifer, and Guifré Vidal. Tensor network decompositions in the presence of a global symmetry. *Phys. Rev. A*, 82(5):050301, Nov 2010.

[60] Sukhwinder Singh, Robert N. C. Pfeifer, and Guifre Vidal. Tensor network states and algorithms in the presence of a global u(1) symmetry. *Phys. Rev. B*, 83:115125, Mar 2011.

[61] Hakan E. Türeci, M. Hanl, M. Claassen, A. Weichselbaum, T. Hecht, B. Braunecker, A. Govorov, L. Glazman, A. Imamoglu, and J. von Delft. Many-body dynamics of exciton creation in a quantum dot by optical absorption: A quantum quench towards kondo correlations. *Phys. Rev. Lett.*, 106(10):107402, Mar 2011.

[62] F. Verstraete, A. Weichselbaum, U. Schollwock, J. I. Cirac, and Jan von Delft. Variational matrix-product-state approach to quantum impurity models. *arXiv:0504305v2 [cond-mat]*, 2005.

[63] F. Verstraete, V. Murg, and J.I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 2008.

[64] G. Vidal. Entanglement renormalization. *Phys. Rev. Lett.*, 99:220405, Nov 2007.

[65] A. Weichselbaum, F. Verstraete, U. Schollwock, J. I. Cirac, and Jan von Delft. Variational matrix-product-state approach to quantum impurity models. *Phys. Rev. B*, 80 (16):165117, 2009.

[66] Andreas Weichselbaum. Discarded weight and entanglement spectra in the numerical renormalization group. *Phys. Rev. B*, 84:125130, Sep 2011.

[67] Andreas Weichselbaum. General framework for non-abelian symmetries in tensor networks. *arXiv:1202.5664v1 [cond-mat.str-el]*, February 2012.

[68] Andreas Weichselbaum and Jan von Delft. Sum-rule conserving spectral functions from the numerical renormalization group. *Phys. Rev. Lett.*, 99(7):076402, 2007.

[69] Andreas Weichselbaum, Wolfgang Münder, and Jan von Delft. Anderson orthogonality and the numerical renormalization group. *Phys. Rev. B*, 84(7):075137, Aug 2011.

[70] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863–2866, Nov 1992.

[71] Steven R. White and Adrian E. Feiguin. Real-time evolution using the density matrix renormalization group. *Phys. Rev. Lett.*, 93:076401, 2004.

[72] Kenneth G. Wilson. The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47(4):773–840, Oct 1975.

[73] K. Winkler, G. Thalhammer, F. Lang, R. Grimm, J. Hecker Denschlag, A. J. Daley, A. Kantian, H. P. Büchler, and P. Zoller. Repulsively bound atom pairs in an optical lattice. *Nature*, 441:853–856, 2006.

[74] Michael M. Wolf, Frank Verstraete, Matthew B. Hastings, and J. Ignacio Cirac. Area laws in quantum systems: Mutual information and correlations. *Phys. Rev. Lett.*, 100 (7):070502, Feb 2008.

[75] Gergely Zaránd and Jan von Delft. Analytical calculation of the finite-size crossover spectrum of the anisotropic two-channel kondo model. *Phys. Rev. B*, 61:6918–6933, Mar 2000.

[76] Rok Žitko and Thomas Pruschke. Energy resolution and discretization artifacts in the numerical renormalization group. *Phys. Rev. B*, 79(8):085106, 2009.

Appendix P

Selected Publications

In this appendix, a few of the papers highlighted in Sec. 1.1 are reprinted, preceded a short introduction for each. The sections are numbered by Pn, indicating paper number n. The thumb marks **Pn** are intended as a visual guide to separate the individual papers. Similarly, references such as Eq. (Pn-m) refer to equation (m) of the subsequently included paper n.

P1. Sum-Rule Conserving Spectral Functions from the NRG

▶ Andreas Weichselbaum, and Jan von Delft, Phys. Rev. Lett. **99**, 076402 (2007)

Inspired by the treatment of real-time evolution of quantum quenches through the complete basis sets by Anders and Schiller (2005), the similarity with the DM-NRG by Hofstetter (2000) had been stunning. In particular, the complete basis sets naturally generated the reduced density matrices for the environment $n' > n$ for a given iteration n. It was soon clear then within the MPS framework, how complete basis sets also can be related to the calculation of spectral functions. This has lead to an independent publication Peters et al. (2006) which, however, still treats temperature in a single shell-approximation. Importantly, however, the complete basis sets also provide a natural definition for the full density matrix (FDM) in Eq. (P1-8) for arbitrary but fixed temperature.[1] This leads to a multi-shell calculation within the NRG, in that the full density matrix is naturally built from a range of Wilson shells around the energy scale of temperature (*cf.* Sec. 4.2). Overall then, this results in a black-box algorithm without any further requirement for specific patching schemes prevalent until then.

With my codes for abelian symmetries fully developed by then, the setup of the FDM-NRG was straightforward and finished within a few weeks. Using the discrete FDM-NRG

---

[1]By definition of a thermal density matrix, all energies in the exponent necessarily must be taken on the same footing, *i.e.* taken in *un*rescaled energy units relative to a *common* energy reference. This is obvious, of course, but given the usual procedure of rescaling of energies within the NRG, this has initially led to some confusion in the community.

data for spectral functions, one of the first observation was, that spectral sum rules are fulfilled exactly up to numerical double precision noise (10^{-16}). This clearly indicated, that from a numerical point of view (i) the insertion of complete basis sets was setup correctly. More generally, this implies that (ii) despite having exponentially large quantum-many-body spaces, due to the quasi-one-dimensional structure of the problem, complete basis sets can be dealt with exactly in an efficient systematic manner. Note however, that spectral sum rules are fulfilled by construction, irrespective of the accuracy of the calculation. Therefore the latter still needs to be checked independently. Nevertheless, exact sum rules are an important prerequisite, for example, for dynamic mean field calculations (DMFT).[48]

The price to pay for the black-box algorithm of the FDM-NRG is, that the discarded state spaces must be explicitly included in the calculation of spectral properties, while conventional patching schemes only operate within kept spaces. Finally, it may be argued that the discarded state spaces are somewhat more strongly affected by the sharp truncation of the high-energy states. Hence this may affect higher-moment spectral sum rules using FDM-NRG,[76] the detailed analysis of which is left for future studies. With respect to typical calculations within the NRG, however, FDM-NRG turned out a very valuable, systematic, and accurate approach to the simulation of quantum impurity models at arbitrary temperatures. Moreover, due to its clean derivation, also the spectral properties for frequencies comparable or smaller than temperature show clearly improved behavior.

P1.

Sum-Rule Conserving Spectral Functions from the Numerical Renormalization Group

Andreas Weichselbaum and Jan von Delft

Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience,
Ludwig-Maximilians-Universität München, D-80333 München, Germany

We show how spectral functions for quantum impurity models can be calculated very accurately using a complete set of discarded numerical renormalization group eigenstates, recently introduced by Anders and Schiller. The only approximation is to judiciously exploit energy scale separation. Our derivation avoids both the overcounting ambiguities and the single-shell approximation for the equilibrium density matrix prevalent in current methods, ensuring that relevant sum rules hold rigorously and spectral features at energies below the temperature can be described accurately.

Quantum impurity models describe a quantum system with a small number of discrete states, the "impurity," coupled to a continuous bath of fermionic or bosonic excitations. Such models are relevant for describing transport through quantum dots, for the treatment of correlated lattice models using dynamical mean field theory, or for the modeling of the decoherence of qubits.

The impurity's dynamics in thermal equilibrium can be characterized by spectral functions of the type $\mathcal{A}^{\mathcal{BC}}(\omega) = \int \frac{dt}{2\pi} e^{i\omega t} \langle \hat{\mathcal{B}}(t)\hat{\mathcal{C}} \rangle_T$. Their Lehmann representation reads

$$\mathcal{A}^{\mathcal{BC}}(\omega) = \sum_{a,b} \langle b|\hat{\mathcal{C}}|a\rangle \frac{e^{-\beta E_a}}{Z} \langle a|\hat{\mathcal{B}}|b\rangle \delta(\omega - E_{ba}), \quad (1)$$

with $Z = \sum_a e^{-\beta E_a}$ and $E_{ba} = E_b - E_a$, which implies the sum rule $\int d\omega \, \mathcal{A}^{\mathcal{BC}}(\omega) = \langle \hat{\mathcal{B}}\hat{\mathcal{C}} \rangle_T$. In this Letter, we describe a strategy for numerically calculating $\mathcal{A}^{\mathcal{BC}}(\omega)$ that, in contrast to previous methods, rigorously satisfies this sum rule and accurately describes both high *and* low frequencies, including $\omega \lesssim T$, which we test by checking our results against exact Fermi-liquid relations.

Our work builds on Wilson's numerical renormalization group (NRG) method [1]. Wilson discretized the environmental spectrum on a logarithmic grid of energies Λ^{-n} (with $\Lambda > 1$, $1 \leq n \leq N \to \infty$), with exponentially high resolution of low-energy excitations, and mapped the impurity model onto a "Wilson tight-binding chain," with hopping matrix elements that decrease exponentially as $\Lambda^{-n/2}$ with site index n. Because of this separation of energy scales, the Hamiltonian can be diagonalized iteratively: adding one site at a time, a new "shell" of eigenstates is constructed from the new site's states and the M_K lowest-lying eigenstates of the previous shell (the so-called "kept" states), while "discarding" the rest.

Subsequent authors [2–10] have shown that spectral functions such as $\mathcal{A}^{\mathcal{BC}}(\omega)$ can be calculated via the Lehmann sum, using NRG states (kept and discarded) of those shells n for which $\omega \sim \Lambda^{-n/2}$. Though plausible on heuristic grounds, this strategy entails double-counting

ambiguities [5] about how to combine data from successive shells. Patching schemes [9] for addressing such ambiguities involve arbitrariness. As a result, the relevant sum rule is not satisfied rigorously, with typical errors of a few percent. Also, the thermal density matrix (DM) $\hat{\rho} = e^{-\beta \hat{H}}/Z$ has until now been represented rather crudely using only the single N_Tth shell for which $T \simeq \Lambda^{-1/2(N_T-1)}$ [8], with a chain of length $N = N_T$, resulting in inaccurate spectral information for $\omega \lesssim T$. In this Letter we avoid these problems by using in the Lehmann sum an approximate but *complete* set of eigenstates, introduced recently by Anders and Schiller (AS) [11].

Wilson's truncation scheme.—The Wilson chain's zeroth site represents the bare impurity Hamiltonian $\hat{h}_0$ with a set of d_0 impurity states $|\sigma_0\rangle$. It is coupled to a fermionic chain, whose nth site ($1 \leq n \leq N$) represents a set of d states $|\sigma_n\rangle$, responsible for providing energy resolution to the spectrum at scale $\Lambda^{-n/2}$. For a spinful fermionic band, for example, $\sigma_n \in \{0, \uparrow, \downarrow, \uparrow\downarrow\}$, hence $d = 4$. (Bosonic chains can be treated similarly [10].) The Hamiltonian $\hat{H} = \hat{H}_N$ for the full chain is constructed iteratively by adding one site at a time, using $\hat{H}_n = \hat{H}_{n-1} + \hat{h}_n$ (acting in a $d^n d_0$-dimensional Fock space $\mathcal{F}_n$ spanned by the basis states $\{|\sigma_n\rangle \otimes \cdots \otimes |\sigma_0\rangle\}$), where $\hat{h}_n$ links sites n and $n-1$ with hopping strength $\sim \Lambda^{-n/2}$. Since the number of eigenstates of $\hat{H}_n$ grows exponentially with n, Wilson proposed the following iterative truncation scheme to numerically diagonalize the Hamiltonian: Let n_0 be the last iteration for which a complete set $\{|s\rangle_{n_0}^K\}$ of kept eigenstates of $\hat{H}_{n_0}$ can be calculated without truncation. For $n > n_0$, construct the orthonormal eigenstates $\{|s\rangle_n^X\}$ of $\hat{H}_n$ (the nth "shell"), with eigenvalues E_s^n, as linear combinations of the kept eigenstates $|s\rangle_{n-1}^K$ of $\hat{H}_{n-1}$ and the states $|\sigma_n\rangle$ of site n,

$$|s'\rangle_n^X = \sum_{\sigma_n s}^K |\sigma_n\rangle \otimes |s\rangle_{n-1}^K [A_{KX}^{[\sigma_n]}]_{ss'}, \quad (2)$$

with coefficients arranged into a matrix $A_{KX}^{[\sigma_n]}$ whose ele-

P1.

ments are labeled by ss'. The superscript $X = K$ or D indicates that the new shell has been partitioned into "kept" states (say the M_K lowest-lying eigenstates of $\hat{H}_n$) to be retained for the next iteration and "discarded" states (the remaining ones). Since $\hat{h}_n$ acts as a weak perturbation (of relative size $\Lambda^{-1/2}$) on $\hat{H}_{n-1}$, the d-fold degeneracy of the states $|\sigma_n\rangle \otimes |s\rangle_{n-1}^X$ is lifted, resulting in a characteristic energy spacing $\Lambda^{-n/2}$ for shell n. Iterating until the spectrum of low-lying eigenvalues has reached a fixed point (for $n = N$, say), one generates a set of eigenstates $\{|s\rangle_n^X\}$ with the structure of matrix product states [12] (Fig. 1). The states generated for the last Nth shell will all be regarded as discarded [11].

Anders-Schiller basis.—Recently, AS have shown [11] that the discarded states can be used to build a complete basis for the whole Wilson chain: the states $\{|s\rangle_n^X\}$ describing the nth shell are supplemented by a set of d^{N-n} *degenerate* "environmental" states $\{|e_n\rangle = |\sigma_N\rangle \otimes \cdots \otimes |\sigma_{n+1}\rangle\}$ spanning the rest of the chain to construct the set of states $\{|se\rangle_n^X \equiv |e_n\rangle \otimes |s\rangle_n^X\}$. These reside in the complete Fock space $\mathcal{F}_N$ of the full chain, spanning $\mathcal{F}_N$ if $n \leq n_0$. Ignoring the degeneracy-lifting effect of the rest of the chain, these states become approximate eigenstates of the Hamiltonian $\hat{H}_N$ of the *full* chain ("NRG approximation"),

$$\hat{H}_N |se\rangle_n^X \simeq E_s^n |se\rangle_n^X, \qquad (3)$$

with eigenenergies *independent* of the (d^{N-n})-fold degenerate environmental index e_n. (This will facilitate tracing out the environment below.) By construction, we have $_m^D\langle se|s'e'\rangle_n^D = \delta_{mn}\delta_{e_n e_n'}\delta_{ss'}$ and

$$_m^K\langle se|s'e'\rangle_n^D = \begin{cases} 0, & m \geq n \\ \delta_{e_n e_n'}[A_{KK}^{\sigma_{m+1}} \ldots A_{KD}^{\sigma_n}]_{ss'}, & m < n. \end{cases} \quad (4)$$

The discarded states of shell n are orthogonal to the discarded states of any other shell, and to the kept states of that or any later shell. Combining the discarded states from all shells thus yields a complete set of NRG eigenstates of $\hat{H}_N$, the "Anders-Schiller basis," that span the full Fock space $\mathcal{F}_N$ ($\sum_n$ henceforth stands for $\sum_{n>n_0}^N$):

$$\mathbf{1}^{(d_0 d^N)} = \sum_{se} |se\rangle_{n_0}^K {}_{n_0}^K\langle se| = \sum_n \sum_{se} |se\rangle_n^{DD}{}_n\langle se|. \quad (5)$$

Local operators.—Let us now consider a "local" operator $\hat{\mathcal{B}}$ acting nontrivially only on sites up to n_0. Two
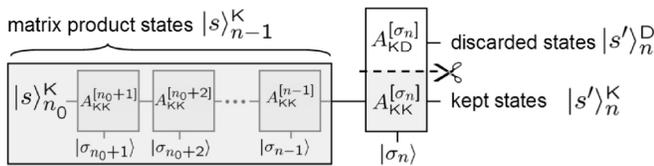

FIG. 1. Diagram for the kept (or discarded) matrix product state $|s'\rangle_n^K$ (or $|s'\rangle_n^D$): the nth box represents the matrix block $A_{KX}^{[\sigma_n]}$, its left, bottom, and right legs carry the labels of the states $|s\rangle_{n-1}^K$, $|\sigma_n\rangle$, and $|s'\rangle_n^K$ (or $|s'\rangle_n^D$), respectively.

particularly useful representations are

$$\hat{\mathcal{B}} = \sum_{ss'e} |se\rangle_{n_0}^K [\mathcal{B}_{KK}^{[n_0]}]_{ss'}{}_{n_0}^K\langle s'e| = \sum_n \sum_{XX'}^{\neq KK} \hat{\mathcal{B}}_{XX'}^{[n]}. \quad (6)$$

The left equality, written $\hat{\mathcal{B}} = \hat{\mathcal{B}}_{KK}^{[n_0]}$ in brief, represents the operator in the complete basis set $\{|se\rangle_{n_0}^K\}$, with matrix elements known exactly numerically (possibly up to fermionic minus signs depending on the environmental states, but these enter quadratically in correlation functions and hence cancel). The right-hand side (RHS) of Eq. (6) expresses $\hat{\mathcal{B}}$ in the AS basis and is obtained as follows: starting from $\hat{\mathcal{B}}_{KK}^{[n_0]}$, one iteratively refines the "kept-kept" part of $\hat{\mathcal{B}}$ from, say, the $(n-1)$th iteration in terms of the NRG eigenstates $\{|se\rangle_n^x\}$ of the next shell, including both kept and discarded states ($X = K, D$),

$$\hat{\mathcal{B}}_{KK}^{[n-1]} = \sum_{XX'} \sum_{ss'e} |se\rangle_n^X [\mathcal{B}_{XX'}^{[n]}]_{ss'}{}_n^{X'}\langle s'e| = \sum_{XX'} \hat{\mathcal{B}}_{XX'}^{[n]}, \quad (7)$$

thereby defining the operators $\hat{\mathcal{B}}_{XX'}^{[n]}$, with matrix elements $[\mathcal{B}_{XX'}^{[n]}]_{ss'} = [A_{XK}^{[\sigma_n]\dagger} \mathcal{B}_{KK}^{[n-1]} A_{KX'}^{[\sigma_n]}]_{ss'}$. Splitting off all $XX' \neq KK$ terms (DD, KD, DK) and iteratively refining each KK term until $n = N$, we obtain the RHS of Eq. (6). It has two important features. First, the matrix elements of the time-dependent operator $\hat{\mathcal{B}}(t) = e^{i\hat{H}t}\hat{\mathcal{B}}e^{-i\hat{H}t}$, evaluated within the NRG approximation, $[\mathcal{B}_n^{XX'}(t)]_{ss'} \simeq [\mathcal{B}_n^{XX'}]_{ss'}e^{it(E_s^n - E_{s'}^n)}$, contain differences of eigenenergies from the *same* shell only, i.e., calculated with the same level of accuracy. Second, by *excluding KK* terms it rigorously avoids the double-counting ambiguities and heuristic patching rules plaguing previous approaches [2–10].

Thermal averages.—To calculate thermal averages $\langle \ldots \rangle_T = \text{Tr}[\hat{\rho} \ldots]$, we write the full density matrix (FDM) $\hat{\rho} = e^{-\beta\hat{H}}/Z$ using the NRG approximation Eq. (3),

$$\hat{\rho} \simeq \sum_n \sum_{se} |se\rangle_n^D \frac{e^{-\beta E_s^n}}{Z} {}_n^D\langle se| = \sum_n w_n \hat{\rho}_{DD}^{[n]}, \quad (8)$$

where $w_n \equiv d^{N-n}Z_n^D/Z$ and $Z_n^D \equiv \sum_s^D e^{-\beta E_s^n}$. The RHS of Eq. (8) expresses $\hat{\rho}$ as sum over $\hat{\rho}_{DD}^{[n]}$, the density matrix for the *discarded* states of shell n, properly normalized as $\text{Tr}[\hat{\rho}_{DD}^{[n]}] = 1$, and entering with relative weight w_n, with $\sum_n w_n = 1$. Similarly, for spectral functions we have

$$\langle \ldots \rangle_T = \sum_n w_n \langle \ldots \rangle_n, \quad \mathcal{A}(\omega) = \sum_n w_n \mathcal{A}_n(\omega), \quad (9)$$

where the averages $\langle \cdots \rangle_n$ and spectral functions $\mathcal{A}_n(\omega)$ are calculated with respect to $\hat{\rho}_{DD}^{[n]}$ of shell n only.

Previous strategies [4–11] for thermal averaging amount to using a "single-shell approximation" $w_n = \delta_{nN_T}$ for the density matrix and terminating the chain at a length $N = N_T$ set by $T \simeq \Lambda^{-1/2(N_T-1)}$. As a result, spectral features on scales $\omega \leq T$, which would require a longer chain, are described less accurately [see Figs. 2(a) and 2(b)]. Our

70

novel approach avoids these problems by using the *full* density matrix (FDM), summed over *all* shells, letting the weighting function w_n select the shells relevant for a given temperature yielding a smooth T dependence [see Fig. 2(c)]. Since w_n has a peak width of five to ten shells depending on Λ, d and M_K and peaks at n values somewhat above N_T [arrow Fig. 2(b)], spectral information from energies well below T is retained.

Let us now consider the spectral function $\mathcal{A}^{\mathcal{BC}}(\omega)$, for local operators $\hat{B}$ and $\hat{C}$. Equations (4), (6), (8), and (9) can be used to evaluate $\langle \hat{B}(t)\hat{C}\rangle_n$. Fourier transforming the result we find (sums over ss' and σ_n implied)

$$\mathcal{A}_n^{\mathcal{BC}}(\omega) = \sum_{m>n_0}^{n} \sum_{XX'}^{\neq KK} [\mathcal{C}_{X'X}^{[m]}\rho_{XX}^{[mn]}]_{s's}[\mathcal{B}_{XX'}^{[m]}]_{ss'}\delta(\omega - E_{s's}^m),$$

$$[\rho_{DD}^{[m=n]}]_{ss'} = \delta_{ss'}\frac{e^{-\beta E_s^n}}{Z_n}, \qquad (10)$$

$$[\rho_{KK}^{[m<n]}]_{ss'} = [A_{KK}^{[\sigma_{m+1}]}\ldots A_{KD}^{[\sigma_n]}\rho_{DD}^{[nn]}A_{DK}^{[\sigma_n]\dagger}\ldots A_{KK}^{[\sigma_{m+1}]\dagger}]_{ss'}.$$

Similarly, the static quantity $\langle \hat{B}\hat{C}\rangle_n$ equals the first line's RHS without the δ function. The matrix elements $[\rho_{XX}^{[mn]}]_{ss'} \equiv \sum_{em}^{X}\langle se|\hat{\rho}_{DD}^{[n]}|s'e\rangle_m^X$ are given by the second and third lines, together with $\rho_{KK}^{[m=n]} = \rho_{DD}^{[m<n]} = 0$. After performing a "forward run" to generate all relevant NRG eigenenergies and matrix elements, $\mathcal{A}^{\mathcal{BC}}(\omega)$ can be calculated in a single "backward run," performing a sum with the structure $\sum_{m>n_0}^{N}[\mathcal{C}\rho^{\text{red}}\mathcal{B}\cdot\delta()]^{[m]}$, starting from $m = N$. Here $\rho_{XX}^{[m],\text{red}} \equiv \sum_{n\geq m}^{N}w_n\rho_{XX}^{[mn]}$ (updated one site at a time during the backward run) is the *full* reduced density matrix for shell m, obtained iteratively by tracing out all shells at smaller scales $\Lambda^{-n/2}$ ($n \geq m$).

Equations (8)–(10) are the main results of our "FDM-NRG" approach. They rigorously generalize Hofstetter's DM-NRG [8] (which leads to similar expressions, but using $w_n = \delta_{nN_T}$ and without excluding KK matrix elements), and provide a concise prescription, free from double-counting ambiguities, for how to combine NRG data from different shells when calculating $\mathcal{A}^{\mathcal{BC}}(\omega)$. The relevant sum rule is satisfied *identically*, since by construction $\int d\omega\, \mathcal{A}_n^{\mathcal{BC}}(\omega) = \langle \hat{B}\hat{C}\rangle_n$ holds for every n and arbitrary temperature and NRG parameters Λ and M_K.

Smoothing discrete data.—We obtain smooth curves for $\mathcal{A}^{\mathcal{BC}}(\omega)$ by broadening the discrete δ functions in Eq. (10) using a broadening kernel that smoothly interpolates from a log-Gaussian form (of width α) [2,4] for $|\omega| \gtrsim \omega_0$, to a regular Gaussian (of width ω_0) for $|\omega| < \omega_0$, where ω_0 is a "smearing parameter" whose significance is explained below. To obtain high-quality data, we combine small choices of α with an average over N_z slightly shifted discretizations [3] (see [13] for more details).

Application to Anderson model.—We illustrate our method for the standard single-impurity Anderson model (SIAM). Its local Hamiltonian $\hat{h}_0 \equiv \sum_{\sigma}\epsilon_0 c_{0\sigma}^{\dagger}c_{0\sigma} +$
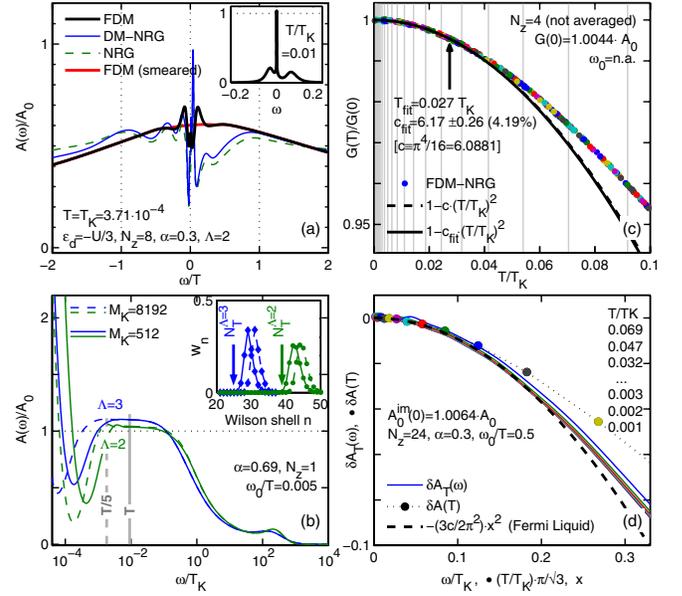

FIG. 2 (color online). FDM-NRG results for the spectral function $\mathcal{A}_T(\omega)$ of the SIAM, with $U = 0.12$, $\Gamma = 0.01$, $\epsilon_d = -U/2$ ($T_K = 2.185 \times 10^{-4}$), $\Lambda = 1.7$, and $M_K = 1024$, unless indicated otherwise. Inset of (a): FDM-NRG result for $A_T(\omega)$ with ω in units of bandwidth. For (a),(b), an unconventionally small smearing parameter was used, $\omega_0 = 0.005T$ [except for thick gray (red) curve in (a)], with $\omega_0 = 0.5T$), leading to spurious low-frequency oscillations. These illustrate the differences (a) between NRG (dashed green curve), DM-NRG [solid thin (blue) curve], and FDM-NRG (black curve) results for the regime $\omega \lesssim T$, and (b) between different choices of M_K and Λ for FDM-NRG, which yield different shapes for the weights w_n [shown in inset of (b)]: larger Λ reduces the scale δ_T at which oscillations set in, but yields less accurate values for the Kondo peak height in the regime $\delta_T \lesssim \omega \lesssim T_K$. (c),(d) Comparison of high-quality FDM-NRG data (dots, solid curves) with exact Fermi-liquid results (black dashed lines) for (c) the conductance $G(T)$ for $T \ll T_K$, and (d) for $\mathcal{A}_T^{\text{im}}(\omega)$ for T, $\omega \ll T_K$. In (c), c_{fit} was found from a data fit to $c_{\text{fit}}(T/T_K)^2$ for $T < T_{\text{fit}}$ (arrow). In (d) we plot $\delta\mathcal{A}_T(\omega) = [A_T^{\text{im}}(\omega) - A_T^{\text{im}}(0)]/A_0^{\text{im}}(0)$ vs ω/T_K (curves) and $\delta\mathcal{A}(T) = [A_T^{\text{im}}(0)/A_0^{\text{im}}(0) - 1]$ vs $(T/T_K)\pi/\sqrt{3}$ (dots), for a set of 12 temperatures between 0.001 and 0.069T_K (with curves and dots having same T in the same color), to illustrate the leading ω and T behavior of $\mathcal{A}_T^{\text{im}}(\omega)$; the dashed black line represents the expected Fermi-liquid behavior in both cases, $-(3c/2\pi^2)x^2$ vs x.

$U c_{0\uparrow}^{\dagger}c_{0\uparrow}c_{0\downarrow}^{\dagger}c_{0\downarrow}$ describes a localized state with energy ϵ_0, with a Coulomb penalty U for double occupancy. It is coupled to a Wilson chain $\sum_{n\sigma}\lambda_n(c_{n+1\sigma}^{\dagger}c_{n\sigma} + \text{H.c.})$, which generates a local level width Γ. We calculated $\mathcal{A}^{<}(\omega) \equiv \mathcal{A}^{c_{0\sigma}^{\dagger}c_{0\sigma}}(-\omega)$, $\mathcal{A}^{>}(\omega) \equiv \mathcal{A}^{c_{0\sigma}c_{0\sigma}^{\dagger}}(\omega)$ and $\mathcal{A} \equiv \mathcal{A}^{>} + \mathcal{A}^{<}$. An "improved" version $\mathcal{A}^{\text{im}}$ thereof can be obtained by calculating the impurity self-energy $\Sigma(\omega, T)$ [6,13] via FDM-NRG, which is less sensitive to smoothening details and yields more accurate results for the Kondo peak height $\mathcal{A}_{T\simeq 0}(0)$ at zero temperature.

Sum rules.—As expected, we find FDM-NRG to be significantly more accurate at lower computational cost

71

than NRG or DM-NRG [8,15]. The sum rules

$$\int d\omega \,\mathcal{A}^{c_{0\sigma}^\dagger c_{0\sigma}}(\omega) = \langle c_{0\sigma}^\dagger c_{0\sigma}\rangle_T, \quad \int d\omega \,\mathcal{A}(\omega) = 1 \quad (11)$$

hold exactly to 10^{-15} for our discrete data, and to 10^{-4} after smoothing (due to numerical integration inaccuracies). Moreover, even for M_K as small as 256, our results for $\mathcal{A}_{T\simeq0}(0)$ and $\mathcal{A}_{T\simeq0}^{\rm im}(0)$ typically agree to within 2% and 0.2%, respectively, with the Friedel sum rule, which requires $\pi\Gamma\mathcal{A}_{T=0}^{\rm exact} = \sin^2\pi\langle c_{0\sigma}^\dagger c_{0\sigma}\rangle_0$. The exact relation $\mathcal{A}^<(\omega) = f(\omega)\mathcal{A}(\omega)$ (f is the Fermi function), which follows from detailed balance, is likewise satisfied well (though not rigorously so): the left-hand side of Eq. (11) typically equals $\int d\omega f(\omega)\mathcal{A}(\omega)$ to better than 10^{-4}.

Low-frequency data.—Because of the underlying logarithmic discretization, all NRG-based schemes for calculating finite-temperature spectral functions inevitably produce spurious oscillations at very low frequencies $|\omega| \ll T$. The scale δ_T at which these set in can be understood as follows: the Lehmann sum in Eq. (1) is dominated by contributions from initial states $|a\rangle$ with energy $E_a \simeq T$, represented by NRG shells with n near N_T. The characteristic energy scale of these states limits the accuracy obtainable for energy differences E_{ba} to accessible final states $|b\rangle$. Thus the scale δ_T is set by those shells which contribute with largest weight w_n to the density matrix.

We analyze this in more detail in Figs. 2(a) and 2(b) by purposefully choosing the smearing parameter to be unconventionally small, $\omega_0 \ll T$. The resulting spurious oscillations are usually smeared out using $\omega_0 \gtrsim \delta_T$ [Fig. 2(a), thick gray (red) curve], resulting in quantitatively accurate spectral functions only for $|\omega| \gtrsim \omega_0 \simeq \delta_T$. For conventional NRG approaches, the "single-shell" approximation $w_n = \delta_{nN_T}$ typically leads to $\delta_T \simeq T$, as can be seen in Fig. 2(a) [dashed (green) line and thin solid (blue) line]. In contrast, FDM-NRG yields a significantly reduced value of $\delta_T \simeq T/5$ [Fig. 2(a), black line, and Fig. 2(b)], since the weighting functions w_n [inset of Fig. 2(b)] retain weight over several shells below N_T, so that lower-frequency information is included.

Fermi-liquid relations.—To illustrate the accuracy of our low-frequency results, we calculated $\mathcal{A}_T^{\rm im}(\omega)$ for $\omega, T \ll T_K$ for the symmetric SIAM, and made quantitative comparisons to the exact Fermi-liquid relations [14],

$$A_T(\omega) \simeq A_0\left[1 - \frac{c}{2}\left(\frac{T}{T_K}\right)^2 - \frac{3c}{2\pi^2}\left(\frac{\omega}{T_K}\right)^2\right],$$

$$G(T) \equiv \int_{-\infty}^{\infty} d\omega A(\omega, T)\left(-\frac{\partial f}{\partial\omega}\right) \simeq A_0\left[1 - c\left(\frac{T}{T_K}\right)^2\right].$$

Here $A_0 \equiv 1/\pi\Gamma$, $c \equiv \pi^4/16$, and the Kondo temperature T_K is defined via the static magnetic susceptibility [4] $\chi_0|_{T=0} \equiv 1/4T_K$. Figures 2(c) and 2(d) show the FDM-NRG data [gray (colored) dots and lines] to be in remarkably good quantitative agreement with these relations (black dashed curves). The results for the "conductance"

$G(T)$, being a frequency integrated quantity obtained by summing over discrete data directly without the need for broadening, are more accurate than for $\mathcal{A}_T^{\rm im}(\omega)$, and reproduce the prefactor c with an accuracy consistently within 5% (until now, accuracies of the order of 10%–30% had been customary). The smoothness of the data in Fig. 2(c), obtained using temperatures not confined to the logarithmic grid $\Lambda^{-n/2}$ [gray vertical lines in Fig. 2(b)], together with the remarkable stability with respect to different z shifts illustrate the accuracy of our approach.

Conclusions.—Our FDM-NRG method offers a transparent framework for the calculation of spectral functions of quantum impurity models, with much improved accuracy at reduced complicational cost. Its results satisfy frequency sum rules rigorously and give excellent agreement with other consistency checks such as the Friedel sum rule, detailed balance, or Fermi-liquid relations, including the regime $\omega \lesssim T$.

Note added.—Just before completion of this work we learned that Peters, Pruschke, and Anders had followed up on the same idea [15].

[1] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975); H. R. Krishnamurti, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980); **21** 1044 (1980).

[2] O. Sakai, Y. Shimizu, and T. Kasuya, J. Phys. Soc. Jpn. **58**, 3666 (1989).

[3] M. Yoshida, M. A. Whitaker, and L. N. Oliveira, Phys. Rev. B **41**, 9403 (1990).

[4] T. A. Costi, A. C. Hewson, and V. Zlatic, J. Phys. C **6**, 2519 (1994).

[5] T. A. Costi, Phys. Rev. B **55**, 3003 (1997).

[6] R. Bulla, A. C. Hewson, and T. Pruschke, J. Phys. C **10**, 8365 (1998).

[7] R. Bulla, T. Costi, and T. Pruschke, arXiv:cond-mat/0701105.

[8] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000).

[9] R. Bulla, T. A. Costi, and D. Vollhardt, Phys. Rev. B **64**, 045103 (2001).

[10] R. Bulla, N. H. Tong, and M. Vojta, Phys. Rev. Lett. **91**, 170601 (2003).

[11] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005); Phys. Rev. B **74**, 245113 (2006).

[12] F. Verstraete *et al.*, arXiv:cond-mat/0504305.

[13] See EPAPS Document No. E-PRLTAO-99-025733 for appendices which give more details. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.

[14] A. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, New York, 1993).

[15] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006). We recommend this paper for a more detailed comparison of the new and previous approaches.

P2.   Variational matrix-product-state approach to quantum impurity models

▶ A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and Jan von Delft
   Phys. Rev. B **80**, 165117 (2009); [cond-mat/0504305v2 (2005)].

In this paper we demonstrated for the first time that DMRG can be directly used on the Wilson chain to reproduce prototypical spectral data obtained by the NRG otherwise. The initial proof of principle concerned the spectral function for the single impurity Anderson model (SIAM) [cond-mat/0504305v1 (2005)]. This was complimented by extensive further numerical calculations on the Kondo model itself, which allowed for a direct comparison to analytical calculations. The results were finally published in the PRB attached.

In order to calculate spectral data within the DMRG, we used the well-established correction vector method in Eq. (P2-5). The major complication, however, was that the broadening η required by the DMRG must be adjusted (1) dynamically, and (2) shall not exceed and thus smear out sharp physical features such as the (exponentially) narrow Kondo peak at zero frequency. The strictly variational approach in Eq. (P2-6) while appealing from a theoretical point of view, turned out insufficient, in practice, to obtain a converged Kondo peak: squaring the resolvent operator of an already ill-conditioned matrix inversion problem is fatal. Instead, using a linearized biconjugate gradient method allowed us to obtain well-converged spectral data throughout the entire spectrum.

Figure P2-1 shows spectral data for a logarithmic discretization using $\Lambda = 1.2$. While this small a Λ is clearly out of reach for the NRG, it is accessible nevertheless to the quasi-variational setting of the DMRG. Using slightly larger broadening η during the calculation with subsequent deconvolution leads to excellent agreement with Friedel sum rule [see inset to Fig. P2-1(d)]. For this calculation, the Wilson chains for spin-up and spin-down had been unfolded, thus allowing for an altered MPS geometry, with the impurity in the center rather than at the end of a semi-infinite Wilson chain. This allowed us to argue in favor of substantially smaller MPS dimensions. Note that the sweeping on the unfolded Wilson chain is again only allowed within a variational setting, as it clearly violates the principle of energy scale separation that is essential to the NRG.

Figure P2-2 shows the spectral function for the Kondo model in the presence of a large magnetic field, $B \gg T_K$, leading to a sharp shoulder-like structure in the spectral data in agreement with analytical studies. Using an *adapted* logarithmic discretization with *uniform* resolution for energies $|\omega| \lesssim B$, an excellent prediction for the spectral function was obtained: the result in Fig. P2-2 shows (1) significantly better resolution at finite frequency as compared to the resolution accessible to the NRG. Moreover, (2) the final spectral data closely resembles the line shape predicted by analytical calculations, *except* for a shift in the peak position, *i.e.* a renormalized $B \to B_{\text{eff}}$. This potential shift had been estimated already in the original analytical study, and indeed, is nicely confirmed by the DMRG results.

In order to obtain a spectral function with good resolution for frequencies $|\omega| \simeq B$,

the *linearized* discretization within the dynamical window $|\omega| \lesssim B$ turned out crucial. In contrast, using fine discretization of the bath around the frequencies with pronounced structure only, *e.g.* around $\omega \simeq B$ while again using more coarse-grained discretization for small frequencies, $|\omega| \ll B$, this immediately would lead to artificial irregular oscillations in the resulting spectral function *within* the window $|\omega| \lesssim B$. The underlying reason can be intuitively understood as follows. By definition of a spectral function, a local excitation is created at the impurity (*e.g.* described by a composite operator in case of the Kondo model). Now having a large magnetic field, this implies that enforcing a certain spin at the impurity opposite to the direction preferred by the magnetic field, loads a large amount of excitation energy into the system. This energy will be dissipated in the system through relaxation. Therefore within the energy window $|\omega| \lesssim B$ all possible transitions and relaxations must be allowed *without any artificial bias*. The latter is provided by a uniform linear discretization of the bath within this window. A similar kind of argument is clearly also expected to have implications for true out-of-equilibrium simulations at finite voltage bias V (*e.g.* replacing the magnetic field B) in the presence of two thermal reservoirs.

The DMRG represents an accurate and highly flexible framework to deal with the problem presented. Nevertheless, it should be emphasized that this also comes with a significant price tag.[2] The reason is that for every single frequency data point ω_i in the final spectral function, the corresponding correction vector must be calculated. Using the correction vector from a neighboring data point in frequency as an initial guess for the correction vector, this usually leads to slow convergence, to the extent that there is no gain in efficiency as compared to restarting the calculation from scratch from a randomly initialized starting vector for each frequency. Therefore the latter was the method of choice for the calculations presented in the paper attached which also can be easily parallelized on a cluster.

In the time since above paper was published, however, we developed an alternative strong method for calculating spectral functions based on Chebyshev expansion, akin to what is typically used with exact diagonalization [Holzner et al. (2011)]. This allows to calculate spectral functions in a single-shot calculation over the entire frequency range, the cost of which, amazingly, for the same accuracy is *comparable or less* as compared to the cost of a *single* frequency data point using the correction vector ansatz!

---

[2]I had been asked once to estimate the cost of the curve shown in Fig. P2-2 in terms of electricity. A conservative estimate was about € 100.

Variational matrix-product-state approach to quantum impurity models

A. Weichselbaum,[1] F. Verstraete,[2] U. Schollwöck,[1] J. I. Cirac,[3] and Jan von Delft[1]

[1]*Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience,*
Ludwig-Maximilians-Universität München, 80333 Munich, Germany
[2]*Institut der Theoretischen Physik, Universität Wien, Boltzmanngasse 3, A-1090 Vienna, Austria*
[3]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, Garching D-85748, Germany*

We present a unified framework for renormalization group methods, including Wilson's numerical renormalization group (NRG), and White's density-matrix renormalization group (DMRG), within the language of matrix-product-states. This allows improvements over Wilson's NRG for quantum impurity models, as we illustrate for the one-channel Kondo model. Moreover, we use a variational method for evaluating Green's functions. The proposed method is more flexible in its description of spectral properties at finite frequencies, opening the way to time-dependent, out-of-equilibrium impurity problems. It also substantially improves computational efficiency for one-channel impurity problems, suggesting potentially *linear* scaling of complexity for *n*-channel problems.

Wilson's numerical renormalization group (NRG) is a key method[1] for solving quantum impurity models such as the Kondo, Anderson, or spin-boson models, in which a local degree of freedom, the "impurity," is coupled to a continuous bath of excitations. These models are of high relevance in the description of magnetic impurities, quantum dots, and problems of decoherence. NRG has been used with great success to calculate both thermodynamic[1,2] and dynamical[3–6] properties. It is, however, of limited use in more complex situations: computational cost grows exponentially for a coupling to multiple bands in the bath. In systems out-of-equilibrium or with time-dependent external parameters, such as occur in the tuning of quantum dots, difficulties arise due to NRG's focus on low-energy properties through its logarithmic discretization scheme which looses accuracy at high spectral frequencies.

In the present paper, we draw attention to the fact that states generated by the NRG have the structure of *matrix-product-states* (MPS)[7,8] on a one-dimensional geometry. This is a simple observation, which however has important conceptual and practical implications:

(i) As White's density-matrix renormalization group (DMRG)[9] for treating quantum chain models is in its single-site version identical to variational MPS,[8] NRG, and DMRG are now seen to have the same formal basis of matrix-product-states, resolving a long-standing question about the connection between both methods. (ii) *All* NRG results can be improved upon systematically by *variational optimization* in the space of variational matrix-product-states (VMPS) of the same structure as those used by NRG. This does not lead to major improvements at $\omega=0$ where NRG works very well, but leads to the inclusion of feedback from low-to-high-energy states, also allowing the relaxation of the logarithmic bath discretization of NRG: spectra away from $\omega=0$ can be described more accurately and with higher resolution. (iii) Recent algorithmic advances using VMPS,[8] in particular those treating time-dependent problems,[10,11] can now be exploited to tackle quantum impurity models involving time dependence or nonequilibrium; this includes applications to the description of driven qubits coupled to decohering baths,

as relevant in the field of quantum computation. (iv) The VMPS algorithm allows ground state properties of quantum impurity models to be treated more efficiently than NRG: the same accuracy is reached in much smaller ansatz spaces (roughly of square-root size). Moreover, our results suggest that for many (if not all) *n* channel impurity problems it should be feasible to use an *unfolded* geometry, for which the complexity will only grow linearly with *n*.

The present paper provides a "proof of principle" for the VMPS approach to quantum impurity models by applying it to the one-channel Kondo model. We reproduce the NRG flow of the finite size spectrum,[2] and introduce a VMPS approach for calculating Green's functions, as we illustrate for the impurity spectral function,[3] which yields a significant improvement over existing alternative techniques.[12–15] Our results illustrate in which sense the VMPS approach is numerically more efficient than the NRG.

I. NRG GENERATES MATRIX-PRODUCT-STATES

To be specific, we consider Wilson's treatment of the Kondo model, describing a local spin-1/2 impurity in an external magnetic field B coupled to a fermionic bath. To achieve a separation of energy scales, the bath excitations are represented by a set of logarithmically spaced, discrete energies $\omega_n=\Lambda^{-n}$, where $\Lambda>1$ is a "discretization parameter."[1] By tridiagonalization, the model is then mapped onto the form of a semi-infinite chain $\mathcal{H}=\lim_{N\to\infty}\mathcal{H}^N$ where[1]

$$\mathcal{H}^N = BS_z - 2J\mathbf{s}\cdot\mathbf{S} + \sum_{n=1}^{N-1} \xi_n(c^\dagger_{n\mu}c_{n+1,\mu} + c^\dagger_{n+1,\mu}c_{n\mu}). \quad (1)$$

$\mathcal{H}^N$ describes an impurity spin $\mathbf{S}$ in a Zeeman field B, exchange coupled to the spin $\mathbf{s}=\frac{1}{2}c^\dagger_1\boldsymbol{\sigma}c_1$ of the first site of a chain of length N of fermions with spin μ and exponentially decreasing hopping matrix elements along the chain ($\xi_n \sim \Lambda^{-n/2}$). $\mathcal{H}^N$ lives on a Hilbert space spanned by the set of $d_I d^N$ basis states $\{|i_0,i_1,i_2,\ldots i_N\rangle\}$, where i_0 labels the d_I possible impurity states and i_n (for $n=1,\ldots,N$) the d possible

states of site n (for the Kondo model, $i_0 = \{\uparrow, \downarrow\}$ and for all other sites $i_n = \{0, \uparrow, \downarrow, \uparrow\downarrow\}$, i.e., $d_I = 2$ and $d = 4$).

To diagonalize the model, NRG starts with a chain of length $(\bar{n} - 1)$, chosen sufficiently small that $\mathcal{H}^{\bar{n}-1}$ can be diagonalized exactly, yielding a set of eigenstates $|\psi_\alpha^{\bar{n}-1}\rangle$. One continues with the subsequent iterative prescription: project $\mathcal{H}^{\bar{n}-1}$ onto the subspace spanned by its lowest D eigenstates, where $D < d_I d^{\bar{n}-1}$ is a control parameter (typically between 500 and 2000); add site $\bar{n}$ to the chain and diagonalize $\mathcal{H}^{\bar{n}}$ in the enlarged (Dd)-dimensional Hilbert space, writing the eigenstates as

$$|\psi_\beta^{\bar{n}}\rangle = \sum_{i_{\bar{n}}=1}^{d} \sum_{\alpha=1}^{D} |\psi_\alpha^{\bar{n}-1}\rangle |i_{\bar{n}}\rangle P_{\alpha\beta}^{[i_{\bar{n}}]}, \qquad (2)$$

where the coefficients have been arranged in a matrix $P_{\alpha\beta}^{[i_{\bar{n}}]}$ with matrix indices α, β, labeled by the site index $\bar{n}$, and state index $i_{\bar{n}}$; rescale the eigenenergies by a factor $\Lambda^{1/2}$; and repeat, until the eigenspectrum converges, typically for chain lengths N of order 40 to 60. At each step of the iteration, the eigenstates of $\mathcal{H}^N$ can thus be written [by repeated use of Eq. (2)] in the form of a so-called matrix-product-state,

$$|\psi_\alpha^N\rangle = P_{\alpha_0}^{[i_0]} P_{\alpha_0\alpha_1}^{[i_1]} P_{\alpha_1\alpha_2}^{[i_2]} \dots P_{\alpha_{N-1}\alpha}^{[i_N]} |i_0, i_1, \dots, i_N\rangle \qquad (3)$$

(summation over repeated indices implied). The ground state is then the lowest eigenstate of the effective Hamiltonian $\mathcal{H}_{\alpha\beta}^N = \langle \psi_\alpha^N | \mathcal{H}^N | \psi_\beta^N \rangle$, i.e., the projection of the original $\mathcal{H}$ on the subspace of MPS of the form (3).

II. VMPS OPTIMIZATION

Let us now be more ambitious and aim to find the *best possible* description of the ground state within the space of all MPSs of the form (3), using the matrix elements of the matrices $\{P^{[n]}\}$ with $P^{[n]} \equiv \{P^{[i_n]}\}$ as *variational parameters* to minimize the energy. Using a Lagrange multiplier to ensure normalization, we thus study the following optimization problem:

$$\min_{|\psi^N\rangle \in \{\text{MPS}\}} [\langle \psi^N | \mathcal{H}^N | \psi^N \rangle - \lambda \langle \psi^N | \psi^N \rangle]. \qquad (4)$$

This cost function is multiquadratic in the $d_I + d(N-1)$ matrices $\{P^{[n]}\}$ with a multiquadratic constraint. Such problems can be solved efficiently using an iterative method in which one fixes all but one (let's say the $\bar{n}$'th) of the matrices $\{P^{[n]}\}$ at each step; the optimal $P^{[\bar{n}]}$ minimizing the cost function given the fixed values of the other matrices can then be found by solving an eigenvalue problem.[8] With $P^{[\bar{n}]}$ optimized, one proceeds the same way with $P^{[\bar{n}+1]}$ and so on. When all matrices have been optimized locally, one sweeps back again, and so forth. By construction, the method is guaranteed to converge as the energy goes down at every step of the iteration, having the ground-state energy as a global lower bound. Given the rather monotonic hopping amplitudes, we did not encounter problems with local minima.

In contrast, NRG constructs the ground state in a single *one-way* sweep along the chain: each $P^{[n]}$ is thus calculated

only once, without allowing for possible feedback of P's calculated later. Yet viewed in the above context, the ground-state energy can be lowered further by MPS optimization sweeps. This accounts for the *feedback* of information from low-to-high-energy scales. This feedback may be small in practice, but it is not strictly zero, and its importance increases as the logarithmic discretization is refined by taking $\Lambda \to 1$. Note that the computational complexity of both VMPS optimization and NRG scales as NdD^3,[8,9] and symmetries can be exploited (with similar effort) in both approaches. The inclusion of feedback leads to a better description of spectral features at high frequencies, which are of importance in out-of-equilibrium and time-dependent impurity problems. Moreover, it also allows to relax the logarithmic discretization scheme, further improving the description of structures at high frequency as illustrated below.

The result of a converged set of optimization sweeps is a VMPS ground state $|\tilde{\psi}_0^N\rangle$ of the form (3); exploiting a gauge degree of freedom,[8] the $\tilde{P}$'s occurring therein can always be chosen such that all vectors $|\tilde{\psi}_\alpha^n\rangle = [\tilde{P}^{[i_0]} \dots \tilde{P}^{[i_n]}]_\alpha |i_0, \dots, i_n\rangle$ are orthonormal. The effective Hamiltonian at chain length n, the central object in NRG, is then $\tilde{\mathcal{H}}_{\alpha\beta}^n = \langle \tilde{\psi}_\alpha^n | \Lambda^{n/2} \mathcal{H}^n | \tilde{\psi}_\beta^n \rangle$. Its eigenspectrum can be monitored as n increases, resulting in an energy level flow along the chain.

III. GREEN'S FUNCTIONS

Similar techniques also allow Green's functions to be calculated variationally.[15] The typical Green's functions of interest are of the form $G_\eta^c(\omega) = \langle \psi_0 | c | \chi \rangle$ where $|\chi\rangle$, commonly called a correction vector,[16] is defined by

$$|\chi\rangle \equiv \frac{1}{\omega - \mathcal{H} + i\eta} c^\dagger |\psi_0\rangle, \qquad (5)$$

with $|\psi_0\rangle$ the ground state of the system, e.g. calculated using the VMPS approach and thus represented as MPS. The spectral density is then given by $\mathcal{A}(\omega) = -\lim_{\eta \to 0} \frac{1}{\pi} \mathfrak{Im}(G_\eta^c(\omega))$. The (unnormalized) state $|\chi\rangle$ may be calculated *variationally* within the set of MPS by optimizing the weighted norm

$$\mathcal{N} = \left\| \chi > - \frac{1}{\mathcal{H} - \omega - i\eta} c^\dagger |\psi_0\rangle \right\|_{W = (\mathcal{H} - \omega)^2 + \eta^2}, \qquad (6)$$

where $\|\xi>\|_W^2 \equiv \langle \xi | W | \xi \rangle$, and weight $W > 0$ such that it yields a quadratic equation. Writing $|\chi\rangle \equiv |\chi_r\rangle + i |\chi_i\rangle$ and assuming $\mathcal{H}$, $|\psi_0\rangle$, $|\chi_r\rangle \equiv \mathfrak{Re}|\chi\rangle$, and $|\chi_i\rangle \equiv \mathfrak{Im}|\chi\rangle$ real, this norm can be written as (compare Ref. 14) $\mathcal{N}^2 = \langle \chi_r | (\mathcal{H} - \omega)^2 + \eta^2 |\chi_r\rangle - 2\langle \chi_r | (\mathcal{H} - \omega) c^\dagger |\psi_0\rangle + \langle \chi_i | (\mathcal{H} - \omega)^2 + \eta^2 |\chi_i\rangle - 2\eta \langle \chi_i | c^\dagger |\psi_0\rangle + \langle \psi_0 | cc^\dagger |\psi_0\rangle$. Minimizing $\mathcal{N}$ efficiently by optimizing one P at a time leads to two independent optimizations over $|\chi_r\rangle$ and $|\chi_i\rangle$, respectively. Both involve only multilinear terms such that each iteration step requires to solve a sparse linear set of equations.[10]

Minimizing $\mathcal{N}$ involves the calculation of $\mathcal{H}^2$, which can be done efficiently as follows. Generally speaking, $\mathcal{H}$ has the structure of a spin chain with only nearest-neighbor couplings as shown in Eq. (1). Naively one expects that one will have to evaluate on the order of N^2 expectation values of
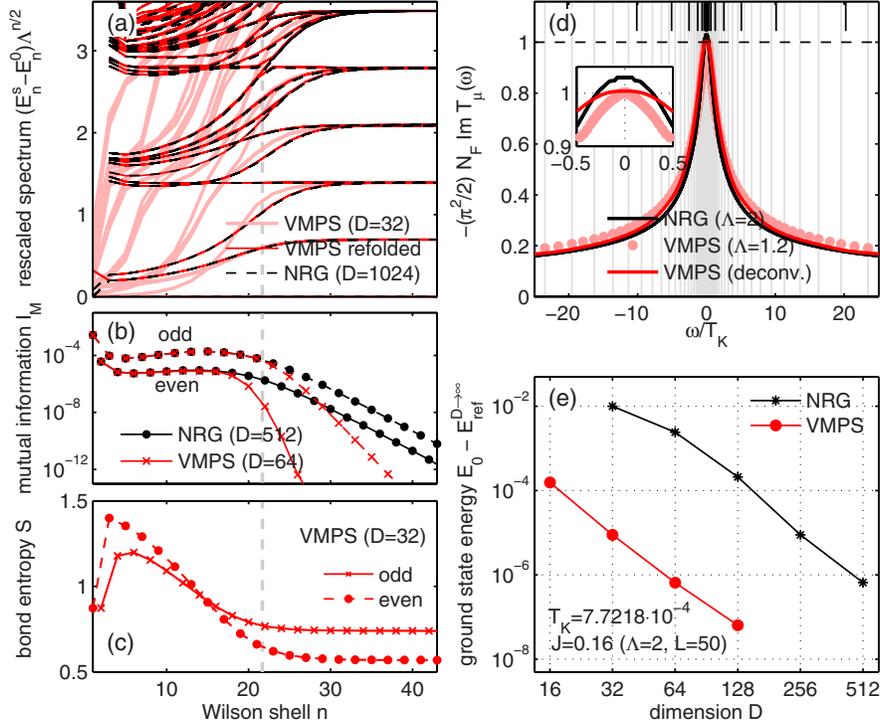
FIG. 1. (Color online) Comparison of VMPS and NRG data for logarithmic discretization of the Kondo model as in Eq. (1) for J =0.16 and Λ=2 if not specified otherwise. (a) Energy level flow of the Kondo model as a function of site index n obtained from $\mathcal{H}_\mu^{\mathrm{eff}}$ of a variationally optimized MPS with D_{MPS}=32 (light red/light solid), of the corresponding recombined spin chains (red/dark solid) (Ref. 17), and from NRG using D_{NRG}=32^2 states (dashed black). The Wilson shell corresponding to $T_K \equiv \sqrt{J}e^{-1/J}$ is indicated by the vertical dashed line through panels (a) to (c). (b) Correlation along the Wilson chain between spin up and spin down at site n in terms of mutual information $I_M(n) \equiv S(n_\uparrow)+S(n_\downarrow)-S(n_\uparrow,n_\downarrow)$. Here S is the entropy of the reduced density matrix of the groundstate with respect to the indicated subspace (Ref. 17) (solid for even, dashed for odd sites n). (c) Bond entropy S along the unfolded Wilson chain, where S is the usual von Neumann entropy of the VMPS reduced density matrix when going from site n to $n+1$, plotted for even and odd iterations, respectively. (d) Comparison of T matrix ($\Im m T_\mu$, see also Fig. 2) for B=0 between VMPS and NRG, including deconvoluted VMPS data (see Appendix). Inset shows zoom into peak at ω=0. The significantly smaller Λ=1.2 applicable for VMPS (discretization intervals are indicated by vertical lines) shows clearly improved agreement with the Friedel sum rule $T(0)\pi^2/2$=1. (e) Comparison of ground-state energy of the Kondo Hamiltonian Eq. (1) for fixed chain length relative to the extrapolated energy for $D \to \infty$ for VMPS and NRG as function of the dimension D of states kept.

local observables. There is, however, a clever linear scheme that only requires iterative update of a set of effective operators. Therefore the computational complexity of calculating ϵ scales as ND^3 similar to just evaluating the energy $\langle \psi | \mathcal{H} | \psi \rangle$. As a side product, this leads to efficient algorithms for calculating excited eigenstates of Hamiltonians close to a fixed energy E by minimizing $\langle \psi | (\mathcal{H}-E)^2 | \psi \rangle$. Moreover, it can be used to estimate errors on eigenenergies as it can be shown that there exists an exact eigenvalue E_{ex} within an interval around E specified by $\epsilon = \sqrt{\langle \psi | (\mathcal{H}-E)^2 | \psi \rangle}$.

For quantum impurity systems with sharp features such as the Kondo model discussed below, it should be noted, however, that the broadening η may have to be chosen extremely small. In this case, the minimization of $\mathcal{N}$ in Eq. (6) can become increasingly ill conditioned as $\eta \to 0$ (see Appendix), with conditioning deteriorating quadratically in η. If one directly solves $\delta / \delta \langle P^{[n]} | [\langle \chi | (\mathcal{H}-\omega-i\eta) | \chi \rangle - \langle \chi | c^\dagger | \psi \rangle] \equiv 0$ by a nonhermitian equation solver such as the biconjugate gradient method, conditioning deteriorates only linearly. This is the strategy that has been followed to obtain the results reported below.

IV. APPLICATION TO KONDO MODEL

Let us now illustrate above strategies by applying them to the Kondo model. Since the Hamiltonian in Eq. (1) couples $\uparrow$ and $\downarrow$ band electrons only via the impurity spin, it is possible (see also Refs. 5 and 17) to "unfold" the semi-infinite Wilson chain into an infinite one, with $\uparrow$ band states to the left of the impurity and $\downarrow$ states to the right, and hopping amplitudes decreasing in both directions as $\Lambda^{-|n|/2}$. Since the left and right end regions of the chain, which describe the model's low-energy properties, are far apart and hence interact only weakly with each other [analyzed quantitatively in terms of mutual information in Fig. 1(b)], the effective Hamiltonian for these low energies will be of the form $\mathcal{H}_\uparrow^{\mathrm{eff}} \otimes 1\!\!1_\downarrow + 1\!\!1_\uparrow \otimes \mathcal{H}_\downarrow^{\mathrm{eff}}$. Due to the symmetry of the Kondo coupling, $\mathcal{H}_\uparrow^{\mathrm{eff}}$ and $\mathcal{H}_\downarrow^{\mathrm{eff}}$ have the same eigenspectrum for $n \geq 1$, such that the fixed point spectrum is already well reflected by analyzing either one, as illustrated in Fig. 1(a). Note that for a direct comparison with NRG, the spin chains can be *recombined* within VMPS.[17] The resulting standard energy flow diagram presented in panel (a) for VMPS and NRG, respectively,

show excellent agreement for low energies for all n including the fixed point spectrum.

In addition to the energy flow diagram in panel (a), convergence with and hence sensitivity on the Kondo energy scale is also nicely seen in other quantities typically calculated within VMPS: in the internal entanglement of the Wilson chain as function of site n in terms of mutual information, shown in panel (b), and the entanglement of Wilson chain up to site n with the remainder of the Wilson chain, shown in panel (c). Note that due to intrinsic even/odd effects of the model, the data from even/odd Wilson sites is plotted separately.

The dimensions of the effective Hilbert spaces needed for VMPS for the unfolded Wilson chain and NRG for the inevitably folded chain to capture the low-energy properties (here energy resolution better than T_K) are roughly related by $D_{\mathrm{MPS}} \sim \sqrt{D_{\mathrm{NRG}}}$,[17] implying significant computational gain with VMPS, as calculation time scales as D^3 for both. Indeed, Fig. 1(e) shows that VMPS has three orders of magnitude of better precision for the *same* D. More generally, if the impurity couples to n electronic bands (channels), the Wilson chain may be unfolded into a starlike structure of $2n$ branches, with $D_{\mathrm{MPS}} \sim D_{\mathrm{NRG}}^{1/2n}$. This implies that for maintaining a desired precision in going from 1 to n channels, D_{MPS} will stay roughly constant, and calculation time for all sites other than the impurity will scale merely *linearly* with the number of channels. Whether the chains can be unfolded in practice can easily be established by checking whether or not the correlation between them, characterized, e.g., in terms of mutual information, decays rapidly with increasing n [cf. Fig. 1(b) and caption].

V. ADAPTIVE DISCRETIZATION

Through its variational character, VMPS does not rely on logarithmic discretization crucial for NRG. The potential of greatly enhanced energy resolution using VMPS is already indicated by the $\Lambda = 1.2$ data in Fig. 1(d). It is illustrated to full extent in Fig. 2, showing the splitting of the Kondo peak in the presence of a strong magnetic field calculated using VMPS (bare: dots, deconvoluted: red solid), standard NRG (blue dashed), and perturbatively[18] (black).

To obtain the VMPS results of Fig. 2, we used an *adapted* discretization scheme for the energies ε_k of the conduction band Hamiltonian $H_{\mathrm{band}} = \Sigma_{k\mu} \varepsilon_k c_{k\mu}^{\dagger} c_{k\mu}$ that forms the starting point for deriving the Wilson chain like Hamiltonian of Eq. (1).[1] Namely, we use a linear or logarithmic discretization scheme for $|\varepsilon_k| < B$ or $> B$, respectively, (as illustrated by light vertical lines in Fig. 2). The nearest-neighbor coupling amplitudes ξ_n of the resulting, modified Wilson chain decay only very slowly with n once the energy scale of site n, namely, $\Lambda^{-n/2}$, drops below B, in contrast to their usual exponential decay for a standard Wilson chain (see inset of Fig. 2, VMPS vs $\Lambda = 1.7$ NRG coupling). The slow decay of ξ_n implies an increased energy resolution at energies up to B, at the cost of a loss of energy scale separation. While the latter fact implies that NRG cannot be used on such a chain, the VMPS approach does not suffer from this limitation. Indeed, it exploits the enhanced energy resolution at energies of or-
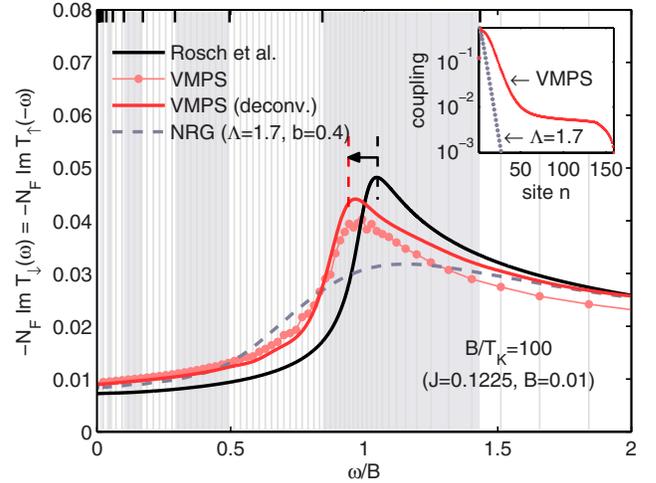

FIG. 2. (Color online) Impurity spectral function for the Kondo model $-N_F \Im m T_{\mu}(\omega) = J^2 \langle\langle \mathcal{O}_{\mu}^{\dagger} | \mathcal{O}_{\mu} \rangle\rangle_{\omega}$ for $B \gg T_K$, where $\mathcal{O}_{\mu} \equiv \mathbf{S} \cdot \boldsymbol{\sigma}_{\mu\mu'} c_{\mu'}^{\dagger}$, and N_F is the density of states at the Fermi energy, calculated with VMPS (dots: raw data, red solid: deconvoluted), NRG (dashed), and perturbative (black solid) (Ref. 18). For NRG, $D = 1024$ states were kept, using a log-Gauss broadening parameter (Ref. 19) of $b = 0.4$. According to Ref. 18, the peak of the perturbative result should be shifted in ω by $B/2 \log(B/T_K)$ (arrow). NRG and VMPS discretization intervals are indicated by shaded areas and gray vertical lines, respectively. Due to the increased linear resolution for $|\omega| \leq B$, the number of states retained within VMPS needed to be increased, and was dynamically governed by either a threshold of $4 \cdot 10^{-8}$ in discarded weight or a maximum number of states of $D = 512$. The latter was required only for frequencies around $\omega \sim B$. The inset shows the hopping amplitudes corresponding to standard logarithmic ($\Lambda = 1.7$) and adapted (VMPS) discretization schemes. The required Lorentzian broadening η of the VMPS data smears out sharper features. Deconvolution (targeting with adaptive spline) together with subsequent GAUSSIAN broadening was applied to obtain the solid light line (see Appendix).

der B to yield spectral peaks around $\omega \simeq B$ that are significantly sharper than those obtainable by NRG (Fig. 2, compare dotted data points to dashed line). The resolution can be enhanced even further (Fig. 2, solid thick light line) by applying a deconvolution scheme to the VMPS data, detailed in the appendix, to account for the broadening effects of using an *a priori* finite η required within the VMPS approach.

Note that such a resolution is out of reach for conventional NRG, whose discretization intervals (shaded intervals), even for comparatively small choice of $\Lambda = 1.7$, are much broader than the spectral features of interest. Note that the NRG data shown here is, by conventional NRG standards, of high quality: first, we used a rather small value of $\Lambda = 1.7$, implying high-energy resolution, by NRG standards; second, we employed the recently developed full density-matrix (FDM) approach,[19,20] which incorporates systematic improvements relative to previous NRG implementations. We have tried extensively to improve the quality of our NRG data via z-averaging,[21] but have found this to be of limited use.

The line shape of our deconvoluted data (red solid line) agrees well with the analytic RG calculation[18] (black solid

line), perturbative in $1/\log(B/T_K)$. The peak positions agree well also after a shift in ω by $-B/2\log(B/T_K)$ of the perturbative result suggested by[18] is taken into account.

As pointed out in the context of Fig. 1, numerical resources in terms of matrix dimensions can be drastically reduced within VMPS when applied to the NRG discretized model, in that the variational freedom provides a highly adaptive method. However, this does come at a price. While the calculation of spectral functions within NRG for the full frequency range are obtained in a single run about as expensive numerically as the iterative diagonalization of the NRG Hamiltonian and as such highly efficient, the correction vector method [cf. Eq. (5)] provides an optimal setting for one frequency at a time. This is highly tailored toward analyzing certain features in frequency space, but implies that for every spectral data point a new correction vector must be obtained, which is itself equally expensive numerically as the calculation of the ground state. Nevertheless, in situations that become computationally hard in NRG or are simply out of reach for NRG due to the required rather crude coarse graining of the conduction band, VMPS does provide a well-controlled technique that can clearly compete and in certain cases outperform NRG. In the example given in Fig. 2, the logarithmic discretization scheme was adapted by introducing an energy interval from $-B$ to B in which the level spacing was chosen to be essentially uniform. It appears, indeed, that for describing dynamical features at frequencies $\omega \sim B$, *all* states with frequencies $|\omega| \lesssim B$ are equally important, which necessitates the use of a uniform level spacing from $-B$ to B. Of course, this does break energy scale separation from the very outset.

VI. OUTLOOK

Let us finish by pointing out that the MPS approach can readily be extended to the case of finite temperatures by using matrix-product density operators[10] instead of MPS, and to time-dependent problems [such as $\mathcal{H} = \mathcal{H}(t)$ or nonequilibrium initial conditions], by using the recently developed adaptive time-dependent DMRG[11] and MPS analogs thereof.[10] Exploratory work in this direction has been very encouraging.[22]

In conclusion, the MPS approach provides a natural language for simulating quantum impurity models. The underlying reason is that these models, when formulated on the Wilson chain, feature only nearest-neighbor interactions. Their low-energy states are thus determined mainly by their nearest-neighbor reduced density matrices, for which very good approximations can be obtained by suitably optimizing the set of matrices constituting a MPS.[23] We also showed how these could be used for a direct (quasi) variational evaluation of Green's functions.

Recently, it has come to our attention that two recent papers by Freyn and Florens[24] and Zitko and Pruschke[25] who claimed improved resolution of NRG spectral functions. Besides properly accounting for the wave function renormalization A_Λ due to discretization,[26] Ref. 25 is heavily based on z-averaging[21] with modest success for finite frequencies—see for example Fig. 8 in Ref. 25 which shows spurious

oscillations. It is exactly these spurious oscillations we had also seen in the z-averaging done excessively for our model. In order to get rid of these spurious oscillations in a systematic unbiased manner, however, one would have to rebroaden the data to get discretization-independent correlation function. Hence, although z-averaging does show modest improvements (as known since Ref. 21), it cannot be expected to cure in much detail the rather crude coarse graining of the conduction band put into the model from the beginning. Reference 24 introduced a procedure for broadening the raw NRG data to obtain smooth spectral peaks, employing a frequency-dependent broadening parameter $b(\omega)$. This led to significantly increased resolution for spectral peaks of the spin-boson model with very weak damping. However, when we tried this method for the present Kondo model, the improvements over conventional NRG broadening techniques were also found to be rather modest.

APPENDIX: DECONVOLUTION OF SPECTRAL DATA

DMRG obtains spectral data from a discretized model Hamiltonian. In order for the spectral data to be smooth, an intrinsic frequency-dependent Lorentzian broadening η is applied during the calculation of the correction vector $|\chi\rangle_k$ at frequency ω_k [cf. Eq. (5)],

$$\delta_{\eta_k}(\omega - \omega_k) \equiv \frac{\eta_k}{\pi} \frac{1}{(\omega - \omega_k)^2 + \eta_k^2}. \tag{7}$$

Since the original model has a continuous spectrum, the broadening η_k should be chosen of the order or larger than the artificial coarse grained discretization intervals δ_ω. Larger η of course improves numerical convergence. However, since Lorentzian broadening produces longer tails than for example GAUSSIAN broadening, this makes it more susceptible to pronounced spectral features close by. Our general strategy for more efficient numerical treatment was then as follows. (i) Choose somewhat larger η ($\eta \simeq 2\delta_\omega$) throughout the calculation. (ii) Deconvolve the raw data to such an extent that the underlying discrete structure already becomes visible again, (iii) followed by a GAUSSIAN smoothening procedure which then acts more locally. Let us describe step (ii) in more detail.

Broadening, by construction, *looses* information. Hence trying to obtain the original data from the broadened data via deconvolution is intrinsically ill conditioned. In literature there are several ways of dealing with this problem, most prominently maximum entropy algorithms (see Refs. 5 and 27). Our approach is targeting the actual spectral function using the knowledge about the Lorentzian broadening used during the VMPS calculation, combined with adaptive spline. Given the data $\widetilde{A}(\omega)$ obtained through VMPS, let us propose the existence of some smooth but *a priori* unknown target curve $A(\omega)$, which when broadened the *same* way as the VMPS data using exactly the same η_k via a Lorenzian broadening kernel

$$\widetilde{A}_k \equiv \widetilde{A}(\omega_k) = \int_{-\infty}^{\infty} d\omega' A(\omega') \delta_{\eta_k}(\omega' - \omega_k), \qquad (8)$$

reproduces the original data $\widetilde{A}(\omega)$. Direct inversion of above equation as it is ill conditioned, as already mentioned, and not useful in practice.

Let us assume the unknown target curve $A(\omega)$ is smooth and parametrized by piecewise polynomials. Given the data points ω_k with $k=1\ldots N$, the intervals in between these values will be approximated in the spirit of adaptive spline functions[27] by 3rd order polynomials ($k=1\ldots N-1$)

$$f_k(\omega) \equiv \begin{cases} a_k + b_k(\omega - \omega_k) + c_k(\omega - \omega_k)^2 + d_k(\omega - \omega_k)^3 & \text{for } \omega \in [\omega_k, \omega_{k+1}] \\ 0 & \text{otherwise}. \end{cases} \qquad (9)$$

Since spectral functions decay as $1/\omega^2$ for large ω, for our purposes the ends are extrapolated asymptotically to infinity, allowing both $1/\omega$ and $1/\omega^2$ polynomials

$$f_0(\omega) \equiv \begin{cases} \dfrac{a_0}{\omega} + \dfrac{b_0}{\omega^2} & \omega \le \omega_1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_N(\omega) \equiv \begin{cases} \dfrac{a_N}{\omega} + \dfrac{b_N}{\omega^2} & \omega \ge \omega_N \\ 0 & \text{otherwise}. \end{cases} \qquad (10)$$

In total, this results in $4(N-1)+2\times 2 = 4N$ parameters, with the target function parametrized piecewise as $A(\omega) \equiv f(\omega) \equiv \Sigma_{k=0}^{N} f_k(\omega)$. In cases where one has not approached the asymptotic limit yet, the ends may simply be modeled also by Eq. (9), taking $c_0 = d_0 = c_N = d_N = 0$. Moreover, if information about the gradient $f'(\omega)$ is known, it can be built in straightforwardly in the present scheme by replacing b_k.

The parameters for the piecewise parametrization are solved for by requiring the following set of conditions:

(i) The function f should be continuous and smooth by requiring that f, f', and f'' are continuous ($3N$ equations).

(ii) The function f, when broadened as in Eq. (8), should reproduce the VMPS data $\widetilde{A}_k$

$$\widetilde{A}_k^c \equiv \sum_{k'=0}^{N} \int_{\omega_{k'}}^{\omega_{k'+1}} d\omega' f_{k'}(\omega') \frac{\eta_k/\pi}{(\omega' - \omega_k)^2 + \eta_k^2} \qquad (11)$$

$$\widetilde{A}_k - \widetilde{A}_k^c = p_k r_k \qquad (12)$$

where $r_k \equiv f_k^{(3)}(\omega_k) - f_{k-1}^{(3)}(\omega_k)$ and $\omega_0 \equiv -\infty$, $\omega_{N+1} \equiv +\infty$ (N equations).

In the spirit of adaptive spline, the third derivative of the piecewise polynomials is no longer required to be continuous. Its jump r_k is set proportional to the change in $\widetilde{A}_k - \widetilde{A}_k^c$ introducing the additional prespecified parameter set p_k, kept small for our purposes (note that enforcing the strict equality $\widetilde{A}_k^c = \widetilde{A}_k$ by setting $p_k = 0$ would result in an ill-conditioned problem).

If interval spacings specified by ω_k are nonuniform, the p_k have to be adapted accordingly. For this paper we used $p_k = p \cdot (\omega_{k+1} - \omega_k)^{\alpha}$ with p on the order of 10^{-6} and $\alpha \simeq 1$. With p_k fixed, Eqs. (11) and (12) determine all spline parameters uniquely in terms of the original VMPS data $\widetilde{A}_k$. The integrals emerging out of Eq. (11) can all be evaluated analytically. The final inversion of Eq. (11) to obtain the parameters for $f(\omega)$ is well behaved for small but finite p, small enough to clearly sharpen spectral features.

[1] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[2] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980).

[3] T. A. Costi, A. C. Hewson, and V. Zlatic, J. Phys.: Condens. Matter **6**, 2519 (1994).

[4] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000).

[5] C. Raas, G. S. Uhrig and F. B. Anders, Phys. Rev. B **69**, 041102(R) (2004); C. Raas and G. S. Uhrig, Eur. Phys. J. B **45**, 293 (2005).

[6] R. Bulla, A. C. Hewson, and Th. Pruschke, J. Phys.: Condens. Matter **10**, 8365 (1998); R. Bulla, N. H. Tong, and M. Vojta, Phys. Rev. Lett. **91**, 170601 (2003).

[7] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992); S. Ostlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995); J. Dukelsky, M. A. Martín-Delgado, T. Nishino, and G. Sierra, Europhys. Lett. **43**, 457 (1998); H. Takasaki, Toshiya Hikihara, and Tomotoshi NishinoJ. Phys. Soc. Jpn. **68**, 1537 (1999).

[8] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004).

[9] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[10] F. Verstraete, J.-J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).

[11] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004); A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech. Theor. Exp. (2004) P04005; S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).

[12] K. A. Hallberg, Phys. Rev. B **52**, R9827 (1995).

[13] T. D. Kühner and S. R. White, Phys. Rev. B **60**, 335 (1999).

[14] E. Jeckelmann, Phys. Rev. B **66**, 045114 (2002).

[15] Compared to other techniques for calculating Green's functions (Refs. 3–5, 12–14, and 28), the VMPS approach proposed here has the advantage that it is variational, and hence in principle optimal within the set of MPS. It is more efficient than the continued fraction method (Ref. 12), the correction vector method (Ref. 13) and dynamical DMRG (Ref. 14), because each of these methods require several states to be calculated simulta-

neously, thus requiring larger D for the same precision.

[16] Z. G. Soos and S. Ramasesha, J. Chem. Phys. **90**, 1067 (1989).

[17] H. Saberi, A. Weichselbaum, and J. von Delft, Phys. Rev. B **78**, 035124 (2008).

[18] A. Rosch, T. A. Costi, J. Paaske, and P. Wolfle, Phys. Rev. B **68**, 014430 (2003); M. Garst, P. Wolfle, L. Borda, J. von Delft, and L. Glazman, *ibid.* **72**, 205125 (2005).

[19] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[20] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006).

[21] M. Yoshida, M. A. Whitaker, and L. N. Oliveira, Phys. Rev. B **41**, 9403 (1990).

[22] C. Guo, A. Weichselbaum, S. Kehrein, T. Xiang, and J. von Delft, Phys. Rev. B **79**, 115137 (2009).

[23] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).

[24] A. Freyn and S. Florens, Phys. Rev. B **79**, 121102(R) (2009).

[25] R. Zitko and T. Pruschke, Phys. Rev. B **79**, 085106 (2009).

[26] V. L. Campo and L. N. Oliveira, Phys. Rev. B **72**, 104432 (2005).

[27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipies in C*, 2nd ed. (Cambridge University Press, Cambridge, 1993).

[28] S. Nishimoto and E. Jeckelmann, J. Phys.: Condens. Matter **16**, 613 (2004).

P3. Discarded weight and entanglement spectra in the NRG

▶ Andreas Weichselbaum Phys. Rev. B **84**, 125130 (2011).

Discarded weight

Inspired by the DMRG and its common algebraic basis with the NRG, many of the concepts, well-established within the DMRG, can also be transferred to the NRG. A very prominent example in that respect is the concept of discarded weight. Within the DMRG, this appears naturally through the truncation based on Schmidt-decomposition given effective orthonormal basis sets. As such it endows the DMRG with a *direct absolute measure* for the accuracy of a given calculation. In contrast, the NRG lacks such a quantitative measure for accuracy. The only way to demonstrate convergence within the NRG had been backup runs at different number of states kept, showing that the results are stable. Compared to the DMRG, this situation appeared unsatisfactory. Specifically, in cases where one *cannot* afford significantly larger number of states kept due to the complexity of a given model, it appears desirable to have a *predictive measure* that also allows to get a good estimate on the accuracy of a given calculation within the NRG.

Now reduced density matrices [*cf.* Eq. (P3-11)] can be obtained within the NRG without much extra cost, since reduced density matrices are defined within the kept state spaces, by construction. The latter, however, can be used to motivate a *sensible estimate* for the discarded weight within the NRG. As shown in the paper, quite generally, a global discarded weight of $\varepsilon^{\mathrm{D}} \lesssim 10^{-12}$ suggests well-converged NRG data. By now, this argument has also been extensively applied to more complex three-channel setups (which served as the original motivation for this paper), with the same conclusion.

Entanglement spectra

Entanglement spectra, on the other hand, at first glance represent a completely disconnected discussion. Nevertheless, by definition, entanglement spectra are derived from the eigenspectra [*cf.* Eq. (P3-23)] of exactly the same reduced density matrices. While for discarded weight the emphasis is on the low-energy spectrum of these reduced density matrices, entanglement spectra, in contrast, analyze their high-energy spectrum. Therefore similar to the discarded weight, entanglement spectra are easily accessible within the NRG framework.

The entanglement spectra are calculated *w.r.t.* a specific iteration n. Consequently, given the iterative RG interpretation of the NRG, the entanglement spectra can also be combined into *entanglement flow diagrams* (in complete analogy to a standard energy flow diagram), which was demonstrated the first time in the attached paper. As it turns out, despite being a non-gapped system and despite being a pure analysis of the ground state wave function only, given energy scale separation, the *lowest energies* in the entanglement

flow diagram often nicely replicate features seen the energy flow diagram otherwise! On the other hand, there are cases where the energy- and the entanglement-flow-diagram clearly show *qualitatively* different structure even at low energies. The latter directly suggests the importance of the physics at much smaller energy scales when discussing the dynamics at higher energies, *i.e.* much larger than temperature. The latter had been pointed by Hofstetter (2000) for a specific model, namely the Kondo model in the presence of a magnetic field. The prescription of entanglement spectra, however, is more general, in that it allows to *predict* when non-trivial low-energy physics feeds back to larger energies.

Discarded weight and entanglement spectra in the numerical renormalization group

A. Weichselbaum

Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience,
Ludwig-Maximilians-Universität, D-80333 Munich, Germany

A quantitative criterion to prove and analyze convergence within the numerical renormalization group (NRG) is introduced. By tracing out a few further NRG shells, the resulting reduced density matrices carry relevant information on numerical accuracy as well as entanglement. Their spectra can be analyzed twofold. The smallest eigenvalues provide a sensitive estimate of how much weight is discarded in the low-energy description of later iterations. As such, the discarded weight indicates in a site-specific manner whether sufficiently many states have been kept within a single NRG run. The largest eigenvalues of the reduced density matrices, on the other hand, lend themselves to a straightforward analysis in terms of entanglement spectra, which can be combined into entanglement flow diagrams. The latter show strong similarities with the well-known standard energy flow diagram of the NRG, supporting the prevalent usage of entanglement spectra to characterize different physical regimes.

I. INTRODUCTION

The numerical renormalization group (NRG)[1] is a powerful method that provides a highly systematic nonperturbative approach to the wide realm of so-called quantum impurity systems. These consist of an arbitrary small quantum system (the *impurity*) in contact with a macroscopic noninteracting usually fermionic bath. Each part is simple to solve exactly on its own. In the presence of interaction at the location of the impurity, however, the combination of both gives rise to strongly correlated quantum-many-body phenomena.[2] Wilson's logarithmic coarse-graining of the bath leads to a semi-infinite chain with exponentially decaying couplings, which justifies the concept of energy scale separation. That is, the Wilson chain can be diagonalized iteratively by adding one site at a time and retaining the lowest M_K states only. The obvious question, however, is how many states should one keep on average for convergence in this procedure? At a given iteration there is no quantitative *a priori* measure that indicates how many low-energy states are required for a proper description of the remaining low-energy physics. Usually, the only way to check convergence within the NRG is by repeating the entire calculation and showing that the results no longer change when further increasing M_K. Therefore an NRG calculation is typically run somewhat blindly for some predetermined M_K.

This somewhat uncontrolled truncation in the NRG is in stark contrast to the situation in the density matrix renormalization group (DMRG).[3–5] DMRG is based on a (strictly) variational principle, and as such has a clean well-defined truncation of the state space for part of the system through the discarded weight in its reduced density matrix.[4] In contrast to the less suggestive plain number M_K of states kept, the discarded weight represents a reliable *quantitative* measure for the accuracy of a calculation. Within the DMRG, M_K can be easily adjusted according to some predefined threshold in the discarded weight instead. Motivated by DMRG then, an approximate similar criterion can be established within the NRG as will be shown in the following. The analysis requires

a slightly longer chain, as shown schematically in Fig. 1. With the extra n_0 sites traced out again from the ground state space of the enlarged system, this allows to estimate the discarded weight. The latter offers a quantitative convergence measure that is specifically of interest for numerically expensive models such as multichannel models, or models where the energy scale separation along the Wilson chain might be in question due to modifications in the discretized Hamiltonian. In either case, a small discarded weight provides a strong indication for converged NRG data.

Furthermore, the reduced density matrices generated for the evaluation of the discarded weight also allow a quite different analysis in terms of their *dominant* correlations. In particular, combining their entanglement spectra into *entanglement flow diagrams* offers a complementary view to the usual NRG energy flow diagram, which is entirely based on the analysis of the low-energy state space of a prior NRG run.

The paper is thus organized as follows. In Sec. I the essentials of the numerical renormalization group are revisited, including the construction of reduced density matrices. Section II then uses a specific set of reduced density matrices in the definition and analysis of the discarded weight within the NRG. Section III offers a complementary view on these reduced density matrices by analyzing their entanglement content in terms of entanglement spectra. Section IV, finally, summarizes and presents a brief outlook.
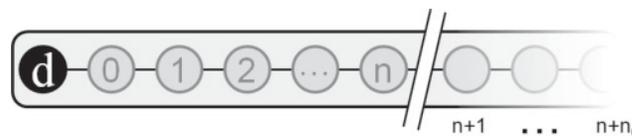

FIG. 1. Schematic depiction of tracing out the low-energy sector of the Wilson chain at iteration n by including and analyzing n_0 more NRG iterations. The impurity (dot) is entirely contained in the first site, while the bath is coarse-grained and mapped onto the remaining semi-infinite tight-binding chain of sites $n = 0,1,2,\ldots$.

P3.

A. Numerical renormalization group

Within the NRG, the continuum of the noninteracting bath of half-bandwidth W is logarithmically coarse-grained in energy space, followed by an exact mapping onto a semi-infinite so-called Wilson-chain.[1,6] The impurity space is coupled to the first site of this chain only, as depicted schematically in Fig. 1. The logarithmic coarse-graining is defined through the dimensionless discretization parameter $\Lambda > 1$. With the chemical potential at energy zero, the continuum of states in the energy intervals $\pm W[\Lambda^{-(n-z+1)}, \Lambda^{-(n-z)}]$ is effectively represented by single fermionic levels (coarse-graining), including an arbitrary z-shift with $z \in [0,1]$.[7–9] The subsequent exact mapping onto the semi-infinite chain (Lanczos tridiagonalization)[10] results in an effective tight-binding chain with the *exponentially* decaying hopping $t_n \sim \Lambda^{-n/2}$ between sites n and $n+1$. For sufficiently large Λ, typically $\Lambda \gtrsim 1.7$, this then justifies the essential NRG assumption of *energy scale separation*: by iterative diagonalization of the Wilson chain by adding one site at a time, large energies are considered first, with the (approximate) eigenstates at large energies *discarded* and considered unimportant in the description of the lower energy scales still to follow. Thus each site of the Wilson chain corresponds to an energy shell with a characteristic energy scale

$$\omega_n \equiv \frac{\Lambda^{z-1}(\Lambda - 1)}{\log \Lambda} W \Lambda^{-\frac{n}{2}}. \tag{1}$$

Here, the prefactor was chosen such that the rescaled couplings $\lim_{n\to\infty}(t_n/\omega_n) = 1$ quickly approach unity for longer Wilson chains for arbitrary Λ and z-shift, with the discretization following the prescription of Ref. 9 for a flat hybridization, i.e. $\Gamma(E) = \Gamma \theta(W - |E|)$.

With $\hat{H}_n$ the full Hamiltonian $\hat{H}$ of the Wilson chain up to and including site n, its low-energy eigenstates, are given by the NRG eigenstates $\hat{H}_n|s_n\rangle = E_s^n|s_n\rangle$. Complemented by an arbitrary state $|e_n\rangle$ for the remainder of the system following site n, the NRG assumption of energy scale separation can be summarized then in the following approximation[11]

$$\hat{H}|se\rangle_n \simeq E_s^n|se\rangle_n, \tag{2}$$

that is, the states $|se\rangle_n \equiv |s_n\rangle \otimes |e_n\rangle$ are, to a good approximation, also eigenstates of the entire Wilson chain. The energies E_s^n at iteration n are usually expressed relative to the ground state energy of that iteration, and rescaled by a factor $\frac{W}{2}(\Lambda + 1)\Lambda^{-n/2} \propto \omega_n$ to resolve the energy shell at iteration n. The resulting energies are referred to as *rescaled energies*. For fully fermionic systems, they typically show an intrinsic even-odd behavior. Thus combining the rescaled energies vs. even and odd iterations n separately, this results in the standard *energy flow diagrams* of the NRG.[1,6]

The approximate many-body eigenstates $|se\rangle_n$ are constructed iteratively, and therefore described in terms of matrix-product states.[5,12–14] Each iterative step results in a basis transformation, encoded in an A-tensor, that combines an existing effective basis $|s_n\rangle$ for the system up to and including site n with the state space $|\sigma\rangle$ of site $n+1$

$$|s_{n+1}\rangle = \sum_{s_n', \sigma_{n+1}} |s_n', \sigma_{n+1}\rangle \underbrace{\langle s_n', \sigma_{n+1}|s_{n+1}\rangle}_{\equiv A^{[\sigma_{n+1}]}_{s_n' s_{n+1}}}, \tag{3}$$

with $|s_n', \sigma_{n+1}\rangle \equiv |s_n'\rangle \otimes |\sigma_{n+1}\rangle$. The orthogonality of state spaces, $\langle s_{n+1}|s_{n+1}'\rangle = \delta_{ss'}$, directly implies the orthonormality relation for A-tensors[4]

$$\sum_{\sigma_{n+1}} A^{[\sigma_{n+1}]\dagger} A^{[\sigma_{n+1}]} = \mathbf{1}. \tag{4}$$

Without truncation, the dimension M_n of the state space $|s_n\rangle$ increases exponentially with the number of sites included, $M_n \sim d^n$, with d the dimension of a local Wilson site. Therefore the maximum number of states M_K, that one can maintain in a calculation, is quickly reached after $n_0 \simeq \log(M_K)/\log(d)$ iterations. For every subsequent iteration, the state space $|s_n\rangle$ is truncated by retaining the lowest M_K states in energy only. This leads to the distinction between $|s_n^K\rangle$ and $|s_n^D\rangle$ for kept and discarded states at iteration n, respectively. Correspondingly, this also splits the A-tensor into two parts, A_{KK} and A_{KD}, that propagate the state kept space from the previous iteration into the newly generated kept or discarded space, respectively.

The truncation criteria with respect to a fixed prespecified M_K can be softened in terms of an energy cutoff,[2] E_K, that is taken constant in rescaled energies. For a fair comparison for different z-shifts, it will be specified in units of the energy scale ω_n in Eq. (1). Since NRG data typically appears bunched at certain energies (e.g., see Fig. 3), E_K may hit a "gap" in the NRG spectrum at some iteration, and the last "bunch" of states included may lie, on average, at clearly smaller energies than E_K. Given the empirical importance of the first few NRG iterations, therefore as a safety measure, by default, E_K was taken by 20% larger for the very first iteration where truncation occurred, i.e. using $1.2E_K$ there with E_K specified in context. Typical values are in the range $E_K = 5, \ldots, 8$.

The model system considered in this paper is the well known standard single impurity Anderson model (SIAM)

$$H_N^{\text{SIAM}} = \sum_\sigma \varepsilon_{d\sigma} \hat{n}_\sigma + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow} + \sum_\sigma \sqrt{\frac{2\Gamma}{\pi}}(\hat{d}_\sigma^\dagger \hat{f}_{0\sigma} + \text{H.c.})$$

$$+ \sum_\sigma \sum_{n=0}^{N-1} t_n(\hat{f}_{n,\sigma}^\dagger \hat{f}_{n+1,\sigma} + \text{H.c.}), \tag{5}$$

with the operators $\hat{d}_\sigma^\dagger$ ($\hat{f}_{n\sigma}^\dagger$) creating a particle with spin $\sigma \in \{\uparrow, \downarrow\}$ at the impurity (at site n in the bath), respectively, having $\hat{n}_{d\sigma} \equiv \hat{d}_\sigma^\dagger \hat{d}_\sigma$. The energy $\varepsilon_{d\sigma} \equiv \varepsilon_d - \frac{B}{2}(\hat{n}_{d\uparrow} - \hat{n}_{d\downarrow})$ is the spin-dependent level position of the impurity in the presence of a magnetic field B. Furthermore, U is the on-site Coulomb interaction and Γ the hybridization of the impurity with the bath. All parameters will be specified in units of the bandwidth $W := 1$ in context with the figure panels. The bath in Eq. (5) is already represented in terms of a Wilson chain,[1] described by the semi-infinite tight binding chain ($N \to \infty$) with exponentially decaying hopping amplitudes $t_n \sim \Lambda^{-n/2}$. In practice, N can be taken finite, with $\hat{H}_n$ describing the Wilson chain up to and including site $n \leqslant N$.

Charge and spin are conserved in the SIAM in Eq. (5), where, however, only the Abelian part of the symmetries is included in the calculations. Hence the number of states M_K directly refers to the actual number of states kept in a calculation (in contrast to the dimension of reduced multiplet

spaces with non-Abelian symmetries). Similarly, also the discussion of the entanglement spectra further below will refer to the Abelian symmetry labels which also applies when non-Abelian symmetries are broken. Note that while, in general, a particle-hole symmetric impurity setting will be used, this can be easily broken by applying a (small) gating potential to the impurity level. Moreover, the $SU(2)$ spin symmetry, in fact, will be broken explicitly by the application of an external magnetic field.

B. Density matrices

The NRG eigenbasis of Eq. (2) with respect to the discarded space forms a complete many-body eigenbasis.[11] Initially introduced for the feat of real-time evolution within the NRG, this eigenbasis is actually applicable and tractable more generally within the NRG framework.[15] In particular, this allows the clean calculation of correlation functions in terms of the full density matrix (FDM) in the many-body eigenbasis,[12] in that

$$\hat{\rho}(T) \equiv \frac{1}{Z} e^{-\beta \hat{H}} \cong \frac{1}{Z} \sum_{nse} e^{-\beta E_s^n} |se\rangle_{nn}^{\text{DD}}\langle se|, \qquad (6)$$

with $\beta \equiv 1/k_B T$ for arbitrary temperatures T, using *non-rescaled* energies E_s^n relative to a common energy reference, by construction of a thermal density matrix. Equation (6) can be rewritten as $\hat{\rho}(T) \equiv \sum_n w_n(T) \hat{\rho}_n(T)$, i.e. a normal-ized distribution $\sum_n w_n = 1$ of the density matrices $\hat{\rho}_n(T)$ generated in the basis of iteration n.[12] For a given tem-perature T, the distribution w_n is strongly peaked around iteration n_T that corresponds to the energy scale of temper-ature. Hence temperature essentially terminates the Wilson chain.

In this paper, however, mainly reduced density matrices de-rived from ground states will be considered, hence temperature is essentially zero. More generally then, consider an arbitrary density matrix defined in the many-body basis $|s_n\rangle$ of iteration n in either kept or discarded space, $X \in \{K, D\}$,

$$\hat{\rho}_n^{[X]} \equiv \sum_{s_n s_n' \in X} \rho_{s_n s_n'}^{[X]} |s_n\rangle \langle s_n'|, \qquad (7)$$

where $\rho_n^{[X]}$ (i.e. without the hat) represents the space of matrix elements $\rho_{s_n s_n'}^{[X]}$. The prototypical and well-known operation on such a density matrix is tracing out the last site n,[11,12,15–17]

$$\hat{\rho}_{n-1}^{[K]} = \sum_{\substack{s_{n-1}, s_{n-1}' \\ \sigma_n}} \left(A_{KX}^{[\sigma_n]} \rho_n^{[X]} A_{KX}^{[\sigma_n]\dagger} \right)_{s_{n-1} s_{n-1}'} |s_{n-1}\rangle \langle s_{n-1}'|$$

$$\equiv \hat{\mathcal{P}}_n \hat{\rho}_n^{[X]}, \qquad (8)$$

written as a matrix product of the matrices $A_{KX}^{[\sigma_n](\dagger)}$ and $\rho_n^{[X]}$ in the first line. Equation (8), in the following referred to as *backward update*, introduces the notational shorthand $\hat{\mathcal{P}}_n$ for the bilinear product of the A- and A^*-tensor at site n, that acts as a linear superoperator on the density matrix $\hat{\rho}_n$. The corresponding contraction pattern is shown in a simple graphical depiction in Fig. 2. By construction, the backward update of a density matrix in Eq. (8) always results in a density matrix in the kept space of the earlier iteration, and with Eq. (4)
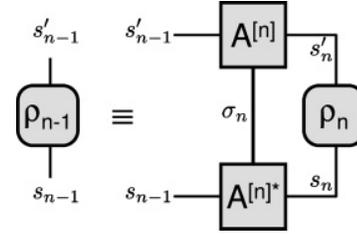

FIG. 2. *Backward update* of a given density matrix ρ_n at iteration n. Blocks represent data spaces, lines correspond to indices. The lines connecting different blocks are contracted indices (i.e. indices summed over), such as σ_n, s_n, and s_n', while open lines represent open indices (e.g., the indices s_{n-1} and s_{n-1}').

representing a complete positive map, Eq. (8) clearly also preserves the properties of a density matrix.

II. DISCARDED WEIGHT WITHIN THE NRG

The standard notion of NRG is that it zooms in toward the low-energy sector of a given many-body Hamiltonian, while iteratively discarding states at higher energies. Having a semi-infinite chain, this can continue to arbitrarily small energy scales, which enables NRG to resolve dynamically generated small energy scales as they appear, for example, in the context of Kondo physics. From a variational point of view for matrix-product states, this implies that the cost function can be identified as

$$\lim_{N \to \infty} \langle s_N | H_N | s_N \rangle \to \text{MIN}, \qquad (9)$$

yielding the ground state $|0\rangle_\infty$ of the semi-infinite Wilson chain. For a sufficiently long chain of total length N then included in a given calculation, the state $|0\rangle_N$ will be referred to as the overall ground state of this Wilson chain. In fact, the cost function in Eq. (9) is well captured within the NRG through its principle of energy scale separation.[18]

If at a given iteration within the NRG states essentially decouple with respect to the low-energy state space still to follow, these states will quickly and efficiently be discarded as high-energy states. The truncation toward the low-energy sector also implies that the state space at large energies is necessarily more crudely resolved, consistent with the coarser discretization there. The lowest M_K states kept at a given iteration n then are important for the correct description of the low-energy sector still to come. However, there is no real quantitative *a priori* measure to indicate whether the number M_K of states to be kept is appropriate. Conversely, however, at a given iteration n one can ask whether all states kept a few iterations earlier were actually important. This question can be answered entirely within the kept spaces of these iterations, and hence is numerically cheap to analyze.

A. Construction of reduced density matrices

Consider the actual ground state space G at some arbitrary but fixed iteration n'. In general, it may be $g_{n'}$-fold degenerate,

87

hence consider its fully mixed density matrix,

$$\hat{\rho}_{0,n'} \equiv \frac{1}{g_{n'}} \sum_{s \in G} |s_{n'}\rangle\langle s_{n'}|. \qquad (10)$$

By construction, the number of eigenvalues of $\hat{\rho}_{0,n'}$ unequal zero, i.e. its Schmidt rank, is equal to $g_{n'}$. Now, tracing out the last iteration n', i.e. the lowest-energy scale included in $\hat{\rho}_{0,n'}$, is equivalent to the back-propagation $\hat{\rho}_0^{[n'-1;1]} \equiv \hat{\mathcal{P}}_{n'}\hat{\rho}_{0,n'}$ in Eq. (8). Through this operation, the Schmidt rank will rise, in general, by a factor of d, with d the state space dimension of a Wilson site. Repeating this process iteratively, this allows to trace out the n_0 smallest energy shells in $\hat{\rho}_{0,n'}$. Thus with $n' = n + n_0$, this leads to the reduced density matrix,

$$\hat{\rho}_0^{[n;n_0]} \equiv \left(\prod_{l=n+1}^{n+n_0} \hat{\mathcal{P}}_l \right) \hat{\rho}_{0,n+n_0} \equiv \sum_{ss'}^{M_K} \rho_{ss'}^{[n;n_0]} |s_n^K\rangle\langle s_n'^K|, \quad (11)$$

which, by construction, is defined in the *kept* space of iteration n. The Schmidt rank will grow quickly (i.e., exponentially) in this process, until after n_0 iterations, with

$$n_0 \gtrsim \mathrm{ceil}[\log(M_K)/\log(d)] \quad (n_0 \ll N), \qquad (12)$$

it reaches the full dimension M_K of the kept space. Typically, n_0 is much smaller compared to the full length N of the Wilson chain considered, and conversely also specifies the initial number of NRG iterations in a forward direction that can be typically performed without truncation. For the definition of the discarded weight below, it is sufficient to stop the back propagation of $\hat{\rho}_{0,n+n_0}$ at this point.

The reduced density matrix $\hat{\rho}_0^{[n;n_0]}$ generated in Eq. (11) is, in general, not diagonal in the energy eigenbasis $|s_n^K\rangle$, since through the traced out lower-energy sites it does know about an enlarged system. Its eigenvectors are described by a unitary transformation $u_{rs'}^{[n;n_0]}$ within the NRG eigenstates kept at iteration n,

$$|r_{n;n_0}\rangle \equiv \sum_{s'} u_{rs'}^{[n;n_0]} |s_n'^K\rangle,$$

$$\text{with} \qquad \hat{\rho}_0^{[n;n_0]} |r_{n;n_0}\rangle = \rho_r^{[n;n_0]} |r_{n;n_0}\rangle, \qquad (13)$$

where the index r shall refer to the eigenstates of the reduced density matrix, in contrast to the index s for the energy eigenstates. Here, the eigenvalue $\rho_r^{[n;n_0]}$ describes the importance of a specific linear superposition of NRG eigenstates at iteration n for the low-energy description of later iterations.

This offers two routes for the analysis of the density matrices $\hat{\rho}_0^{[n;n_0]}$. (i) Adhering to the energy eigenbasis of the NRG, the importance of the kept state $|s_n^K\rangle$ at eigenenergy E_s^n for the later low-energy physics is given by the expectation value

$$\rho_s^{[n;n_0]} \equiv \langle s_n^K | \hat{\rho}_0^{[n;n_0]} | s_n^K \rangle, \qquad (14)$$

i.e. the diagonal matrix elements $\rho_{ss}^{[n;n_0]}$. Alternatively, (ii) using the eigenbasis of the reduced density matrices, the weights of these states are given by the eigenvalues $\rho_r^{[n;n_0]}$, while now their energies are given by the expectation values

$$E_r^{[n;n_0]} \equiv \langle r_{n;n_0} | \hat{H}_n | r_{n;n_0} \rangle. \qquad (15)$$

Both routes will be analyzed and compared in the following. However, the actual eigendecomposition of the reduced density matrices will be preferred for the remainder of the paper as explained.

In either case, a set of states i with (average) energy E_i is given together with their respective (average) weight ρ_i that represents the states importance for later iterations. For the first [second] route above this data is given by $(E_s^n, \rho_s^{[n;n_0]})$ $[(E_r^{[n;n_0]}, \rho_r^{[n;n_0]})]$, respectively. Given that the reduced density matrix $\hat{\rho}_0^{[n;n_0]}$, by construction, exists in the kept space only, therefore all states i refer to the kept space or a linear superpositions thereof. Moreover, for every iteration, the weights ρ_i are normalized, that is, they are positive and add up to 1, while by combining data from different iterations, the energies E_i are always specified in rescaled units.

The resulting data (E_i, ρ_i) then is clearly correlated. It is analyzed threefold, (i) in terms of the average distribution of the rescaled energies E_i

$$\nu(E) \cong \frac{1}{N'} \sum_{n=1}^{N}{}' \sum_{E < E_i < E+dE} 1, \qquad (16)$$

(ii) the average distribution of the weights ρ_i,

$$\nu(\rho) \cong \frac{1}{N'} \sum_{n=1}^{N}{}' \sum_{\rho < \rho_i < \rho+d\rho} 1, \qquad (17)$$

and (iii) their average dependence on each other

$$\rho(E) \cong \frac{1}{N' dE} \sum_{n}{}' \sum_{E < E_i < E+dE} \rho_i \qquad (18a)$$

$$\cong \kappa e^{-\kappa E}. \qquad (18b)$$

Here some appropriate linear (logarithmic) binning of the data is assumed with energy (weight) intervals dE ($d\rho$), respectively. In particular, the densities in Eqs. (16) and (17) are clearly dependent on these binning intervals, which therefore will be properly indicated in the subsequent plots. The prime in the summation and the normalization indicates that only those iterations n are included where state space truncation occurred, i.e. typically $n \gtrsim n_0$. The total number of these iterations is given by N'. With chosen normalization then, the sum over the binned $\nu(E)$ and $\nu(\rho)$ data both yield the average number of kept states, while the integrated weight distribution $\rho(E)$ in Eq. (18a) is normalized to 1 since $\mathrm{tr}(\rho) \sim \int_0^\infty \rho(E)\,dE = 1$. As will be seen later, the weight distribution $\rho(E)$ typically shows a clear exponential decay with a characteristic exponent κ, as indicated already in Eq. (18b), with the prefactor chosen such that it also preserves normalization.

1. Energy eigenbasis

The correlation between the eigenenergies E_s^n and their corresponding weights $\rho_s^{[n;n_0]}$ is plotted as a scatterplot in the main panel of Fig. 3. The model analyzed is the SIAM in Eq. (5) in the Kondo regime using a fixed number of kept states, with all parameters specified in the figure caption. The weights $\rho_s^{[n;n_0]}$ clearly diminish exponentially with energy, which is intuitively expected as a consequence of energy scale separation within the NRG. The integrated weight
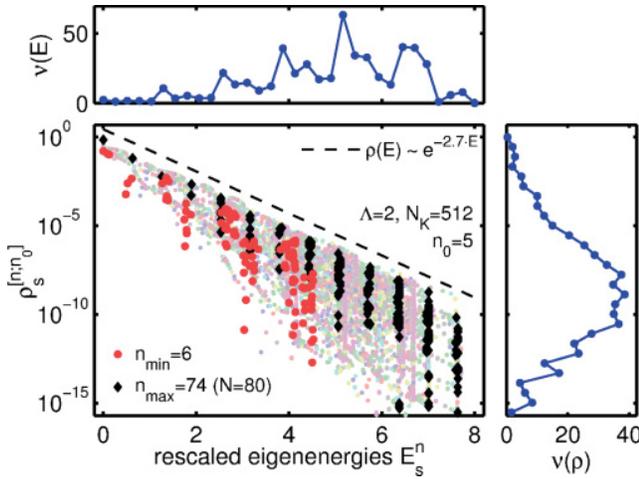
P3.

FIG. 3. (Color online) Weight distribution of energy eigenstates over full NRG run at fixed $M_K = 512$ for the SIAM [Eq. (5): $U = 0.20, \epsilon_d = -U/2, \Gamma = 0.01$]. The main panel shows the rescaled eigenenergies E_s^n vs. their weights $\rho_s^{[n;n_0]}$ as in Eq. (14). Data is shown only for those iterations where truncation occurred, with data from the same iteration shown in the same color. The two iterations with smallest (largest) energy range, n_{max} (n_{min}), are highlighted in strong colors [black diamonds (red crosses)], respectively, while light colors are used for all other iterations. The top [right] panel shows the energy [weight] distribution $\nu(E)$ [$\nu(\rho)$], Eq. (16) [Eq. (17)], respectively, for the data in the main panel, with matching energy E [weight ρ] axis. The binning referred to in the text to Eqs. (16) and (17) is indicated by the intervals between the data points in the top and right panel.

distribution $\rho(E)$ [dashed black line, cf. Eq. (18a)] shows a clear exponential decay with an exponent $\kappa \simeq 2.7$. As seen in Fig. 3, this distribution clearly also serves as an upper bound of the weights $\rho_s^{[n;n_0]}$ at a given energy.

The upper panel in Fig. 3 shows the distribution $\nu(E)$ in Eq. (16) of the energies E_s^n plotted in the main panel (matching horizontal axis). This distribution shows a strong increase with energy E, consistent with the notion that the many-body phase space grows quickly as the available energy for excitations becomes larger. Toward large energies, eventually, the data is necessarily truncated to the finite number M_K of kept states, which leads to a drop in the density $\nu(E)$. The exact boundary with respect to energy is somewhat blurred, though, since in given case fixed M_K allows the energy range to vary for different iterations n. The right panel of Fig. 3, on the other hand, shows the distribution $\nu(\rho)$ in Eq. (17) of the weights $\rho_s^{[n;n_0]}$ plotted in the main panel (matching vertical axis). This distribution is peaked around the largest weights $\rho_s^{[n;n_0]}$ for the largest energies E_s^n.

The data in the main panel of Fig. 3 is typically bunched around a set of energies for a fixed iteration n. This is also reflected in the distribution $\nu(E)$ in the upper panel of Fig. 3, and is due to the discretization of the model. Moreover, two iterations are highlighted in strong colors. These correspond to the iterations whose energy range is smallest [$n_{min} = 6$, red bullets] or largest [$n_{max} = 74$, black diamonds]. Intuitively, the largest numerical error is expected from iterations such as n_{min} since, through Eq. (18b), stopping at premature energies

directly translates to the largest missing (i.e. *discarded*) weight in the density matrix. As an aside, this serves as a strong argument in favor of truncation with respect to a fixed energy cutoff E_K rather than a fixed number M_K of states. Fixed E_K, however, also introduces more noise to the data in particular for higher-lying states. Hence both truncations will be used and pointed out in context.

The weights $\rho_s^{[n;n_0]}$ in the main panel of Fig. 3 show significant vertical spread, which translates into a pronounced tail toward exponentially smaller ρ in the distribution $\nu(\rho)$ in the right panel. For a given energy E therefore, many of the states have *orders of magnitude* lower weight than the topmost weights close to $\rho(E)$ in the main panel. This indicates that the energy representation with its corresponding diagonal weights $\rho_s^{[n;n_0]}$ is not necessarily the optimal basis to analyze accuracy. Moreover, note that using the energy eigenbasis $|s_n\rangle$ with energies E_s^n in the analysis of the reduced density matrices actually mingles the energy scales of an effectively larger system $\hat{H}_{n+n_0}$ with the basis generated w.r.t. $\hat{H}_n$ only.

2. Eigenbasis of reduced density matrices

From the point of view of a variationally optimal representation of the ground state space of an enlarged system, on the other hand, one is directly led to the *eigenspectrum* of the reduced density matrix, as exemplified within DMRG.[3] The analysis of Fig. 3 therefore is repeated for the same underlying Wilson chain, yet with two modifications: (i) the eigen decomposition of the reduced density matrices in Eq. (13) together with Eq. (15) is used instead of the energy eigenbasis, and furthermore (ii) the NRG truncation criterion is based on a fixed energy cutoff, $E_K = 6$. The results are shown in Fig. 4, with striking quantitative differences compared to Fig. 3. The spread in the scatterplot is significantly
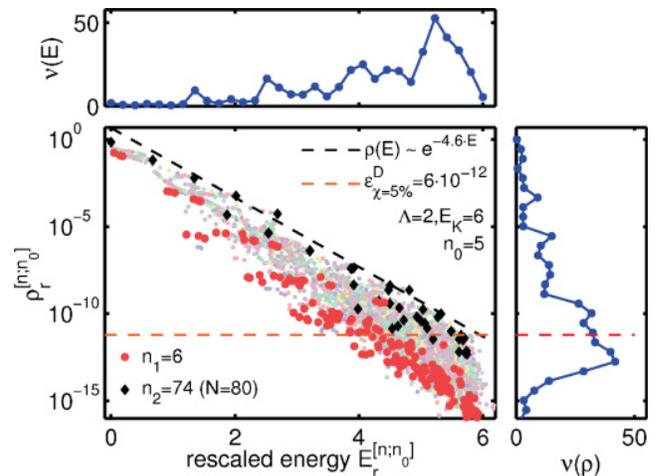
FIG. 4. (Color online) Similar analysis as in Fig. 3 (see caption there for further information) for the same underlying Hamiltonian, except that the eigenspectrum of the reduced density matrices in Eq. (13) was used together with Eq. (15) and a fixed energy cutoff $E_K = 6$. Similar to Fig. 3, only those iterations are shown where truncation occurred (same color for data from the same iteration), with the same two iterations highlighted as in Fig. 3, indicated by n_1 and n_2. The estimate for the overall discarded weight $\varepsilon_{\chi=5\%}^D \simeq 6 \cdot 10^{-12}$ as defined in Eq. (21) is indicated by the horizontal dashed line.

narrowed, and overall, the data decays much faster with $\kappa \simeq 4.6$ [cf. Eq. (18b)]. Therefore this leads to a clearly improved separation of the actually relevant states for the subsequent description of the lower-energy scales. This suggests that many of the NRG eigenstates, as their energy increases, loose importance much faster as compared to Fig. 3, despite the relatively large diagonal weights ρ_s in the density matrix still seen there. In a sense, the weights there represent mere matrix-elements in a nondiagonal representation.

The iterations highlighted in Fig. 4 are the same iterations as in Fig. 3. Given a fixed energy cutoff $E_K = 6$ here, however, both have a comparable energy range (hence the altered notation n_1 and n_2), with the number M_K of kept states varying from ~1000 at very early iterations (in particular iteration n_1), down to ~250 at late iterations (such as iteration n_2). Note also the markedly fewer data points seen for iteration n_2. This is only partly due to the reduced number of states, as there are also large systematic (approximate) degeneracies at the strong-coupling Kondo fixed point already reached at this iteration. This results in many of the black diamonds lying indistinguishably on top of each other (see also the discussion on entanglement spectra later).

As seen from above discussion, rather than taking the energy eigenstates $|s_n\rangle$ and the corresponding diagonal matrix elements $\rho_s^{[n;n_0]}$ (Fig. 3), the eigenvalues $\rho_r^{[n;n_0]}$ of the reduced density matrix $\hat{\rho}_0^{[n;n_0]}$ do represent a clearly better choice for the analysis of accuracy or entanglement in the system (Fig. 4), and thus will be used henceforth. This prescription, in general, also shows a more systematic exponential decay all the way down to numerical double precision noise (10^{-16}), with the decay rate κ of $\rho(E)$ roughly independent of the discretization parameter Λ.

3. Definition of discarded weight

With the motivation above, the definition of the discarded weight is based on the eigen decomposition of the reduced density matrices $\hat{\rho}_0^{[n;n_0]}$ in Eq. (11), using the combined data of Eqs. (13) and (15). In terms of Fig. 4, adding more states to the calculation essentially extends the data to larger energies and smaller weights, while the large-weight low-energy sector already remains widely intact. Therefore the largest discarded weight, i.e. the weight missing by states not included and hence not available, can be estimated to a good approximation, up to an overall prefactor, by the smallest weights in the kept space which are easily accessible. Given the exponential decay of the weights together with the residual spread in the data as seen in Fig. 4, the discarded weight at given iteration n can thus be defined through the *average* weights $\rho_r^{[n;n_0]}$ for the highest energies $E_r^{[n;n_0]}$ in the kept space,

$$\varepsilon_{n;n_0}^{K_\chi} \equiv \left\langle \rho_r^{[n;n_0]} \right\rangle_{E_r^{[n;n_0]} \geqslant (1-\chi)\max(E_r^{[n;n_0]})}. \tag{19a}$$

The parameter $\chi \ll 1$ is considered small, yet is chosen large enough (typically $\chi \simeq 0.05$) to average over the residual spread of weights. Alternatively and for comparison, an even simpler measure in terms of the minimum eigenvalue of $\hat{\rho}_0^{[n;n_0]}$ will be considered,

$$\varepsilon_{n;n_0}^{K} \equiv \min\left(\rho_r^{[n;n_0]}\right), \tag{19b}$$

which no longer makes any explicit reference to energies. Note that even though $\varepsilon_{n;n_0}^{K_{(\chi)}}$ or $\varepsilon_{n;n_0}^{K}$, written $\varepsilon_{n;n_0}^{K_{(\chi)}}$ in short, are purely determined within the kept space, they clearly represent a sensible estimate for the discarded weight at iteration n (i.e., $\varepsilon_n^{D_{(\chi)}} \sim \varepsilon_{n;n_0}^{K_{(\chi)}}$) defined as the fraction of relevant state space missing from the latter description of the low-energy physics. If no truncation has occurred at iteration n, however, such as typically for the first $n < n_0$ iterations, of course, then there is no truncation error either, hence $\varepsilon_n^{D_{(\chi)}} = 0$ for these iterations.

In summary, the discarded weight ε_n^D at iteration n is defined as follows,

$$\varepsilon_n^{D_{(\chi)}} \equiv \begin{cases} \varepsilon_{n;n_0}^{K_{(\chi)}} & \text{in the presence of truncation} \\ 0 & \text{without truncation at iteration } n. \end{cases} \tag{20}$$

Here $\varepsilon_{n;n_0}^{K_{(\chi)}}$ can be determined efficiently by including and analyzing n_0 further NRG iterations within the kept space, where typically $n_0 \ll N$ [cf. Eq. (12)]. The overall discarded weight $\varepsilon_{(\chi)}^D$ of a full NRG run then is taken, for simplicity, as the largest discarded weight per iteration,

$$\varepsilon_{(\chi)}^D \equiv \max_n \left(\varepsilon_n^{D_{(\chi)}}\right). \tag{21}$$

Using $\chi = 5\%$ as in Eq. (19a), the discarded weight for the NRG run in Fig. 4 is estimated by $\varepsilon_\chi^D \simeq 6 \cdot 10^{-12}$, indicated by the horizontal dashed line. As seen from Fig. 4, the overall discarded weight ε_χ^D for an NRG run essentially coincides with $\rho(E)$ at the largest energies within the kept space. On the other hand, ε^D [i.e. without the usage of χ based on the plain *minimum* eigenvalue of the reduced density matrices $\hat{\rho}_0^{[n;n_0]}$, cf. Eq. (19b)] will, in general, lie a (constant) few orders of magnitude lower, as it happens, for example, for the data in Fig. 4. Nevertheless, as will be shown in the following, up to an overall global prefactor the discarded weight based on either, ε^D or ε_χ^D, both behave in an essentially similar fashion.

B. Application

The discarded weight $\varepsilon_{(\chi)}^D$ defined in Eq. (21) sensitively depends on the number M_K of states kept or the energy threshold E_K. From Fig. 4 one expects a strongly diminishing discarded weight with increasing M_K or E_K, a quantitative analysis of which is presented in Figs. 5 and 6 for the SIAM. Figure 5 analyzes the dependence of the discarded weight $\varepsilon_{(\chi)}^D$ on the number M_K of states kept. As seen in Fig. 5(a), the discarded weight $\varepsilon_{(\chi)}^D$ strongly decays with M_K, with minor variations when a new Wilson shell is fully included without truncation, e.g. at $M_K \in \{256, 1024\}$. With Fig. 5(a) being a log-log plot, the decay of the discarded weight with M_K rather resembles a polynomial convergence, yet with very large power (on the order of 10). The reason for the slower than exponential decay is due to the strong increase in the density of states $\nu(E)$ of the full many-body eigenspectrum with increasing E as discussed with Figs. 3 and 4.

Together with the analysis of the discarded weight in Fig. 5, an independent physical check for convergence is provided by the numerically computed conductance g_0 in units of $2e^2/h$ shown in Fig. 5(b). The conductance was calculated via the (spin-resolved) spectral function $A_{(\sigma)}(\omega) = \int \frac{dt}{2\pi} e^{i\omega t} \langle \{\hat{d}_\sigma(t), \hat{d}_\sigma^\dagger\} \rangle_T$ of the impurity level, with
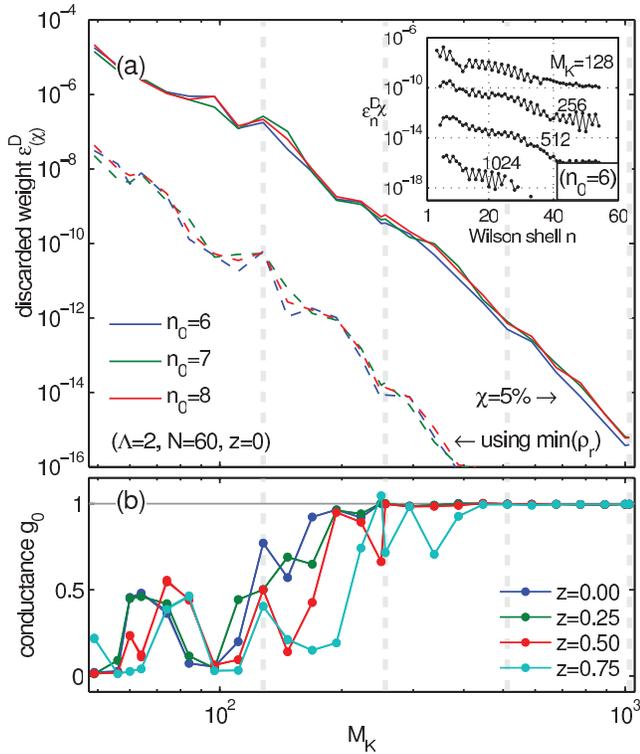
FIG. 5. (Color online) Discarded weight ε^D for the SIAM [Eq. (5): $U = 0.20$, $\epsilon_d = -U/2$, $\Gamma = 0.01$ (same parameters as in Fig. 3), with $T_K \simeq 1.23 \cdot 10^{-5}$]. Panel (a) shows the discarded weight $\varepsilon^D_{(\chi)}$ defined in Eq. (21) vs. M_K using $n_0 \in \{6,7,8\}$. The data $\varepsilon^D_{\chi=5\%}$ is shown in solid lines, while the data based on the minimum eigenvalue of $\hat{\rho}^{[n;n_0]}_0$ [cf., Eq. (19b)] is shown in dashed lines. The distribution of the discarded weight $\varepsilon^{D_\chi}_n$ along the Wilson chain is shown in the inset for $M_K \in \{128,256,512,1024\}$, also marked by the vertical dashed lines in the main panels. Panel (b) shows the conductance g_0 vs. M_K in units of $2e^2/h$ while using a set of shifted discretizations, with the z-values as specified. Convergence in the conductance toward the expected unitary limit is seen for $M_K \gtrsim 400$, i.e. $\varepsilon^D_\chi \lesssim 10^{-12}$.

$g_0 = \pi\Gamma \int d\omega(-\frac{\partial f}{\partial \omega})A(\omega)$. Here the Fermi function $f(\omega)$ and the spectral function $A(\omega)$ are evaluated at small but finite temperature $T \simeq 6 \cdot 10^{-8}$, which is much smaller than the Kondo temperature of $T_K \simeq 1.23 \cdot 10^{-5}$ for given parameter set and corresponds to the energy scale close to the end of the Wilson chain, having $\Lambda = 2$ and $N = 60$. Expecting $g_0 = 1$ for the symmetric SIAM, the data in Fig. 5 indicate convergence for $M_K \gtrsim 400$. The data for smaller M_K is not yet converged, and therefore (strongly) depends on numerical details, such as non-averaged z-shifts.[8,9]

With M_K being constant, the energy of the topmost kept states can vary significantly with Wilson shell n, which directly also leads to a clear dependence of the discarded weight $\varepsilon^D_{(\chi)}$ on n. This is shown in the inset to Fig. 5(a) for the set of different values of M_K marked in the main panels by the vertical dashed lines. The discarded weight ε^D_χ clearly varies over more than three orders of magnitude within a single NRG run, irrespective of the actual M_K. In particular, one can see that earlier iterations dominate the discarded weight ε^D_χ for physical reasons. In the strong-coupling regime for
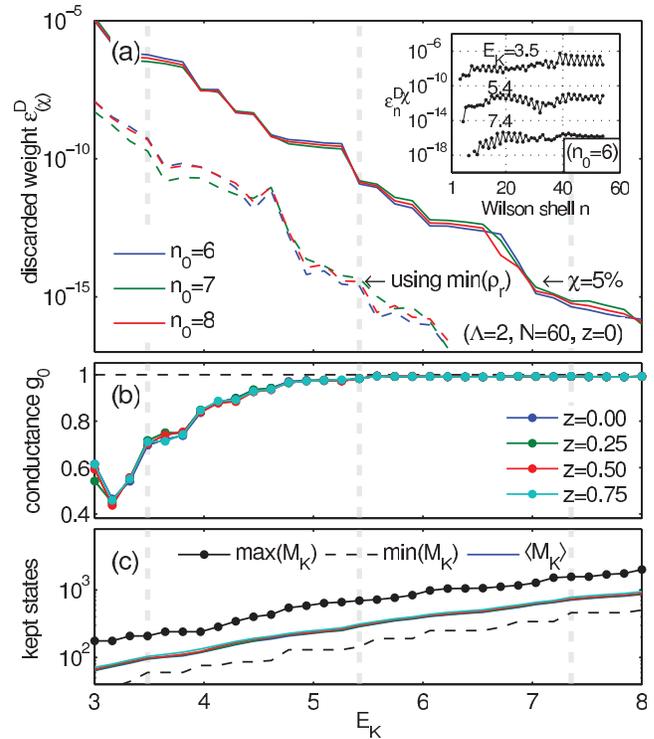
FIG. 6. (Color online) Similar analysis as in Fig. 5, yet for truncation with respect to fixed energy E_K. For several values of E_K, marked by the vertical dashed lines in the main panels, the distribution of the discarded weight $\varepsilon^{D_\chi}_n$ along the Wilson shell n is shown in the inset to panel (a). With M_K allowed to vary over a wider range, panel (c) shows the correlation of M_K with E_K, plotting average, minimum, and maximum of M_K along the Wilson chain. For the average M_K, data for different z-shifts is shown [several lines on top of each other, with same color coding as in panel (b)].

$n \gtrsim n_K$ (with iteration $n_K \simeq 35$ corresponding to the energy scale of T_K), the discarded weight is smallest, while for the intermediate free orbital or local moment regime for $n \lesssim n_K$, these regimes require a *larger* number of states for comparable numerical accuracy from a physical point of view, indeed.

Given the underlying energy scale separation of the NRG, a straightforward way to obtain a more equally distributed $\varepsilon^{D_{(\chi)}}_n$ is achieved using an energy cutoff E_K, as demonstrated in Fig. 6 for exactly the same system as in Fig. 5 otherwise. For the values of E_K indicated by the vertical dashed lines in the main panels, the inset to Fig. 6(a) shows the distribution of $\varepsilon^{D_\chi}_n$. By construction, the discarded weight is, up to even-odd oscillations, clearly more uniformly distributed over the Wilson shells as compared to the case of fixed M_K in Fig. 5(a). The discarded weight in Fig. 6(a) clearly diminishes exponentially with E_K, yet with pronounced intermediate plateaus since the discrete eigenenergies within an NRG run are usually bunched around certain energies. The corresponding average M_K as function of E_K, nevertheless, follows a rather smooth monotonic behavior, as shown in Fig. 6(c). Given fixed E_K, however, clear variations of M_K are seen within a given NRG run, hence also the smallest and largest M_K are shown in Fig. 6(c). Ignoring iterations without truncation, in given example, typically the largest M_K is required at early iterations, while the smallest M_K are

91

encountered in the strong coupling regime at late iterations $n \gtrsim n_K$.

The calculated conductance shown in Fig. 6(b) converges clearly more uniformly with increasing E_K as compared to Fig. 5(b). In particular, it indicates converged NRG data for $E_K \gtrsim 5.5$, which corresponds to $\varepsilon_\chi^D \lesssim 10^{-12}$. Therefore in both settings, for constant M_K in Fig. 5 as well as for constant E_K in Fig. 6, convergence of the physical data is found for a similar discarded weight of $\varepsilon_\chi^D \lesssim 10^{-12}$ with a negligible dependence on n_0. This value therefore is considered a sufficient bound in accuracy to capture the main physics, with other quantities such as the NRG energy flow diagram already also well converged.

Alternatively, using the plain minimum of the eigenvalues of the reduced density matrices in Eq. (19b), this leads to convergence for $\varepsilon^D \lesssim 10^{-16}$. Given that ε^D refers to the minimum eigenvalue in the kept space, ε^D consistently lies about three orders of magnitudes lower than ε_χ^D and is considered a lower bound to the actual discarded weight. While ε^D fluctuates slightly more strongly compared to ε_χ^D owing to the fact that it is not an averaged quantity such as ε_χ^D, it nevertheless follows a similar consistent picture in terms of convergence with the number M_K of states kept or the energy E_K used for truncation. In this sense, either discarded weight, ε^D as well as ε_χ^D, can be used quite generally as a quantitative measure, indeed, to demonstrate accuracy within the NRG. To avoid confusion, however, it shall be made clear which one is used.

III. ENTANGLEMENT SPECTRA

The reduced density matrices $\hat{\rho}_0^{[n;n_0]}$ clearly also carry *physical* information in terms of entanglement along the Wilson chain. This is provided by the high end of their spectral decomposition. There the exact details of the largest eigenvalues of $\hat{\rho}_0^{[n;n_0]}$ are of interest, which do vary with n_0 over a wider range depending on the underlying physics. Hence, in the following, the actual entanglement spectra will be calculated with respect to the reduced density matrices $\hat{\rho}_0^{[n]}$ of the overall ground state of the system,

$$\hat{\rho}_0^{[n]} \equiv \lim_{n_0 \to \infty} \hat{\rho}_0^{[n;n_0]} \simeq \hat{\rho}_0^{[n;N-n]}. \tag{22}$$

The length N of the Wilson chain is taken sufficiently large, such that the energy scale of the last iteration N is much smaller than any other energy scale in the system. Temperature is therefore essentially zero. For comparison, also the *truncated* entanglement spectra will be calculated from $\hat{\rho}_0^{[n;n_0]}$ for finite small n_0, with n_0 specified in context. Motivated by the discussion following Eq. (6), the later analysis can be linked to finite temperature settings.

1. General definition

The partitioning of the Wilson chain into two parts, the chain up to and including site n (part A), and the traced out remainder of the system (part B) is generic. In particular, this allows to make use of the recently introduced entanglement spectra (ES)[19] for the physical characterization of a given wave function. Here these entanglement spectra provide a powerful tool for the systematic analysis of the physical correlations in the reduced density matrices $\hat{\rho}_0^{[n]}$ in Eq. (22).

Consider a given wave function of some system partitioned into parts A and B. The reduced density matrix $\hat{\rho}_A \equiv \mathrm{tr}_B(\rho)$ is obtained by tracing out part B of the overall density matrix ρ. Within this setting, the entanglement spectrum is defined as the spectrum of the fictitious Hamiltonian $\hat{H}_\rho^A$,[19]

$$\hat{\rho}_A =: \exp\left(-\hat{H}_\rho^A\right).$$

One may assume an effective inverse temperature $\beta := 1$ to make contact with a thermal density matrix. This β also sets the (otherwise arbitrary) energy scale in the *per se* dimensionless $\hat{H}_\rho^A$. With $\hat{\rho}_A$ a positive operator, the entanglement spectrum ξ_r is defined as the eigenvalues of $\hat{H}_\rho^A$, that is,

$$\xi_r := -\log \rho_r, \tag{23}$$

with ρ_r the spectral decomposition of the reduced density matrix $\hat{\rho}_A$. Particular information can be read off from the entanglement spectrum as soon as there is a rich amount of quantum numbers specifying the entanglement levels and when entanglement gaps appear which separate a low-lying generic set of levels from irrelevant background correlations.[19–21] The spectra ρ_r and ξ_r are independent of whether A or B is traced out, while, of course, they are dependent on the specific choice of the partitioning. For entanglement spectra, the partitioning typically occurs in real space for gapped systems, analyzing the *edge* of the thus created boundary, while for gapless systems momentum space is preferred.[20] The second case then is consistent with the systematic NRG prescription of energy scales based on the underlying discretization in energy (momentum) space.

By construction, the dominant correlations between systems *A* and *B* correspond to the lowest *entanglement energies* ξ_r, while weaker correlations will rise to higher energies. By tracing out a major part of the system, entanglement spectra provide significantly more information, say, than just the entanglement entropy between *A* and *B*. In particular, it has been shown that it provides fingerprints of the underlying physics, and as such allows to characterize the physical nature of a given wave function.[19,20] This analysis is therefore entirely targeted at a given (ground state) wave function, without any further reference to an underlying physical Hamiltonian that it may have originated from.

2. Application to NRG

The general concept of the entanglement spectra can be readily transferred to the NRG. At each iteration n, the reduced density matrix $\hat{\rho}_0^{[n]}$ in Eq. (22) is computed and diagonalized, with its eigenspectrum mapped onto the entanglement spectrum in Eq. (23). Collecting these spectra and plotting them vs. iteration index n for even and odd iterations separately, the result will be referred to as *entanglement flow diagram*, in complete analogy to the standard energy flow diagrams of the NRG. For comparison, also the truncated entanglement spectra for finite small n_0 will be analyzed, which in their combination will be referred to as *truncated* entanglement flow diagram. In either case, the entanglement spectra are obtained in a *backward* sweep, purely based on the iterative low-energy
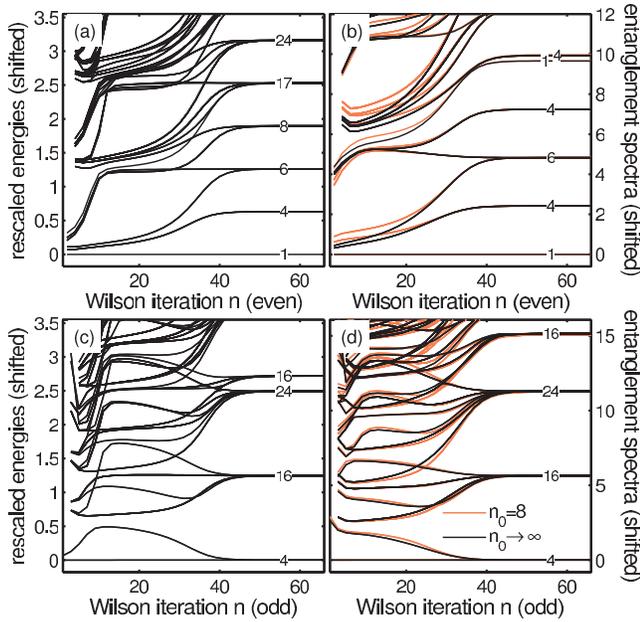
FIG. 7. (Color online) Comparison of the standard NRG energy flow diagram (left panels) to the *entanglement flow diagram* (right panels) for the symmetric SIAM [$U = 0.2$, $\varepsilon_d = -U/2$, $\Gamma = 0.01$, $T_K = 1.2 \cdot 10^{-5}$; $\Lambda = 2$, $M_K = 512$, $N = 80$], with top (bottom) panels for even (odd) iterations, respectively. In addition to the actual entanglement flow diagram obtained from the ground state of the last iteration at $N = 80$ [black lines], the truncated entanglement flow diagram is shown, using $n_0 = 8$ orange (gray) lines. For better comparison with the energy flow diagram, the entanglement spectra (right panels) are also shifted at every iteration with respect to the smallest entanglement energy $\min(\xi)$. The y-scale of the entanglement spectra was adjusted to best match the energy fixed point spectrum in the left panels. Degeneracies of energies at large n (i.e., lines lying indistinguishably on top of each other) are specified by the numbers on top of the lines in all panels.
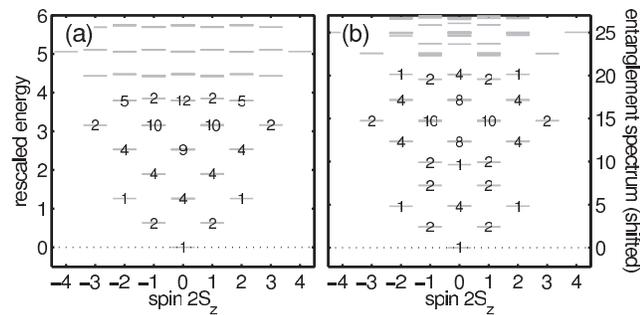

FIG. 8. Comparison of spin-resolved fixed point spectra for the symmetric SIAM in Fig. 7 in the SC regime ($n = 60$). Panel (a),(b) show the energy [entanglement] fixed point spectrum, respectively, vs. spin symmetry quantum number S_z. For all low-energy multiplets the underlying (approximate) degeneracy is indicated. The entanglement spectrum is shifted w.r.t. to its lowest energy and scaled to match the energy fixed point spectrum in panel (a).

Hilbert-space decomposition of a prior NRG run in terms of the A-tensors in Eq. (3). This is in contrast to the energy flow diagram, which is calculated with increasing shell index n in a *forward* sweep making explicit reference to the Hamiltonian.

The entanglement spectra were calculated for the symmetric SIAM in the absence of magnetic field. The resulting entanglement flow diagram is presented in Fig. 7 together with a direct comparison to the standard NRG energy flow diagram. The data is plotted for even (odd) Wilson shells n in the upper (lower) panels, respectively. The energy flow diagram, shown in the left panels, clearly distinguish the well-known physical regimes of the SIAM, namely the free orbital regime (FO; $n \lesssim 10$), the local moment regime (LM; $10 \lesssim n \lesssim n_K$), and the strong coupling regime (SC; $n \gtrsim n_K$), where $n_K \simeq 35$ corresponds to the energy scale of the Kondo temperature $T_K = 1.2 \cdot 10^{-5}$, having $\Lambda = 2$. All degeneracies for $n > n_K$ are explicitly specified in Fig. 7. In particular, for even iterations, the ground state is unique throughout, i.e. the Kondo singlet for $n > n_K$ Fig. 7(a)], while for odd iterations the ground state space at small energies is four-fold degenerate due to the particle-hole symmetric parameter set Fig. 7(c)].

Interestingly, a very similar picture emerges from the entanglement flow diagram in the right panels, Figs. 7(b) and 7(d) (black lines). For comparison, also the truncated entanglement spectra are shown using $n_0 = 8$ [orange (gray) lines], which in given case converge rapidly, in fact, exponentially, with increasing $n_0 \lesssim 10$ toward the actual entanglement flow diagram. These then mimic the energy flow diagram in the left panels over a wide range. For example, the convergence toward the Kondo fixed point occurs around similar iterations, and even the degeneracies of the lowest states of the energy flow diagram are exactly recovered by the entanglement spectra. The same also holds on the symmetry-resolved level, as demonstrated in Fig. 8 for the even iteration $n = 60$ (see later discussion). Nevertheless, looking more closely, a few notable qualitative differences of the entanglement flow diagrams in the right panels of Fig. 7 are seen compared to the energy flow diagrams in the left panels. Overall, the entanglement flow diagrams appear shifted by about five iterations to larger energies. This can be understood, considering that the entanglement spectra are calculated for enlarged systems together with the rapid convergence with increasing n_0 in given case. However, there are further pronounced differences with the energy flow diagram for the earliest iterations in the FO regime, $n \lesssim 10$.

These differences in the entanglement flow diagram can be significantly enhanced by turning on a magnetic field on the order of the Kondo temperature, as shown in Fig. 9 for $B = 1.6 \, T_K$. This corresponds to the energy scale at iteration $n_B \simeq 32$, given $\Lambda = 2$. The magnetic field has been chosen such that, for late iterations $n \gg n_B$, the fixed point spectrum for even and odd iterations become essentially the same [compare the low-energy fixed point spectra in panel a(b) to panel c(d) of Fig. 9, respectively]. Due to the magnetic field, the Kondo singlet (previously the unique state at even iterations) is largely destroyed for $n \gtrsim n_B$ with a sizable magnetization at the impurity. Clearly, the NRG eigenbasis at early iterations $n < n_B$ does not yet know about the small energy physics to come, e.g. the small $B \sim T_K$ applied in given case. Therefore the energy flow diagram essentially remains unaltered there,

when compared to the case without magnetic field in panels (a) and (c) of Fig. 7. The flow changes strongly only starting from the energy scale of the magnetic field value, i.e. for $n > n_B$ where it moves into a different fixed-point spectrum. In particular, there also emerges a unique state now in the energy flow diagram for odd iterations for $n \gg n_B$, i.e. the symmetry broken spinful state favored by the magnetic field.

By including magnetic field, the entanglement flow diagram shows pronounced differences from the energy flow diagram for $n \lesssim n_B$, which includes large portions of the LM regime. While the energy spectrum up to and including site n is ignorant of the low-energy physics to come, this very low-energy physics is captured by the reduced density matrices and thus reflected in the entanglement flow diagram.[22]

Consider the entanglement spectra derived from the overall ground state in Figs. 9(b) and (d) [black lines]. In Fig. 9(d) the ground state remains unique throughout, i.e. remembers the symmetry broken magnetic state, determined at much lower energy scales, all the way up to the largest energies. Within the split-up lowest-energy space with subsequent degeneracies [1-2-1] in Fig. 9(d) for $n \gg n_B$ (to be called [1-2-1] configuration), the first and second excited states cross each other with decreasing n leading to a [1-1-2] configuration for small n, i.e. large energies. Nevertheless, the singly degenerate excited state clearly remains split off, and does not merge with the ground state, which is in strong contrast to the energy flow diagram in Fig. 9(c) with a [2-2] configuration for $n \ll n_B$. This degeneracy in the ground state space that is ignorant of the small magnetic field is partly reflected only in the *truncated* entanglement flow diagram. Using small n_0 [orange (gray) lines in Fig. 9(d)], this eventually also misses the low-energy physics. Therefore these spectra in Fig. 9(d) eventually are also in a [2-2] configuration for the smallest n, with a more irregular transient behavior with increasing n. A similar trend is also observed for even iterations in Figs. 9(a) and (b). While the ground state remains unique for all iterations in both panels, the entanglement flow in Fig. 9(b) tends to split off the excited levels right above the lowest [1-2-1] state space configuration for small n. For the truncated entanglement flow, on the other hand, the lines of these excited levels remain entangled with higher excitations, which is similar to the situation in the energy flow diagram in Fig. 9(a).

Nevertheless, the low-energy fixed-point spectra for $n \gg n_B$ again agree well for both the energy and entanglement flow diagram in Fig. 9, which again also holds for the symmetry-resolved spectra, as demonstrated for the even iteration $n = 60$ in Fig. 10. This agreement in the spectra of the stable low-energy fixed point, present in both the nonmagnetic as well as the magnetic case, is understood as a generic feature. There both the energy eigenstates as well as the reduced density matrices are deeply rooted in the low-energy physics, i.e. of the overall ground state of the system at $T \rightarrow 0$, and hence present a consistent description of the system.

The detailed structure of the energy fixed point spectra provides clear physical information.[1,2] This includes, for example, phase shifts if a Fermi-liquid point of view is supported as is the case for the SIAM. This then directly explains all of the splittings and degeneracies in the low-energy sector of the energy fixed point spectra. For example, consider the energy spectrum in Fig. 8(a) for the fully symmetric SIAM in the

FIG. 9. (Color online) Comparison of the standard NRG energy flow diagram (left panels) to the *entanglement flow diagram* (right panels) for the SIAM at finite magnetic field (same analysis as in Fig. 7, otherwise, see caption there for details, with same model parameters, except $B = 2 \cdot 10^{-5} \simeq 1.6\,T_K$).

FIG. 10. Comparison of spin-resolved fixed point spectra for the SIAM at finite magnetic field in Fig. 9 at the even iteration $n = 60$ (similar analysis as in Fig. 8 otherwise).

nonmagnetic case. Note that while spin-resolved spectra are shown in Fig. 8, in given case the charge-resolved spectra would look exactly the same due to particle-hole symmetry. With the spectra shown for an even iteration, the ground state is unique (i.e., represents the Kondo singlet with $S_z = 0$). The first excited states for $S_z = +\frac{1}{2}$ correspond to an extra particle with spin-up or a hole with spin-down. Given particle-hole symmetry, both processes have the same energy $\delta/2 = 0.63$ (in rescaled energy units), and hence are two-fold degenerate, indicated by the number on top of the level in Fig. 8. By symmetry, the same excitations exist for $2S_z = -1$, leading to the [2-2] degeneracy (four states) in the lowest excitations in Fig. 8(a). The next higher excitation combines two of the above processes. This leads to a total of six excitations, all with energy δ and distributed over $2S_z \in \{-2, 0, +2\}$. Here two of the excitations at $2S_z = 0$ correspond to the extraction or annihilation of two particles with opposite spin.

This fully explains the [1-4-1] degeneracy of the excited states at energy $\delta = 1.26$ in Fig. 8(a), and also the combined six-fold degeneracy seen in the energy flow diagram seen at this energy in Fig. 7(a). The argument can be continued along similar lines to explain the [4-4] (eight states) and [4-9-4] (17 states) degenerate subspaces of the next higher excitations. Excitations with even higher energy eventually have missing levels due to NRG truncation.

The same analysis as for the energy spectra, however, cannot be applied with equal rigor to the entanglement spectra. While the ground state 1 and the lowest [2-2] and [1-4-1] excitations in Fig. 8(b) fully agree in symmetries, degeneracy, and also in the precise relative level spacing, the next higher [4-4] excitation in Fig. 8(a) is broken up in Fig. 8(b), with some of the levels shifting to higher entanglement energy. Nevertheless, the degenerate set [2-10-10-2] further up in energy again equally appears for both energy and entanglement spectra.

The same analysis as in Fig. 8 is repeated for the magnetic case in Fig. 10 for the same even iteration $n = 60$. Despite the rather different level spectrum for large n in the flow diagram in Fig. 9, the actual spin-resolved fixed point spectrum is qualitatively very similar to the nonmagnetic case in Fig. 8. Aside from an overall tilt of the level structure, all degeneracies and level positions of the lower part of the energy spectrum in Fig. 8(a) are again fully described by elementary single-particle excitations. The underlying reason for this similarity of the fixed points spectra in the magnetic and nonmagnetic case is that, apart from the (screened) impurity spin, the system is well described by an effective Fermi-liquid picture. With the low-energy fixed point spectra well reflected in the entanglement spectra, a similar tilt in the level structure is also observed in Fig. 10(b) when compared to Fig. 8(b). Note, for example, that to the lower left of the spectrum the same [1-2-1], as well as the [2-4-2, 2-4-2] state sequence with increasing energy is seen.

IV. SUMMARY AND OUTLOOK

The reduced density matrices of the NRG by tracing out the low-energy sector have been analyzed in detail. The low end of their eigenspectra was used to estimate the discarded weight $\varepsilon_{(\chi)}^{D}$ in Eqs. (19)–(21) as a quantitative and site-resolved measure of the accuracy within the NRG. While, in principle, the same reduced density matrices could also be utilized as the basis for an altered truncation criteria similar to the DMRG, this, however, requires sufficiently large M_K to start with. In practice, this is sufficiently close to a truncation with respect to an energy cutoff E_K. Either way, all of this can be easily and quickly checked using the proposed analysis in terms of the discarded weight which provides a useful quantitative tool.

Furthermore, the dominant correlations of the reduced density matrices were analyzed in terms of their entanglement spectra. Due to the NRG flow toward small energy scales, these spectra can be combined into entanglement flow diagrams. There different physical regimes can be identified similar to the standard NRG energy flow diagrams. Considering that the entanglement spectra are obtained solely based on the wave function, the agreement of the low-energy fixed point spectra is stunning. A possible larger disagreement at higher energies (i.e., for earlier Wilson shells), on the other hand, depends on the specific physical situation. Given the NRG background, this appears to suggest the following. For all energy shells (iterations) n where the entanglement spectrum is quantitatively comparable to the NRG energy spectrum for the lowest set of states, the reduced density matrices themselves are not crucially important in the description of the system. Instead, they may be replaced by thermal density matrices in the NRG eigenbasis. In a sense, by tracing out the low-energy sector, the resulting reduced density matrices maintain an approximate thermal character, with implications to thermalization at a given energy shell.[24] For energy shells with a qualitative difference between the energy and entanglement spectra, however, the reduced density matrices are crucially important to capture the correct physics in the NRG calculation that explicitly uses data from such energy shells.

A detailed analysis of the deeper connection and the explicit differences between the energy and the entanglement spectra appears interesting, yet is out of the scope of this paper. In particular, it also appears instructive to analyze the entanglement spectra for non-Fermi liquid systems such as the symmetric two-channel Kondo model, as the analysis presented in this paper suggests a strong physical connection of the entanglement spectra to the underlying physics.

[1] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[2] R. Bulla, T. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

[3] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[4] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[5] U. Schollwöck, Ann. Phys. (NY) **326**, 96 (2011).

[6] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980).

[7] M. Yoshida, M. A. Whitaker, and L. N. Oliveira, Phys. Rev. B **41**, 9403 (1990).

[8] L. N. Oliveira, Braz. J. Phys. **22**, 155 (1992).

[9] R. Žitko and T. Pruschke, Phys. Rev. B **79**, 085106 (2009).

[10] J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000).

[11] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005).

[12] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[13] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, Phys. Rev. B **80**, 165117 (2009).

[14] S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997).

[15] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006).

[16] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000).

[17] A. I. Tóth, C. P. Moca, Ö. Legeza, and G. Zaránd, Phys. Rev. B **78**, 245109 (2008).

[18] H. Saberi, A. Weichselbaum, and J. von Delft, Phys. Rev. B **78**, 035124 (2008).

[19] H. Li and F. D. M. Haldane, Phys. Rev. Lett. **101**, 010504 (2008).

[20] R. Thomale, D. P. Arovas, and B. A. Bernevig, Phys. Rev. Lett. **105**, 116805 (2010).

[21] R. Thomale, A. Sterdyniak, N. Regnault, and B. A. Bernevig, Phys. Rev. Lett. **104**, 180502 (2010).

[22] Note that it was exactly this kind of reasoning, for example, that led to the success of density-matrix-based NRG methods for dynamical properties, starting with DM-NRG.[16] For the SIAM, for example, the reduced density matrices, while not crucially important in the absence of magnetic field, are absolutely essential for the correct description of spectral correlations at finite magnetic field. There it is intuitively clear that the spin-resolved spectral function $A_\sigma(\omega)$ at the impurity at small temperatures $T \ll T_K$ in the presence of a magnetic field $B > T_K$ redistributes significant spectral weight at *large* energies $|\omega| \gg T_K$, which accounts for the breakup of the Kondo singlet. Within the NRG, this thus translates into a *feedback* from small to large energies, which is captured correctly only after including the reduced density matrices for the remainder of the system.[16] Yet contributions from all NRG shells are required to cover the full spectral range of dynamical correlation functions.[23] A clean combined prescription for this was finally provided by the FDM-NRG approach[12] based on complete basis sets.[11,15]

[23] T. A. Costi, Phys. Rev. B **55**, 3003 (1997).

[24] D. Poilblanc, Phys. Rev. B **84**, 045120 (2011).

P4. Kondo Decoherence: Finding the Right Spin Model for Iron Impurities in Gold and Silver

▶ T. A. Costi, L. Bergqvist, A. Weichselbaum, J. von Delft, T. Micklitz, A. Rosch, P. Mavropoulos, P. H. Dederichs, F. Mallet, L. Saminadayar, and C. Bäuerle
Phys. Rev. Lett. **102**, 056802 (2009).

The very original system where the Kondo effect was observed the first time, was iron impurities in gold or silver [de Haas et al. (1934)]. With the Kondo physics understood since Kondo (1964), nevertheless, the underlying microscopic model of above experimental system has remained controversial until very recently. Having partially filled d-orbital levels, it appears very plausible that the plain simple impurity Anderson model is oversimplistic. Through a range of previous studies, nevertheless, a consensus had been reached that the microscopic model corresponds to a fully screened Kondo model. The strategy of the attached paper then was three-fold.

(1) An experimental study of quasi-one-dimensional samples of gold or silver with a carefully controlled impurity density through ion-implantation. These samples were then analyzed *w.r.t.* two clearly distinctive physical quantities, namely resistivity and dephasing rate as a function of temperature.

(2) Using ab-initio calculations, the environment of the iron impurity as well the crystal field splitting of the otherwise 5-fold degenerate d-orbitals was estimated in the presence of spin-orbit coupling. This allowed to obtain an effective low-energy Hamiltonian. The results showed that the d-orbitals are split into the 2-fold degenerate e_g and the 3-fold degenerate t_{2g} set, with a somewhat larger density of states of the latter at the fermi energy.

(3) NRG calculations finally allowed to simulate the strongly-correlated Kondo regime of the model Hamiltonians derived from the ab-initio calculations.

Given a fully screened Kondo model, the ab-initio calculations therefore permit setups with $n = 2$ (e_g) or $n = 3$ (t_{2g}) spinful bath channels, screening an impurity with spin $S = n/2$. From the NRG point of view, the results were complimented still by $n = 1$, for comparison. The experimental resistivity data was used to calibrate each model with the NRG simulations, which then allowed a parameter-free prediction for the dephasing rate. As seen in Fig. P4-4, the best agreement is seen for $n = 3$, which thus strongly suggests the fully screened three-channel model as the actual microscopic model, with an impurity spin of $S = 3/2$. The latter is also supported by the ab-initio calculations, as well as other experiments using Mössbauer spectroscopy.

For above NRG calculations, the consistent and accurate approach to arbitrary temperatures provided by the FDM-NRG was absolutely essential. Nevertheless, the NRG calculations were performed using abelian symmetries only, which due to the complexity of the model did not provide *fully* converged bare-spectral functions then. Only by using

the self-energy improved spectral functions,[10] a good agreement with the Friedel-sum-rule was obtained which thus served as indication that the physics is well-captured. This has been confirmed recently using the full-fledged non-abelian NRG, with a further detailed comparison to magnetoresistivity data to be published in a separate publication (Markus Hanl et al., in preparation).

Kondo Decoherence: Finding the Right Spin Model for Iron Impurities in Gold and Silver

T. A. Costi,[1,2] L. Bergqvist,[1] A. Weichselbaum,[3] J. von Delft,[3] T. Micklitz,[4,7] A. Rosch,[4] P. Mavropoulos,[1,2]
P. H. Dederichs,[1] F. Mallet,[5] L. Saminadayar,[5,6] and C. Bäuerle[5]

[1]*Institut für Festkörperforschung, Forschungszentrum Jülich, 52425 Jülich, Germany*
[2]*Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany*
[3]*Physics Department, Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience,
Ludwig-Maximilians-Universität München, 80333 München, Germany*
[4]*Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany*
[5]*Institut Néel-CNRS and Université Joseph Fourier, 38042 Grenoble, France*
[6]*Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France*
[7]*Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
(Received 8 October 2008; published 3 February 2009)

We exploit the decoherence of electrons due to magnetic impurities, studied via weak localization, to resolve a long-standing question concerning the classic Kondo systems of Fe impurities in the noble metals gold and silver: which Kondo-type model yields a realistic description of the relevant multiple bands, spin, and orbital degrees of freedom? Previous studies suggest a fully screened spin S Kondo model, but the value of S remained ambiguous. We perform density functional theory calculations that suggest $S = 3/2$. We also compare previous and new measurements of both the resistivity and decoherence rate in quasi-one-dimensional wires to numerical renormalization group predictions for $S = 1/2$, 1, and 3/2, finding excellent agreement for $S = 3/2$.

Introduction.—The Kondo effect of magnetic impurities in nonmagnetic metals, e.g., Mn, Fe, or Co in Cu, Ag, or Au, first manifested itself in the early 1930s as an anomalous rise in resistivity with decreasing temperature, leading to a resistivity minimum [1]. In 1964 Kondo explained this effect [2] as resulting from an antiferromagnetic exchange coupling between the spins of localized magnetic impurities and delocalized conduction electrons.

However, for many dilute magnetic alloys a fundamental question has remained unresolved to this day: which effective low-energy Kondo-type model yields a realistic description of the relevant multiple bands, spin, and orbital degrees of freedom [3]? Cases in point are Fe impurities in Au and Ag, the former being the very first magnetic alloy known to exhibit an anomalous resistivity minimum [1]. Previous attempts to fit experimental data on, for example, Fe impurities in Ag (abbreviated as AgFe) with exact theoretical results for thermodynamics, by assuming a fully screened low-energy effective Kondo model [4,5], have been inconclusive: specific heat data are absent and the local susceptibility of Fe in Ag obtained from Mössbauer spectroscopy [6] indicated a spin of $S = 3/2$ while a fully screened $S = 2$ model has been used to fit the temperature dependence of the local susceptibility [7].

A promising alternative route towards identifying the model for Fe in Au or Ag is offered by studying transport properties of high purity quasi-one-dimensional mesoscopic wires of Au and Ag, doped with a carefully controlled number of Fe impurities by means of ion implantation [8–13]. Magnetic impurities affect these in two different ways. Besides causing the aforementioned

resistivity anomaly, they also make an anomalous contribution $\gamma_m(T)$ to the electronic phase decoherence rate $\gamma_\phi(T)$ measured in weak (anti)localization: an itinerant electron which spin flip scatters off a magnetic impurity, leaves a mark in the environment, and thereby suffers decoherence. By checking model predictions for both effects against experimental observations over several decades in temperature, decoherence can thus be harnessed as a highly sensitive probe of the actual form of the effective exchange coupling. Experiments along these lines [11,12] were consistent with a Kondo model in which the impurity spin is fully screened and inconsistent with underscreened or overscreened Kondo models [11]. A consistent description of both resistivity and decoherence measurements using the simplest fully screened Kondo model, the $S = 1/2$ single-channel Kondo model, was, however, not possible: different Kondo scales were required for fitting the resistivity and decoherence rates [11,12].

In this Letter we address the above problem via the following strategy. (i) We carry out density functional theory calculations within the local density approximation (LDA) for Fe in Au and Ag to obtain information that allows us to prescribe a low-energy effective model featuring three bands coupling to impurities with spin $S = 3/2$. (ii) We calculate the resistivity $\rho_m(T)$ and decoherence rate $\gamma_m(T)$ due to magnetic impurities for three fully screened Kondo models, with $n = 2S = 1$, 2, and 3, using Wilson's numerical renormalization group (NRG) approach. (iii) We compare these predictions to experimental data: extracting the characteristic Kondo temperature T_K^S for each choice of n from fits to $\rho_m(T)$ and using these T_K^S to obtain

parameter-free predictions for $\gamma_m(T)$, we find that the latter agree best with experiment for $n = 3$.

LDA calculations.—Fully relaxed density functional theory calculations employing the VASP code [14] showed that low-symmetry Fe configurations (split interstitials [15]) are energetically unfavorable: Fe impurities prefer an environment with cubic symmetry. As the calculated defect formation energy of an Fe interstitial was found to be about 2 eV higher than the energy of a substitutional defect, we discuss the latter case in the following. This is in line with experiments on Fe implantation in AgAu alloys, where only substitutional Fe-atoms are found [16].

Figure 1 shows the *d*-level local density of states of substitutional Fe in Ag and Au, obtained by spin-polarized calculations using a 108 atom supercell, with similar results being found for a 256 atom supercell. The cubic local symmetry leads to e_g (doublet) and t_{2g} (triplet) components with a e_g-t_{2g} splitting, $\Delta \gtrsim 0.15$ eV in LDA [Figs. 1(a) and 1(b)]. The widths Γ_{e_g} and $\Gamma_{t_{2g}}$ of the e_g and t_{2g} states close to the Fermi level (E_F) are of the order of 1 eV, resulting from a substantial coupling to the conduction electrons. The large t_{2g} component at E_F persists within LDA + U [Figs. 1(c) and 1(d)] using $U = 3$ eV and a Hund's coupling $J_H = 0.8$ eV.

The spin and orbital moments are given in the legends of Fig. 1 (spin-polarized Korringa-Kohn-Rostoker calculations yielded similar values [16]): Within spin-polarized LDA a large spin moment μ_S of approximately 3–3.1μ_B forms spontaneously, consistent with Mössbauer measurements that give 3.1–3.2μ_B for the spin moment for Fe in Ag [6]. In contrast, there is no tendency for a sizable orbital

FIG. 1 (color online). The *d*-level local density of states (LDOS) of substitutional Fe in Ag and Au within spin-polarized LDA (a),(b) and LDA + U (c),(d), with inclusion of spin-orbit interactions, and showing the e_g [gray (red)] and t_{2g} (black) components of the *d*-level LDOS of FeAg (left-hand panels) and FeAu (right-hand panels). Majority (minority) contributions are shown positive (negative). Legends give the spin (μ_S) and orbital (μ_L) magnetic moments in units of the Bohr magneton μ_B and the splitting (Δ) between the e_g and t_{2g} components of the *d*-level LDOS.

moment (or a Jahn-Teller distortion). The small orbital moments μ_L of $<0.1\mu_B$ (consistent with experimental results [17]) arise only due to weak spin-orbit coupling. We therefore conclude that the orbital degree of freedom is quenched on an energy scale set by the width $\Gamma_{t_{2g}}$ of the t_{2g} orbitals. Moreover, since the spin-orbit splitting of the localized spin in the cubic environment is proportional to μ_L^4, it is tiny, well below our numerical precision of 0.01 meV, and, therefore, smaller than the relevant Kondo temperatures.

Low-energy effective models.—The above results justify formulating an effective low-energy model in terms of the spin degree of freedom only. The large spin moment μ_S of 3–3.1μ_B suggests an effective spin $S = 3/2$. Our LDA results thus imply as effective model a spin-3/2 three-channel Kondo model, involving local and band electrons of t_{2g} symmetry. An alternative possibility, partially supported by the large (almost itinerant) t_{2g} component at E_F, would be to model the system as a spin 1 localized in the e_g orbitals, that is perfectly screened by two conduction electron channels of e_g symmetry. This spin is then also coupled to (almost itinerant) t_{2g} degrees of freedom via the ferromagnetic J_H. At high temperature, the latter binds an itinerant t_{2g} spin 1/2 to the local spin 1 to yield an effective spin 3/2, consistent with the spin moment of 3–3.1μ_B obtained within LDA, whereas in the low temperature limit, the irrelevance of J_H under renormalization [4] leads to the stated effective spin-1, two-band model. Though such a model is well justified only for $J_H \ll \Gamma_{t_{2g}}$, which is not the case here where $J_H \sim \Gamma_{t_{2g}}$, our LDA results do not completely exclude such a model. To identify which of the models is most appropriate, we shall confront their predictions with experimental data below.

We thus describe Fe in Ag and Au using the following fully screened Kondo model:

$$H = \sum_{k\alpha\alpha} \varepsilon_k c_{k\alpha\sigma}^\dagger c_{k\alpha\sigma} + J\sum_\alpha \mathbf{S} \cdot \mathbf{s}_\alpha. \tag{1}$$

It describes *n* channels of conduction electrons with wave vector k, spin σ, and channel index α, whose spin density $\Sigma_\alpha \mathbf{s}_\alpha$ at the impurity site is coupled antiferromagnetically to an Fe impurity with spin $S = n/2$. Whereas our LDA results suggest $n = 3$, we shall also consider the cases $n = 1$ and 2.

NRG calculations.—The resistivity $\rho_m(T)$ and decoherence rate $\gamma_m(T)$ induced by magnetic impurities can be obtained from the temperature and frequency dependence of the impurity spectral density [18,19]. We have calculated these quantities using the NRG [20–22]. While such calculations are routine for $n = 1$ and 2 [21], they are challenging for $n = 3$. Exploiting recent advances in the NRG [20] we were able to obtain accurate results also for $n = 3$ (using a discretization parameter of $\Lambda = 2$ and retaining 4500 states per NRG iteration).

Figure 2 shows $\rho_m(T)$ and $\gamma_m(T)$ for $n = 2S = 1$, 2, and 3. For $T \gtrsim T_K^S$, enhanced spin-flip scattering causes both
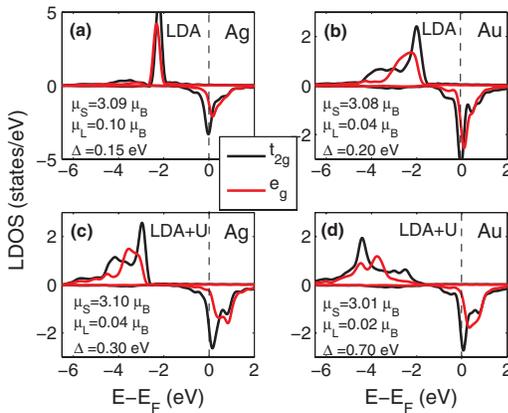
$\rho_m(T)$ and $\gamma_m(T)$ to increase with decreasing temperature. For $T \lesssim T_K^S$ the effective exchange coupling becomes so strong that the impurity spins are fully screened by conduction electrons, forming spin singlets, causing $\rho_m(T)$ to saturate to a constant and $\gamma_m(T)$ to drop to zero. While these effects are well known [2,8–12], it is of central importance for this study that they depend quite significantly on $S = n/2$, in such a way that conduction electrons are scattered and decohered more strongly the larger the local spin S: With increasing S, (i) both resistivities and decoherence rates decay more slowly with T at large temperatures ($\gg T_K^S$), and (ii) the "plateau" near the maximum of $\gamma_m(T)$ increases slightly in maximum height γ_m^{max} and significantly in width. These changes turn out to be sufficient to identify the proper value of S when comparing to experiments below.

Comparison with experiment.—We compared our theoretical results for $\rho_m(T)$ and $\gamma_m(T)$ to measurements on quasi-one-dimensional, disordered wires, for two AgFe samples [11], (AgFe 2 and AgFe 3 having 27 ± 3 and 67.5 ± 7 ppm Fe impurities in Ag, respectively), with a Kondo scale $T_K \approx 5$ K (for $S = 3/2$, see below). These measurements extend up to $T \lesssim T_K$ allowing the region $T/T_K \lesssim 1$ of the scaling curves in Fig. 2 to be compared to experiment. At $T \gtrsim T_K \approx 5$ K (i.e., $T/T_K \geq 1$) the large phonon contribution to the decoherence rate prohibits reliable extraction of $\gamma_m(T)$ for our Ag samples (see below). In order to compare theory and experiment for temperatures $T/T_K \geq 1$, above the maximum in the decoherence

rate, we therefore carried out new measurements on a sample (AuFe 3) with 7 ± 0.7 ppm Fe impurities in Au with a lower Kondo scale $T_K \approx 1.3$ K but, as discussed above, described by the same Kondo model. Combining both sets of measurement thereby allows a large part of the scaling curves in Fig. 2 to be compared with experiment.

Following [11], we subtract the electron-electron contribution [23] from the total resistivity ρ, yielding $\Delta\rho$ due to magnetic impurities (m) and phonons (ph):

$$\Delta\rho(T) = \rho_m(T) + \rho_{ph}(T) + \delta. \tag{2}$$

Here δ is an (unknown) offset [24] and $\Delta\rho(T)$ is expressed per magnetic impurity. For temperatures low enough that $\rho_{ph}(T)$ can be neglected, $\Delta\rho(T) - \delta$ corresponds to the theoretical curve $\rho_m(T) = \rho_m(0)f_S(T/T_K^S)$ (cf. caption of Fig. 2), where $\rho_m(0) = \Delta\rho(0) - \delta$ is the unitary Kondo resistivity. Figure 3 illustrates how we extract the Kondo scale T_K^S and $\rho_m(0)$ from the experimental data, by fitting the Kondo-dominated part of $\Delta\rho(T)$ in a fixed temperature range (specified in the caption of Fig. 3) to the NRG results of Fig. 2(a), using the ansatz

$$\Delta\rho(T) \approx \delta + [\Delta\rho(0) - \delta]f_S(T/T_K^S). \tag{3}$$

Such fits are made for each of the fully screened Kondo models, using T_K^S and δ as fit parameters. Importantly, the values for T_K^S and $\rho_m(0)$ obtained from the fits, given in the inset and caption of Fig. 3, respectively, show a significant S dependence: both T_K^S and $\rho_m(0)$ increase with S, since the slope of the logarithmic Kondo increase of the theory curves for ρ_m (cf. Fig. 2) decreases significantly in magni-

FIG. 2 (color online). (a) Resistivity $\rho_m(T)$ (solid lines) and (b) decoherence rate $\gamma_m(T)$ for $2S = n = 1, 2, 3$; $\rho_m(0) = 2\tau\bar{\rho}/\pi\hbar\nu_0$, $\gamma_m^0 = 2/\pi\hbar\nu_0$, where $\bar{\rho}$ is the residual resistivity, ν_0 the density of states per spin and channel, τ the elastic scattering time, and γ_m^{max} is the maximum value of $\gamma_m(T)$. We defined the Kondo scale T_K^S for each S via $\rho_m(T_K^S) = \rho_m(0)/2$. Dashed lines in (a) show that the empirical form $\rho_m(T)/\rho_m(0) \approx f_S(T/T_K^S)$ with $f_S(x) = [1 + (2^{1/\alpha_S} - 1)x^2]^{-\alpha_S}$, used to fit experimental to NRG results for $S = 1/2$ [25], also adequately fits the NRG results for $S = 1$ and $S = 3/2$.
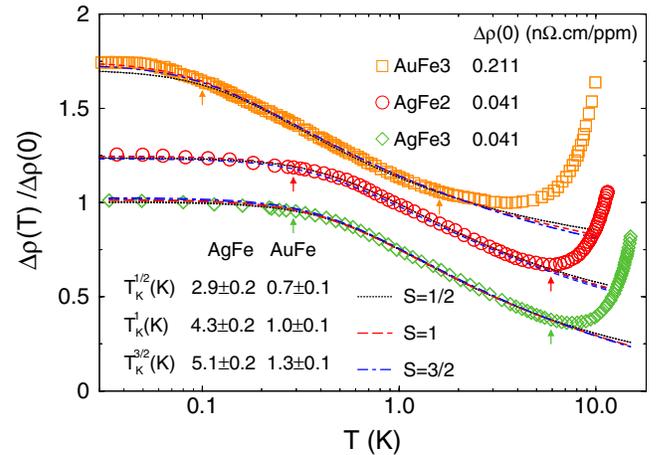
FIG. 3 (color online). Measured resistivities $\Delta\rho(T)$ (symbols) fitted to Eq. (3) (lines), for $n = 2S = 1, 2,$ and 3, in the range below the onset of the phonon contribution, but above 100–200 mK [26]. Specifically, we used 0.1–1.6 K for AuFe and 0.29–5.9 K for AgFe (arrows). The curves for AgFe 2 and AuFe 3 have been offset vertically by 0.25 and 0.75, respectively. The inset gives the Kondo scales T_K^S for AgFe and AuFe extracted from the fits. Estimates of the unitary Kondo resistivities for $n = 1, 2,$ and 3 (in units of $n\Omega \cdot cm/ppm$) yield $\rho_m(0) = 0.041, 0.047,$ and 0.049 for AgFe (averaged over the two samples) and 0.23, 0.26, and 0.27 for AuFe, respectively.

tude with S. Nevertheless, all three models fit the Kondo contribution very well, as shown in Fig. 3, so a determination of the appropriate model from resistivity data alone is not possible.

To break this impasse, we exploit the remarkably sensitive S dependence of the spin-flip-induced decoherence rate $\gamma_m(T)$. Figure 4 shows the measured dimensionless decoherence rate $\gamma_m(T)/\gamma_m^{\max}$ for Ag and Au samples (symbols) as function of T/T_K^S for $S = 1/2$, 1, and 3/2, using the T_K^S values extracted from the resistivities, together with the corresponding parameter-free theoretical predictions (lines), taken from Fig. 2(b). The agreement between theory and experiment is poor for $S = 1/2$, better for $S = 1$, but excellent for $S = 3/2$, confirming the conclusion drawn above from *ab initio* calculations. The dependence on S is most strikingly revealed through the width of the plateau region (in units of T/T_K^S), which grows with S for the theory curves but shrinks with S for the experimental data (for which T_K^S grows with S), with $S = 3/2$ giving the best agreement.

Conclusions.—In this Letter we addressed one of the fundamental unresolved questions of Kondo physics: that of deriving and solving the effective low-energy Kondo model appropriate for a realistic description of Fe impurities in Au and Ag. Remarkably, for both Ag and Au samples, the use of a fully screened $S = 3/2$ three-channel Kondo model allows a quantitatively consistent description of both the resistivity and decoherence rate *with a single* T_K (for each material). Our results set a benchmark for the level of quantitative understanding attainable for the Kondo effect in real materials.
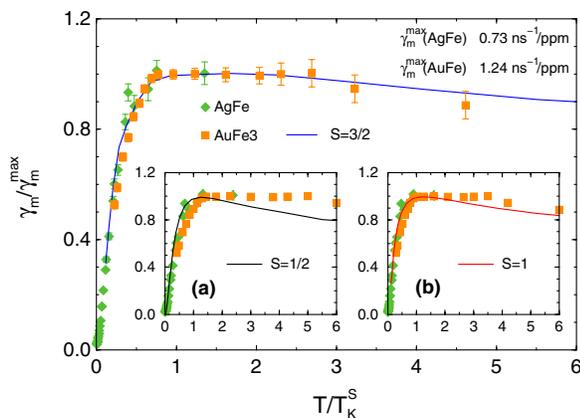
FIG. 4 (color online). Comparison of the measured (symbols) and theoretical (lines) results for the dimensionless decoherence rate $\gamma_m(T)/\gamma_m^{\max}$ as function of T/T_K^S, using $S = 3/2$. Insets show comparisons to $S = 1/2$ (a) and $S = 1$ (b). T_K^S for AgFe and AuFe was extracted from the resistivities (inset of Fig. 3), while $\gamma_m^{\max}$ was determined as the average plateau height in the region $T/T_K^{3/2} \in [0.7, 1.35]$. Typical error bars are shown for $S = 3/2$. They grow with increasing temperatures due to the increasing difficulty of subtracting the growing phonon contribution to the decoherence rate.

[1]   W. J. de Haas, J. de Boer, and G. J. van den Berg, Physica (Amsterdam) **1**, 1115 (1934).
[2]   J. Kondo, Prog. Theor. Phys. **32**, 37 (1964).
[3]   P. Nozières and A. Blandin, J. Phys. (Paris) **41**, 193 (1980).
[4]   A. C. Hewson, *The Kondo Problem to Heavy Fermions*, Cambridge Studies in Magnetism (Cambridge University Press, Cambridge, England, 1997).
[5]   H. U. Desgranges, J. Phys. C **18**, 5481 (1985).
[6]   P. Steiner and S. Hüfner, Phys. Rev. B **12**, 842 (1975).
[7]   P. Schlottmann and P. D. Sacramento, Adv. Phys. **42**, 641 (1993).
[8]   F. Pierre *et al.*, Phys. Rev. B **68**, 085413 (2003).
[9]   F. Schopfer *et al.*, Phys. Rev. Lett. **90**, 056801 (2003).
[10] C. Bäuerle *et al.*, Phys. Rev. Lett. **95**, 266805 (2005).
[11] F. Mallet *et al.*, Phys. Rev. Lett. **97**, 226804 (2006).
[12] G. M. Alzoubi and N. O. Birge, Phys. Rev. Lett. **97**, 226803 (2006).
[13] T. Capron *et al.*, Phys. Rev. B **77**, 033102 (2008).
[14] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11 169 (1996).
[15] G. Vogl, W. Mansel, and P. H. Dederichs, Phys. Rev. Lett. **36**, 1497 (1976).
[16] R. Kirsch *et al.*, Europhys. Lett. **59**, 430 (2002).
[17] W. D. Brewer *et al.*, Phys. Rev. Lett. **93**, 077205 (2004).
[18] T. Micklitz *et al.*, Phys. Rev. Lett. **96**, 226601 (2006).
[19] G. Zaránd *et al.*, Phys. Rev. Lett. **93**, 107204 (2004).
[20] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).
[21] R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008); T. A. Costi, A. C. Hewson, and V. Zlatić, J. Phys. Condens. Matter **6**, 2519 (1994);O. Sakai, Y. Shimizu, and N. Kaneko, Physica (Amsterdam) **186–188B**, 323 (1993).
[22] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006); F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005); W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000); R. Bulla, A. C. Hewson, and Th. Pruschke, J. Phys. Condens. Matter **10**, 8365 (1998).
[23] B. L. Altshuler, A. G. Aronov, and D. E. Khmelnitzky, J. Phys. C **15**, 7367 (1982).
[24] Experiments measure resistance changes on ramping up the temperature, hence the unknown offset δ.
[25] D. Goldhaber-Gordon *et al.*, Phys. Rev. Lett. **81**, 5225 (1998).
[26] Below this temperature, the smaller signal to noise ratio makes the measurements less accurate.

102

P5. Quantum quench of Kondo correlations in optical absorption

▶ Christian Latta, Florian Haupt, Markus Hanl, Andreas Weichselbaum, Martin Claassen, Wolf Wuester, Parisa Fallahi, Stefan Faelt, Leonid Glazman, Jan von Delft, Hakan E. Türeci, and Atac Imamoglu; Nature **474**, 627 (2011).

While I haven been much engaged with methods and method development myself, there have also been many applications on interesting physical systems. This was usually done through projects with diploma, masters, as well as Ph.D. students. A very successful project that I would like to point out in that respect, is work done with the Ph.D. student Markus Hanl in a very fruitful collaboration with Hakan E. Türeci and the experimental group of Atac Imamoglu in the context of absorption spectra of Kondo-excitons in quantum dots.

In this project, the clear challenge from the experimental point of view was to have a large enough Kondo-temperature, such that for reasonably small temperatures within the lifetime of an exciton, that is generated through the absorption of a photon, Kondo correlations can build up. What from a theorists point of view corresponds to a few tunable parameters (here in particular the coupling of a quantum dot through the hybridization Γ to a close-by fermionic reservoir), requires cutting-edge experimental expertise. Amazingly, most of the experimental setting eventually can be understood in terms of a plain single impurity Anderson model together with Anderson orthogonality physics. After significant fine-tuning of the corresponding Fermi-Golden rule calculation, we achieved a good theoretical understanding of the Kondo-exciton dynamics in the experimental setting, fully fitted and understood through FDM-NRG simulations in combination with analytical calculations by Türeci et al. (2011).

LETTER

Quantum quench of Kondo correlations in optical absorption

C. Latta[1], F. Haupt[1], M. Hanl[2], A. Weichselbaum[2], M. Claassen[1], W. Wuester[1], P. Fallahi[1], S. Faelt[1], L. Glazman[3], J. von Delft[2], H. E. Türeci[1,4] & A. Imamoglu[1]

The interaction between a single confined spin and the spins of an electron reservoir leads to one of the most remarkable phenomena of many-body physics—the Kondo effect[1,2]. Electronic transport measurements on single artificial atoms, or quantum dots, have made it possible to study the effect in great detail[3–5]. Here we report optical measurements on a single semiconductor quantum dot tunnel-coupled to a degenerate electron gas which show that absorption of a single photon leads to an abrupt change in the system Hamiltonian and a quantum quench of Kondo correlations. By inferring the characteristic power-law exponents from the experimental absorption line shapes, we find a unique signature of the quench in the form of an Anderson orthogonality catastrophe[6,7], induced by a vanishing overlap between the initial and final many-body wavefunctions. We show that the power-law exponent that determines the degree of orthogonality can be tuned using an external magnetic field[8], which unequivocally demonstrates that the observed absorption line shape originates from Kondo correlations. Our experiments demonstrate that optical measurements on single artificial atoms offer new perspectives on many-body phenomena previously studied using transport spectroscopy only.

Optical spectroscopy of single quantum dots has demonstrated its potential for applications in quantum information processing, particularly in the realization of single- and entangled-photon sources[9,10], coherent spin qubits[11,12] and a spin–photon interface[13,14]. Although recent experiments have established this system as a new model for solid-state quantum optics, all of the striking experimental observations made so far can be understood within the framework of single- or few-particle physics enriched by perturbative coupling to reservoirs involving phonons, a degenerate electron gas[15–17] or nuclear spins[18,19].

We present differential transmission experiments[20] on single, charge-tunable quantum dots that reveal optical signatures of the Kondo effect. By contrast with prior experiments[17,21], the tunnel coupling between the quantum dot and a nearby degenerate electron gas, which we refer to as the fermionic reservoir, is engineered to be so strong that the resulting exchange interactions cannot be treated using a perturbative system–reservoir theory: in the initial state, the 'system'—quantum dot spin—is maximally entangled with the fermionic reservoir, forming a singlet. Various settings have been proposed for finding optical signatures of Kondo physics[8,22–25]; our work is most closely related to the theoretical investigation of refs 8, 25.

The feature that differentiates our results from all prior transport-based investigations of the Kondo effect[3–5] is the realization of a quantum quench of the local Hamiltonian; in our experiments, photon absorption abruptly turns off the exchange interaction between the quantum dot electron and the fermionic reservoir, leading to the destruction of the correlated dot–reservoir singlet that otherwise acts as a local scattering potential for all reservoir electrons. The overlap between N-electron fermionic reservoir states with and without a local scattering potential scales as $N^{-\alpha}$, with $\alpha > 0$ (refs 6, 7). This reduced overlap, called an Anderson orthogonality catastrophe (AOC), leads to

a power-law tail in absorption if the scattering potential is turned on or off by photon absorption. Here we determine the AOC-induced power-law exponents in the absorption line shape that uniquely characterize the quench of Kondo correlations. Moreover, by tuning the applied laser frequency, we observe both the perturbative and the non-perturbative regimes of the Kondo effect in one absorption line shape, without having to change the fermionic reservoir (electron) temperature, T_{FR}. The AOC after a Kondo quench can, in principle, also be probed by core-level X-ray absorption spectroscopy of suitable bulk materials[26], but optical studies of quantum dots offer higher resolution and a tunable local Hamiltonian.

The quantum dot sample we study is shown schematically in Fig. 1a: a gate voltage, V_g, applied between a top Schottky gate and the degenerate electron gas, allows us to tune the charging state of the quantum dot[27]. Figure 1b shows the photoluminescence spectrum of a particular quantum dot (dot 1), as a function of V_g, where different discrete 'charging plateaux' are clearly observable. The dependence of the photoluminescence energy on the quantum dot charging state originates from a Coulomb renormalization of the optical transition energy. In addition to photoluminescence lines (for example X^0) associated with a fixed charging state (for example neutral) of the quantum dot, we also observe spatially indirect transitions with a strong dependence on V_g (refs 8, 17; see Fig. 1b, red arrow).
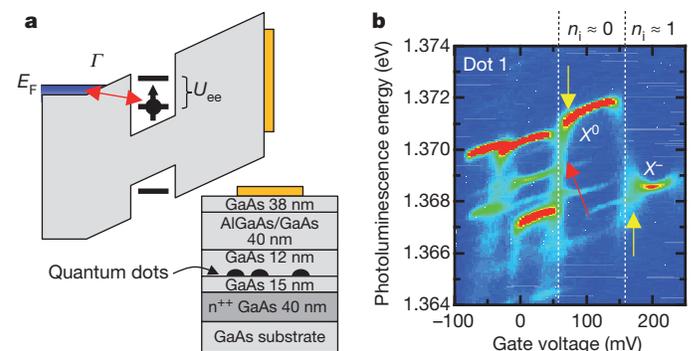

Figure 1 | Single quantum dot strongly coupled to a fermionic reservoir. a, Band structure of the device. The quantum dots are separated by a 15-nm tunnel barrier from an n^{++}-doped GaAs layer (fermionic reservoir). A voltage, V_g, applied between the electron gas and a semi-transparent NiCr gate on the sample surface controls the relative value of the quantum dot single-particle energy levels with respect to the Fermi energy, E_F. **b**, Low-temperature (4 K) photoluminescence spectrum of a single quantum dot (dot 1) as a function of V_g; n_i denotes the initial state electron occupancy of the quantum dot. The interaction of the quantum dot electron with the Fermi sea leads to a broadening of the photoluminescence lines at the plateau edges (yellow arrows) and indirect recombinations of a quantum dot hole and a Fermi sea electron (red arrow). Indirect transitions are identified by the stronger V_g dependence of the transition energy, compared with that for direct transitions. A detailed discussion of the origin of various photoluminescence lines can be found in ref. 17.

[1]Institute of Quantum Electronics, ETH-Zürich, CH-8093 Zürich, Switzerland. [2]Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München, D-80333 München, Germany. [3]Sloane Physics Laboratory, Yale University, New Haven, Connecticut 06520, USA. [4]Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA.

In this Letter, we focus on the X^- plateau, for which the quantum dot carries the charge of a single electron and the influence of the fermionic reservoir on the quantum dot photoluminescence dispersion and linewidth is strongest. The X^- optical transition couples the initial configuration, containing on average one electron in the quantum dot, to a final configuration, containing on average two electrons and a valence-band hole (a negatively charged trion). This transition can be described within the framework of an excitonic Anderson model[8,25] (EAM), depicted schematically in Fig. 2c (and described explicitly in Supplementary Information). It is parameterized by the energy, ε, of the quantum dot electron level with respect to the Fermi level; the on-site Coulomb repulsion, U_{ee}; the tunnelling rate, Γ, between quantum dot and fermionic reservoir; the half-bandwidth, D, of the fermionic reservoir; and the electron–hole Coulomb attraction, U_{eh}. The last is relevant only in the final configuration, where it effectively lowers the electron level energy to $\varepsilon - U_{eh}$, thus ensuring the double occupancy of the electron level. An estimate from the photoluminescence data in Fig. 1b yields $U_{eh} \approx U_{ee} + 4$ meV.

The inset of Fig. 2a shows high-resolution laser absorption spectroscopy on dot 1 across the X^- single-electron charging plateau
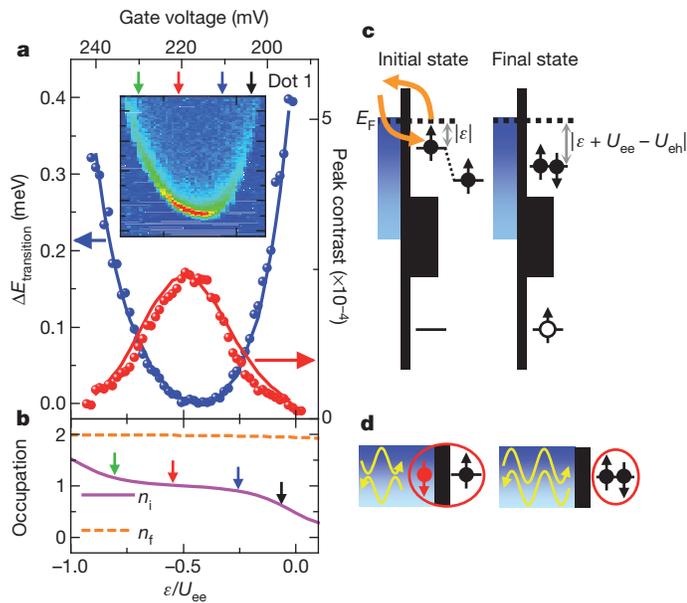
Figure 2 | Gate voltage dependence of the peak absorption strength.
a, Experimental data (symbols) for the ε dependence of the shift in the resonance energy, $\Delta E_{transition}$ (blue, left axis), and the absorption contrast (red, right axis) are well reproduced by NRG calculations (solid lines) for the following parameters: $U_{ee} = 7.5$ meV, $\Gamma = 0.7$ meV, $D = 3.5$ meV, $U_{eh} = 11$ meV, $T_{FR} = 180$ mK. Inset, absorption on the negatively charged exciton X^- transition of dot 1 as a function of the gate voltage, measured at $T_{FR} = 180$ mK. **b**, NRG results for the respective occupancies, n_i and n_f, of the quantum dot electron level in the initial and final ground states. **c**, Energy renormalization process: the initial configuration (left) features a single electron in the quantum dot, whose energy is lowered by virtual tunnelling between the dot and the fermionic reservoir. Because virtual excitations with energy ΔE contribute a shift proportional to $-\Gamma/\Delta E$, the total shift (involving a sum over all possible values of ΔE), is strongest near the edges of the X^- plateau. Towards the right-hand edge (ε near zero), the dominant contribution comes from virtual tunnelling of the quantum dot electron into the fermionic reservoir (as depicted); towards the left-hand edge (ε near $-U_{ee}$), it comes from virtual tunnelling of a fermionic reservoir electron into the quantum dot (not depicted). In the final configuration (right), the quantum dot contains two electrons and a hole. The electron–hole Coulomb attraction, U_{eh}, effectively lowers the quantum dot electron level energy to $\varepsilon - U_{eh}$. This raises the energy cost, ΔE, for virtual excitations by $U_{eh} - U_{ee}$ (which is $\gg \Gamma$), such that the final-state energy renormalization is negligible. The renormalization of the transition energy, probed by a weak laser, is thus mainly due to initial-state energy renormalization. **d**, Anderson orthogonality: the Kondo cloud (left-hand diagram) and local singlet (right-hand diagram) of the initial and final configurations produce strong or weak scattering phase shifts, respectively.

(Supplementary Information). Here we parameterize V_g in terms of ε, normalized and shifted such that $\varepsilon = -U_{ee}/2$ for the gate voltage at which the absorption contrast is maximal. Instead of the usual linear d.c. Stark shift of the absorption peak that is characteristic of charge-tunable quantum dots, we find a strongly nonlinear, ε-dependent shift of the X^- transition energy[15,17], which measures the energy difference between the final and initial ground states. As shown in Fig. 2c, this energy shift arises from a renormalization of the initial state energy[28] due to virtual tunnelling between the singly occupied quantum dot and the fermionic reservoir (analogous to the Lamb shift of atomic ground states). The final trion state energy, on the other hand, is hardly affected by virtual tunnelling processes, owing to $U_{eh} - U_{ee}$ being large. This renormalization-induced redshift of the initial state is strongest at the plateau edges and leads to an ε-dependent blueshift of the optical resonance frequency. The latter can be used to determine the EAM parameters for dot 1: $U_{ee} = 7.5$ meV, $\Gamma = 0.7$ meV and $D = 3.5$ meV. Numerical renormalization group (NRG) calculations for the transition energy (Fig. 2a, blue line) give excellent agreement with the experimental data (blue symbols).

We now consider the detailed form of the absorption line shape, $A(v)$, as function of the detuning, v, between the applied laser frequency and the transition threshold. Figure 3a shows, on a log–log scale, the blue ($v > 0$) tail of $A(v)$ for dot 1, for the four values of gate voltage indicated by arrows in the inset of Fig. 2a. The inset of Fig. 3a compares the full, un-normalized absorption line shapes for the same gate voltages on a linear scale; the red ($v < 0$) absorption tail allows us to determine the temperature of the fermionic reservoir to be
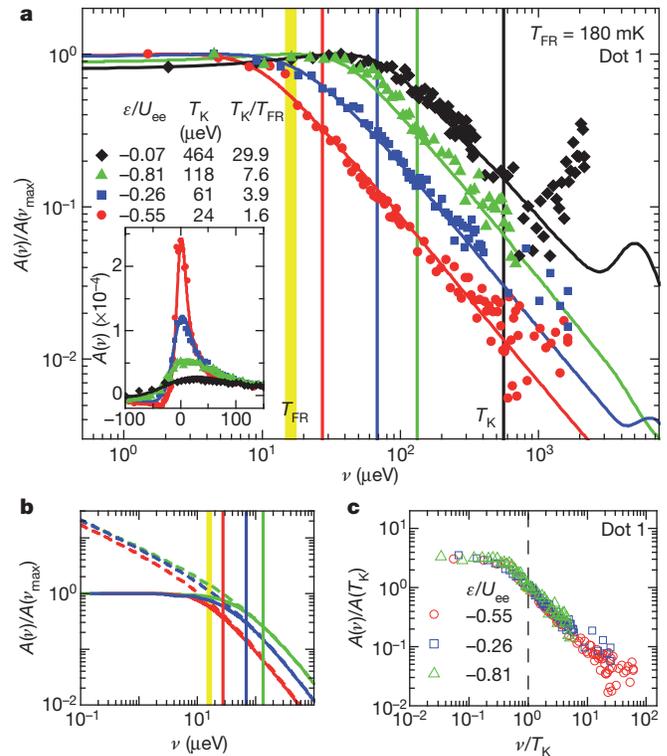
Figure 3 | The absorption line shape $A(v)$. **a**, Blue tail of $A(v)/A(v_{max})$ for dot 1, plotted versus the laser detuning, v, on a log–log scale. Here v_{max} is the threshold frequency for which the absorption strength is maximal. The experimental data were measured at an electron temperature of $T_{FR} = 180$ mK for the four values of gate voltage, ε, indicated by arrows in Fig. 2a; the corresponding Kondo temperatures, $T_K(\varepsilon)$, are indicated by vertical lines in matching colours. The yellow line indicates T_{FR}. NRG results (solid lines), obtained using the parameters from the fit in Fig. 2a, are in remarkable agreement with experiment. Inset, the measured full (un-normalized) absorption line shape for the same ε values, plotted on a linear scale. **b**, NRG results for $T = T_{FR}$ (solid lines) and $T_{FR} = 0$ (dashed lines); the latter show the $v^{-0.5}$ behaviour expected in the strong-coupling regime, $T_{FR} \ll v \ll T_K$. **c**, The rescaled line shape, $A(v)/A(T_K)$, versus v/T_K shows a universal scaling collapse characteristic of Kondo physics.

P5.

106

$T_{FR} = 180$ mK, equivalent to 15.6 μeV (Supplementary Information). The strong variation of the peak absorption strength and width shown in the inset of Fig. 3a is a consequence of the exponential dependence of the Kondo temperature on the gate voltage ε:

$$T_K(\varepsilon) = \sqrt{\Gamma D} \exp\left[-\left(1 - \left(\frac{2\varepsilon}{U_{ee}} + 1\right)^2\right)\frac{\pi U_{ee}}{8\Gamma}\right] \quad (1)$$

For dot 1, T_K varies between 24 and 118 μeV; we emphasize that even though $T_K = 464$ μeV for the black curve (Fig. 3a, inset), the dot–reservoir system is no longer in the local moment regime for this gate voltage. All line shapes carry the signatures of an optical interference effect induced by the sample structure (causing some line shapes to become negative for small red detunings), and of independently measured fluctuations in gate voltage; both effects have been taken into account in the calculated line shapes (Supplementary Information). Calculating the line shapes using NRG (solid lines) without any further fit parameters, we find remarkable agreement with experiment for all four line shapes shown in Fig. 3a, demonstrating the validity of the EAM[8] for the coupled dot–reservoir system.

For blue detunings satisfying $v > \max(T_{FR}, T_K)$, a perturbative description for $A(v)$ is possible. The frequency scale for which the perturbative $\sim v^{-1}$ dependence in Fig. 3a sets in and the peak absorption contrast itself both strongly depend on gate voltage. Remarkably, for gate voltages such that the initial ground state is a Kondo singlet, this dependence is such that it permits a scaling collapse: Fig. 3c shows the normalized absorption line shape, $A(v)/A(T_K(\varepsilon))$, as a function of v/T_K for the red, green and blue curves of Fig. 3a (but omitting the black curve, which is in the mixed valence regime). We find that all three curves collapse to a universal scaling function of v/T_K, as expected[8] for the regime $T_{FR} \ll v \ll U_{ee}$. Thus, the ε dependence of the crossover scale is captured by equation (1) for T_K; this observation is unequivocal proof that the Kondo effect is indeed present in our system.

In the limit $T_{FR} < v < T_K$, a perturbative description of the line shape is no longer valid. In the initial configuration, the exchange interaction between the quantum dot and the fermionic reservoir induces a 'Kondo screening cloud' that forms a singlet with the quantum dot spin. This acts as a scattering potential that induces strong phase shifts for those low-energy fermionic excitations whose energies differ from the Fermi level by T_K or less. In the final configuration after photon absorption, the quantum dot has two electrons in a local singlet state. Therefore, the Kondo screening cloud, and the scattering potential that it constitutes for reservoir electrons, disappears in the long-time limit: the corresponding ground-state wavefunction is a tensor product of the local singlet and free electronic states, with only weak phase shifts. Because the initial and final fermionic reservoir phase shifts differ (as depicted schematically in Fig. 2d), the fermionic reservoir does not remain a spectator during the X^- transition; instead, the transition matrix element between the ground states of the initial and final configurations is vanishingly small. This leads to an AOC that manifests itself by transforming a delta-function resonance (of an uncoupled quantum dot) into a power-law singularity[6] of the form $v^{-\eta}$, where the exponent η characterizes the extent of the AOC. For $T_{FR} \ll v \ll T_K$, the absorption line shape of the X^- transition is expected to show an analogous power-law singularity. The exponent η is predicted[8,25] to range between 0 and 0.5 (assuming no magnetic field), with $\eta \approx 0.5$ being characteristic for a Kondo-correlated initial state and an uncorrelated final state. This line shape modification is a consequence of a redistribution of the optical oscillator strength, associated with the fact that the fermionic reservoir wavefunction in the Kondo-correlated initial state has finite overlap with a range of final states consisting of electron–hole pair excitations out of a non-interacting fermionic reservoir.

If $T_{FR} \ll T_K$ and the optical detuning is reduced below T_K, the line shape is predicted to cross over smoothly from the perturbative $1/v$ tail to the strong-coupling $1/v^{0.5}$ power law just discussed. This crossover is illustrated in Fig. 3b (dashed lines) by NRG calculations, performed at $T_{FR} = 0$ for the three ε values of Fig. 3c: Remarkably, despite drastic

differences in the $v > T_K$ tails due to different values of $T_K(\varepsilon)$, all three line shapes show similar power-law exponents, of around $\eta \approx 0.5$, for $v \ll T_K$. For non-zero temperature, however, the $1/v^{0.5}$ power law is cut off and saturates once v decreases past T_{FR} (Fig. 3b, solid lines), because of thermal averaging over initial states with excitation energies $\lesssim T_{FR}$.

A direct extraction of the $1/v^{0.5}$ power law from the measured data is difficult owing to the small accessible experimental window, $T_{FR} < v < T_K$. Nevertheless, we are able to determine the power-law exponent accurately for a more strongly coupled quantum dot (dot 2) by using the fact that the detailed form of the line shape sensitively depends on the exponent η, which can be tuned using an external magnetic field[8]. This tunability arises because the magnetic field, B_{ext}, changes the initial dot occupancies, favouring spin up over spin down, and hence affects the overlap between the initial and final states of the transition (Supplementary Information). Figure 4a shows the $B_{ext} = 0$ absorption line shape for dot 2 with parameters $U_{ee} = 7.5$ meV, $\Gamma = 1$ meV, $D = 6.5$ meV and $U_{eh} = (3/2)U_{ee}$, measured at $\varepsilon/U_{ee} = -0.43$ (where $T_K = 140$ μeV) and $T_{FR} = 15.6$ μeV. An attempt to obtain a fit to the experimental absorption line shape using a perturbative formula[8]

$$A(v) \propto \frac{v/T_{FR}}{1 - e^{-v/T_{FR}}}\frac{\gamma}{v^2 + \gamma^2/4}$$

where $\gamma \lesssim T_{FR}$ denotes a phenomenological relaxation rate, fails markedly for dot 2 (Fig. 4a, red curve). By striking contrast, Fig. 4b shows that an excellent fit is obtained for a weakly coupled dot (dot 3; Supplementary Information).
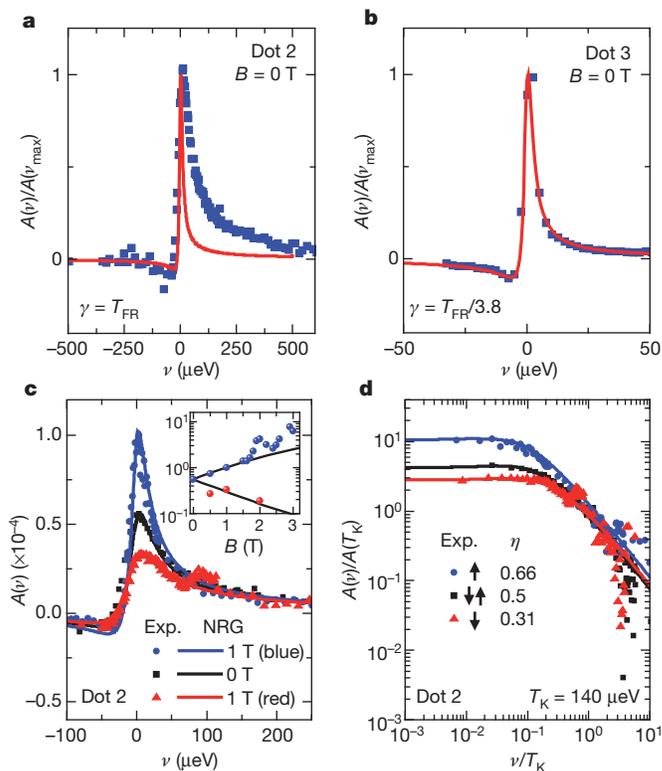
P5.

Figure 4 | Magnetic field dependence of the absorption. a, The absorption line shape of dot 2 for $B = 0$ (symbols) cannot be fit by the perturbative formula (red line) given in the text. **b**, By contrast, for dot 3 such a fit works well. **c**, Absorption line shapes for dot 2, at $B_{ext} = 0$ and 1 T, for the blue–red trion transition. The magnetic field changes the strength of the AOC and the line shape. The small peak that appears at $v \approx 80$ μeV in the red trion absorption is due to incomplete suppression of the laser polarization that couples to the blue trion transition. Inset, the peak absorption contrast shows good agreement with the NRG calculations for $B_{ext} \lesssim 1.5$ T. **d**, Normalized absorption line shape for dot 2 in a log–log plot. These measurements pin the value of $\eta(B_{ext} = 0)$ to ~ 0.5, which is a direct signature of a Kondo singlet in the absorption line shape. In addition, they demonstrate the tunability of an orthogonality exponent.

107

Figure 4c shows the magnetic field dependence of the line shape of dot 2, measured in Faraday geometry, where quantum dot optical selection rules[13] ensure that by choosing right- or left-handed circular polarization of the laser field, it is possible to probe selectively the blue or, respectively, red trion transition that couples exclusively to the spin-up or, respectively, spin-down initial state. In comparison with the $B_{ext} = 0$ (Fig. 4c, black squares) results, the absorption line shapes for the blue (Fig. 4c; blue dots) and red (Fig. 4c; red triangles) trion transitions at $B_{ext} = 1$ T exhibit two striking features: the peak contrast increases (blue) or decreases (red) by a factor of ~2, and the area under the absorption curve increases (blue) or decreases (red) by less than 20%. These observations indicate that the change in the $B_{ext} \leq 1.5$ T line shapes is predominantly due to a line narrowing associated with an increase in the AOC power-law exponent, η, of the blue trion transition and a line broadening associated with a decrease in η for the red trion transition. To quantify the field-induced change in η, we plot in Fig. 4d the corresponding normalized line shapes, $A(\nu)/A(T_K)$, as functions of ν/T_K in a log–log plot, together with the corresponding NRG results (solid lines): the latter yield $\eta = 0.5$ at $B_{ext} = 0$ and $\eta = 0.31$ (red trion) and $\eta = 0.66$ (blue trion) at $B_{ext} = 1$ T, proving the remarkable sensitivity of the measured line shapes to the AOC-determined power-law exponents. By contrast with Fig. 3c, the line shapes in Fig. 4d do not show a scaling collapse. We emphasize that qualitatively similar features are observed for all field values $B_{ext} \leq 1.5$ T; for $B_{ext} > 1.5$ T, the blue trion absorption contrast has oscillations (Fig. 4c, inset), most probably stemming from the modification of the fermionic reservoir density of states at high fields in Faraday geometry.

The area under the (un-normalized) absorption line shape is proportional to the initial occupancy, $n_\uparrow$ or $n_\downarrow$, of the spin-up or, respectively, spin-down state. The small ($\leq 20\%$) field-induced change in the measured areas in Fig. 4c implies a small magnetization, $m = (n_\uparrow - n_\downarrow)/2 \approx 0.16$ (Supplementary Information). By contrast, the corresponding magnetization for a free spin would have been $m = 0.40$. This measurement confirms that the static spin susceptibility of the initial configuration is substantially reduced relative to that of a free spin, providing yet another optical signature of the Kondo screening.

The remarkable agreement between our experimental data depicted in Figs 2–4 and the NRG calculations demonstrates Kondo correlations between a quantum dot electron and the electrons in a fermionic reservoir. The optical probe of these correlations unequivocally shows the signatures of Anderson orthogonality physics associated with the quantum quench of Kondo correlations, with field-tunable power-law exponents. Our experiments establish the potential of single, optically active quantum dots in investigating many-body physics. In addition, they pave the way for a new class of quantum optics experiments in which the influence of the simultaneous presence of non-perturbative cavity or laser coupling and Kondo correlations on electric field and photon correlations could be investigated.

METHODS SUMMARY

The InGaAs quantum dots studied in this work were grown by molecular beam epitaxy; the quantum dot layer was separated by a nominally 15-nm-thick GaAs tunnel barrier from a back gate consisting of a 40-nm-thick n^{++}-doped GaAs layer. This back gate serves as an electron reservoir. The distance from the quantum dot layer to the sample surface was 90 nm. A voltage applied between a 5-nm-thick NiCr top gate and the n^{++} GaAs back gate allows for discrete charging of the quantum dots. The sample was placed inside a fibre-based confocal microscope embedded in a dilution refrigerator with a base temperature of 20 mK in the mixing chamber. The objective was mounted on a stack of low-temperature x–y–z positioners. The cryostat was equipped with a 7-T magnet. The absorption experiments were performed by focusing on a single quantum dot a power- and frequency-stabilized, single-mode tunable laser with an intensity of 15 nW. The objective had a numerical aperture of 0.6, yielding a diffraction-limited spot size. The change in transmission through the sample was recorded using a silicon photodiode. To increase the signal-to-noise ratio, a lock-in technique was used whereby the gate voltage was modulated at 187.195 Hz with a modulation amplitude of 50 mV.

The calculations were carried out using the NRG. The continuous energy spectrum of the Fermi reservoir was logarithmically discretized and mapped onto a semi-infinite chain with exponentially decaying hopping amplitudes. In each

iteration, a new site was added to the chain, which corresponds to including ever lower energy scales of the system. By combining NRG data from all iterations, it was possible to construct a complete set of approximate many-body eigenstates of the full Hamiltonian, which could be used to calculate the physical quantities using the full-density-matrix NRG (Supplementary Information).

1. Kondo, J. Resistance minimum in dilute magnetic alloys. *Prog. Theor. Phys.* **32**, 37–49 (1964).
2. Kouwenhoven, L. P. & Glazman, L. Revival of the Kondo effect. *Phys. World* **14**, 33–38 (January 2001).
3. Goldhaber-Gordon, D. *et al.* Kondo effect in a single-electron transistor. *Nature* **391**, 156–159 (1998).
4. Cronenwett, S. M., Oosterkamp, T. H. & Kouwenhoven, L. P. A tunable Kondo effect in quantum dots. *Science* **281**, 540–544 (1998).
5. van der Wiel, W. G. *et al.* The Kondo effect in the unitary limit. *Science* **289**, 2105–2108 (2000).
6. Mahan, G. *Many-Particle Physics* 612–621 (Kluwer, 2000).
7. Anderson, P. W. Infrared catastrophe in Fermi gases with local scattering potentials. *Phys. Rev. Lett.* **18**, 1049–1051 (1967).
8. Türeci, H. E. *et al.* Many-body dynamics of exciton creation in a quantum dot by optical absorption: a quantum quench towards Kondo correlations. *Phys. Rev. Lett.* **106**, 107402 (2011).
9. Michler, P. *et al.* A quantum dot single-photon turnstile device. *Science* **290**, 2282–2285 (2000).
10. Dousse, A. *et al.* Ultrabright source of entangled photon pairs. *Nature* **466**, 217–220 (2010).
11. Press, D., Ladd, T. D., Zhang, B. & Yamamoto, Y. Complete quantum control of a single quantum dot spin using ultrafast optical pulses. *Nature* **456**, 218–221 (2008).
12. Kim, D. *et al.* Ultrafast optical control of entanglement between two quantum-dot spins. *Nature Phys.* **7**, 223–229 (2011).
13. Yilmaz, S. T., Fallahi, P. & Imamoglu, A. Quantum-dot-spin single-photon interface. *Phys. Rev. Lett.* **105**, 033601 (2010).
14. Claassen, M., Türeci, H. & Imamoglu, A. Solid-state spin-photon quantum interface without spin-orbit coupling. *Phys. Rev. Lett.* **104**, 177403 (2010).
15. Dalgarno, P. A. *et al.* Optically induced hybridization of a quantum dot state with a filled continuum. *Phys. Rev. Lett.* **100**, 176801 (2008).
16. Hilario, L. M. L. & Aligia, A. A. Photoluminescence of a quantum dot hybridized with a continuum of extended states. *Phys. Rev. Lett.* **103**, 156802 (2009).
17. Kleemans, N. A. J. M. *et al.* Many-body exciton states in self-assembled quantum dots coupled to a Fermi sea. *Nature Phys.* **6**, 534–538 (2010).
18. Latta, C. *et al.* Confluence of resonant laser excitation and bidirectional quantum-dot nuclear-spin polarization. *Nature Phys.* **5**, 758–763 (2009).
19. Xu, X. *et al.* Optically controlled locking of the nuclear field via coherent dark-state spectroscopy. *Nature* **459**, 1105–1109 (2009).
20. Högele, A. *et al.* Voltage-controlled optics of a quantum dot. *Phys. Rev. Lett.* **93**, 217401 (2004).
21. Atatüre, M. *et al.* Quantum-dot spin-state preparation with near-unity fidelity. *Science* **312**, 551–553 (2006).
22. Shahbazyan, T. V., Perakis, I. E. & Raikh, M. E. Spin correlations in nonlinear optical response: light-induced Kondo effect. *Phys. Rev. Lett.* **84**, 5896–5899 (2000).
23. Kikoin, K. & Avishai, Y. Many-particle resonances in excited states of semiconductor quantum dots. *Phys. Rev. B* **62**, 4647–4655 (2000).
24. Govorov, A. O., Karrai, K. & Warburton, R. J. Kondo excitons in self-assembled quantum dots. *Phys. Rev. B* **67**, 241307(R) (2003).
25. Helmes, R. W., Sindel, M., Borda, L. & von Delft, J. Absorption and emission in quantum dots: Fermi surface effects of Anderson excitons. *Phys. Rev. B* **72**, 125301 (2005).
26. Gunnarsson, O. & Schönhammer, K. Electron spectroscopies for Ce compounds in the impurity model. *Phys. Rev. B* **28**, 4315–4341 (1983).
27. Warburton, R. J. *et al.* Optical emission from a charge-tunable quantum ring. *Nature* **405**, 926–929 (2000).
28. Anderson, P. W. Localized magnetic states in metals. *Phys. Rev.* **124**, 41–53 (1961).

Author Contributions C.L., F.H., W.W. and P.F. carried out the experiments. M.H. and A.W. performed the numerical analysis. The samples were grown by S.F. C.L. and A.I. planned the experiment. H.E.T., M.C., L.G., A.I. and J.v.D. developed the theoretical framework. C.L., J.v.D. and A.I. supervised the project, carried out the analysis of the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.L. (clatta@phys.ethz.ch) or A.I. (imamoglu@phys.ethz.ch).

P5.

P6. Anderson orthogonality and the numerical renormalization group

► Andreas Weichselbaum, Wolfgang Münder, and Jan von Delft
Phys. Rev. B **84**, 075137 (2011).

Anderson orthogonality (AO) deals with the fact that, given some *local* perturbation of the underlying Hamiltonian, quantum-many-body states in the low-energy-sector quickly become orthogonal to each other in the the thermodynamic limit. Since Anderson orthogonality typically considers an arbitrary local perturbation in some effective non-interacting environment, this setup is ideally suited for the NRG. Anderson orthogonality typically manifests itself in absorption spectra, in that the absorption of a photon changes the local charge configuration. If by creating an electron-hole pair, the hole has a significantly long lifetime such that the entire system can equilibrate prior to recombination, then the latter setting corresponds to a "final Hamiltonian" amenable for AO physics. The related dynamics for complex interacting systems has been studied in much detail in Münder et al. (2011), which was just accepted for publication.

The essential simple observation of the attached paper is Eq. (P6-16) which was inspired by the standard analysis of the finite-size spectra obtained from the NRG:[12,72,75] even though one has a finite Wilson chain of *linear* length k, due to the logarithmic discretization this *mimics an exponentially large system* in terms of mean level-spacing. Starting from this, all AO related relations can be nicely demonstrated. While originally shown for non-interacting fermi ground states,[4] AO nevertheless also holds for interacting systems. For the latter, however, the overlap of ground state wave functions is far from trivial. Yet nevertheless, as we demonstrate in the paper, within the framework of the NRG this becomes a simple exercise in MPS. Overall then, the NRG provides an accurate self-contained framework that allows (1) to *directly calculate the AO exponent* related to the explicit decay of ground state wave functions in the thermodynamic limit, and compare this with known exponents in terms of (2) phase shifts, or (3) overall charge that flows to or comes in from infinity due to the local perturbation. All of this is demonstrated and discussed in detail for a range of interacting models in the attached paper.

Anderson orthogonality and the numerical renormalization group

Andreas Weichselbaum, Wolfgang Münder, and Jan von Delft

Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience, Ludwig-Maximilians-Universität,
DE-80333 Munich, Germany
(Received 20 April 2011; published 10 August 2011)

Anderson orthogonality (AO) refers to the fact that the ground states of two Fermi seas that experience different local scattering potentials, say $|G_{\mathrm{I}}\rangle$ and $|G_{\mathrm{F}}\rangle$, become orthogonal in the thermodynamic limit of large particle number N, in that $|\langle G_{\mathrm{I}}|G_{\mathrm{F}}\rangle| \sim N^{-\frac{1}{2}\Delta_{\mathrm{AO}}}$ for $N \to \infty$. We show that the numerical renormalization group offers a simple and precise way to calculate the exponent Δ_{AO}: the overlap, calculated as a function of Wilson chain length k, decays exponentially $\sim e^{-k\alpha}$, and Δ_{AO} can be extracted directly from the exponent α. The results for Δ_{AO} so obtained are consistent (with relative errors typically smaller than 1%) with two other related quantities that compare how ground-state properties change upon switching from $|G_{\mathrm{I}}\rangle$ to $|G_{\mathrm{F}}\rangle$: the difference in scattering phase shifts at the Fermi energy, and the displaced charge flowing in from infinity. We illustrate this for several nontrivial interacting models, including systems that exhibit population switching.

P6.

I. INTRODUCTION

In 1967, Anderson considered the response of a Fermi sea to a change in local scattering potential and made the following observation[1]: The ground states $|G_{\mathrm{I}}\rangle$ and $|G_{\mathrm{F}}\rangle$ of the Hamiltonians $\hat{H}_{\mathrm{I}}$ and $\hat{H}_{\mathrm{F}}$ describing the system before and after the change, respectively, become orthogonal in the thermodynamic limit, decaying with total particle number N as

$$|\langle G_{\mathrm{I}}|G_{\mathrm{F}}\rangle| \sim N^{-\frac{1}{2}\Delta_{\mathrm{AO}}^2}, \tag{1}$$

because the single-particle states comprising the two Fermi seas are characterized by different phase shifts.

Whenever the Anderson orthogonality (AO) exponent Δ_{AO} is finite, the overlap of the two ground-state wave functions goes to zero as the system size becomes macroscopic. As a consequence, matrix elements of the form $|\langle G_{\mathrm{I}}|\hat{\mathcal{O}}|G_{\mathrm{F}}\rangle|$, where $\hat{\mathcal{O}}$ is a local operator acting at the site of the localized potential, necessarily also vanish in the thermodynamic limit. This fact has far-reaching consequences, underlying several fundamental phenomena in condensed matter physics involving quantum impurity models, i.e., models describing a Fermi sea coupled to localized quantum degrees of freedom. Examples are the Mahan exciton (ME) and the Fermi-edge singularity[2–5] (FES) in absorption spectra, and the Kondo effect[6] arising in magnetic alloys[7] or in transport through quantum dots.[8] For all of these, the low-temperature dynamics is governed by the response of the Fermi sea to a sudden switch of a local scattering potential. More recently, there has also been growing interest in inducing such a sudden switch, or quantum quench, by optical excitations of a quantum dot tunnel-coupled to a Fermi sea, in which case the post-quench dynamics leaves fingerprints, characteristic of AO, in the optical absorption or emission line shape.[9–11]

The intrinsic connection of local quantum quenches to the scaling of the Anderson orthogonality with system size can be intuitively understood as follows. Consider an instantaneous event at the location of the impurity at time $t = 0$ in a system initially in equilibrium. This local perturbation will spread out spatially, such that for $t > 0$, the initial wave function is affected only within a radius $L \simeq v_f t$ of the impurity, with

v_f the Fermi velocity. The AO finite-size scaling in Eq. (1) therefore directly resembles the actual experimental situation and, in particular, allows the exponent Δ_{AO} to be directly related to the exponents seen in experimental observables at long-time scales, or at the threshold frequency in Fourier space.[12]

A powerful numerical tool for studying quantum impurity models is the numerical renormalization group (NRG),[13,14] which allows numerous static and dynamical quantities to be calculated explicitly, also in the thermodynamic limit of infinite bath size. The purpose of this paper is to point out that NRG also offers a completely straightforward way to calculate the overlap $|\langle G_{\mathrm{I}}|G_{\mathrm{F}}\rangle|$ and hence to extract Δ_{AO}. The advantage of using NRG for this purpose is that NRG is able to deal with quantum impurity models that in general also involve local *interactions*, which are usually not tractable analytically. Although Anderson himself did not include local interactions in his considerations,[1] his prediction (1) still applies, provided the ground states $|G_{\mathrm{I,F}}\rangle$ describe Fermi liquids. This is the case for most impurity models (but not all; the two-channel Kondo model is a notable exception). Another useful feature of NRG is that it allows consistency checks on its results for overlap decays since Δ_{AO} is known to be related to a change of scattering phase shifts at the Fermi surface. These phase shifts can be calculated independently, either from NRG energy flow diagrams, or via Friedel's sum rule from the displaced charge, as will be elaborated below.

A further concrete motivation for the present study is to develop a convenient tool for calculating AO exponents for quantum dot models that display the phenomenon of population switching.[15–19] In such models, a quantum dot tunnel-coupled to leads contains levels of different widths, and is capacitively coupled to a gate voltage that shifts the levels energy relative to the Fermi level of the leads. Under suitable conditions, an (adiabatic) sweep of the gate voltage induces an inversion in the population of these levels (a so-called population switch), implying a change in the local potential seen by the Fermi seas in the leads. In this paper, we verify that the method of extracting Δ_{AO} from $\langle G_{\mathrm{I}}|G_{\mathrm{F}}\rangle$ works reliably also for such models. In a separate publication,[12] we will use

this method to analyze whether AO can lead to a quantum phase transition in such models, as suggested in Ref. 19.

The remainder of this paper is structured as follows: In Sec. II, we define the AO exponent Δ_{AO} in general terms, and explain in Sec. III how NRG can be used to calculate it. Section IV presents numerical results for several interacting quantum dot models of increasing complexity: first the spinless interacting resonant level model (IRLM), then the single-impurity Anderson model (SIAM), followed by two models exhibiting population switching, one for spinless and the other for spinful electrons. In all cases, our results for Δ_{AO} satisfy all consistency checks to within less than 1%.

II. DEFINITION OF ANDERSON ORTHOGONALITY

A. AO for a single channel

To set the stage, let us review AO in the context of a free Fermi sea involving a single species or channel of noninteracting electrons experiencing two different local scattering potentials. The initial and final systems are described in full by the Hamiltonians $\hat{H}_\text{I}$ and $\hat{H}_\text{F}$, respectively. Let $\hat{c}^\dagger_{\varepsilon,X}|0\rangle$ be the single-particle eigenstates of $\hat{H}_X$ characterized by the scattering phase shifts $\delta_X(\varepsilon)$, where $X \in \{\text{I,F}\}$ and $\hat{c}^\dagger_{\varepsilon,X}$ are fermion creation operators, and let ε^f be the same Fermi energy for both Fermi seas $|G_X\rangle$. Anderson showed that in the thermodynamic limit of large particle number $N \to \infty$, the overlap

$$\langle G_\text{I}|G_\text{F}\rangle = \langle 0|\prod_{\varepsilon < \varepsilon^f} \hat{c}_{\varepsilon,\text{I}} \prod_{\varepsilon < \varepsilon^f} \hat{c}^\dagger_{\varepsilon,\text{F}}|0\rangle \qquad (2)$$

decays as in Eq. (1),[1,4] where Δ_{AO} is equal to the difference in single-particle phase shifts at the Fermi level

$$\Delta_{\text{AO}} = \Delta_{\text{ph}} \equiv [\delta_\text{F}(\varepsilon^f) - \delta_\text{I}(\varepsilon^f)]/\pi . \qquad (3)$$

The relative sign between Δ_{AO} and Δ_{ph} (+, not −) does not affect the orthogonality exponent Δ_{AO}^2, but follows standard convention [Ref. 20, Eq. (7), or Ref. 21, Eq. (21)].

In this paper, we will compare three independent ways of calculating Δ_{AO}. (i) The first approach calculates the overlap $|\langle G_\text{I}|G_\text{F}\rangle|$ of Eq. (1) explicitly as a function of (effective) system size. The main novelty of this paper is to point out that this can easily be done in the framework of NRG, as will be explained in detail in Sec. III.

(ii) The second approach is to directly calculate Δ_{ph} via Eq. (3), since the extraction of phase shifts $\delta_X(\varepsilon^f)$ from NRG finite-size spectra is well known[13]: Provided that $\hat{H}_X$ describes a Fermi liquid, the (suitably normalized) fixed point spectrum of NRG can be reconstructed in terms of equidistant free-particle levels shifted by an amount determined by $\delta_X(\varepsilon^f)$. The many-body excitation energy of an additional particle, a hole and a particle-hole pair, thus allow the phase shift $\delta_X(\varepsilon^f)$ to be determined unambiguously.

(iii) The third approach exploits Friedel's sum rule,[20] which relates the difference in phase shifts to the so-called *displaced charge* Δ_{ch} via $\Delta_{\text{ch}} = \Delta_{\text{ph}}$. Here the displaced charge Δ_{ch} is defined as the charge in units of e (i.e., the number of electrons) flowing inward from infinity into a region of large but finite

volume, say V_{large}, surrounding the scattering location, upon switching from $\hat{H}_\text{I}$ to $\hat{H}_\text{F}$:

$$\Delta_{\text{ch}} \equiv \langle G_\text{F}|\hat{n}_{\text{tot}}|G_\text{F}\rangle - \langle G_\text{I}|\hat{n}_{\text{tot}}|G_\text{I}\rangle$$
$$\equiv \Delta_{\text{sea}} + \Delta_{\text{dot}} . \qquad (4)$$

Here, $\hat{n}_{\text{tot}} \equiv \hat{n}_{\text{sea}} + \hat{n}_{\text{dot}}$, where $\hat{n}_{\text{sea}}$ is the total number of Fermi-sea electrons within V_{large}, whereas $\hat{n}_{\text{dot}}$ is the local charge of the scattering site, henceforth called "dot."

To summarize, we have the equalities

$$\Delta_{\text{AO}}^2 = \Delta_{\text{ph}}^2 = \Delta_{\text{ch}}^2 , \qquad (5)$$

where all three quantities can be calculated independently and straightforwardly within the NRG. Thus, Eq. (5) constitutes a strong consistency check. We will demonstrate below that NRG results satisfy this check with good accuracy (deviations are typically below 1%).

B. AO for multiple channels

We will also consider models involving several independent and conserved channels (e.g., spin in spin-conserving models). In the absence of interactions, the overall ground-state wave function is the product of those of the individual channels. With respect to AO, this trivially implies that each channel *adds* independently to the AO exponent in Eq. (1),

$$\Delta_{\text{AO}}^2 = \sum_{\mu=1}^{N_c} \Delta_{\text{AO},\mu}^2, \qquad (6)$$

where $\mu = 1, \ldots, N_c$ labels the N_c different channels. We will demonstrate below that the additive character in Eq. (6) generalizes to systems with *local interactions*, provided that the particle number in each channel remains conserved. This is remarkable since interactions may cause the ground-state wave function to involve entanglement between local and Fermi-sea degrees of freedom from different channels. However, our results imply that the asymptotic tails of the ground-state wave function far from the dot still factorize into a product of factors from individual channels. In particular, we will calculate the displaced charge for each individual channel [cf. Eq. (4)]

$$\Delta_{\text{ch},\mu} \equiv \langle G_\text{F}|\hat{n}_{\text{tot},\mu}|G_\text{F}\rangle - \langle G_\text{I}|\hat{n}_{\text{tot},\mu}|G_\text{I}\rangle$$
$$\equiv \Delta_{\text{sea},\mu} + \Delta_{\text{dot},\mu} , \qquad (7)$$

where $\hat{n}_{\text{tot},\mu} = \hat{n}_{\text{sea},\mu} + \hat{n}_{\text{dot},\mu}$. Assuming no interactions in the respective Fermi seas, it follows from Friedel's sum rule that $\Delta_{\text{AO},\mu}^2 = \Delta_{\text{ch},\mu}^2$, and therefore

$$\Delta_{\text{AO}}^2 = \sum_{\mu=1}^{N_c} \Delta_{\text{ch},\mu}^2 \equiv \Delta_{\text{ch}}^2 , \qquad (8)$$

where Δ_{ch}^2 is the total sum of the squares of the displaced charges of the separate channels. Equation (8) holds with great numerical accuracy, too, as will be shown below.

112

III. TREATING ANDERSON ORTHOGONALITY USING NRG

A. General impurity models

The problem of a noninteracting Fermi sea in the presence of a local scatterer belongs to the general class of quantum impurity models treatable by Wilson's NRG.[13] Our proposed approach for calculating Δ_{AO} applies to *any* impurity model treatable by NRG. To be specific, however, we will focus here on generalized Anderson impurity type models. They describe N_c different (and conserved) species or channels of fermions that hybridize with local degrees of freedom at the dot, while all interaction terms are local.

We take both the initial and final ($X \in \{I,F\}$) Hamiltonians to have the generic form $\hat{H}_X = \hat{H}_b + \hat{H}_{d,X} + \hat{H}_{\mathrm{int}}$. The first term

$$\hat{H}_b = \sum_{\mu=1}^{N_c} \sum_{\varepsilon} \varepsilon \, \hat{c}_{\varepsilon\mu}^\dagger \hat{c}_{\varepsilon\mu} \qquad (9)$$

describes a noninteracting Fermi sea involving N_c channels. (N_c includes the spin index, if present.) For simplicity, we assume a constant density of states $\rho_\mu(\varepsilon) = \rho_{0,\mu}\theta(D - |\varepsilon|)$ for each channel with half-bandwidth D. Moreover, when representing numerical results, energies will be measured in units of half-bandwidth, hence $D := 1$. The Fermi sea is assumed to couple to the dot only via the local operators $\hat{f}_{0\mu} = \frac{1}{\sqrt{N_b}} \sum_\epsilon \hat{c}_{\epsilon\mu}$ and $\hat{f}_{0\mu}^\dagger$, that, respectively, annihilate or create a Fermi-sea electron of channel μ at the position of the dot $\vec{r} = 0$, with a proper normalization constant N_b to ensure $[f_{0\mu}, f_{0\mu'}^\dagger] = \delta_{\mu\mu'}$.

The second term $\hat{H}_{d,X}$ contains the noninteracting local part of the Hamiltonian, including the dot-lead hybridization

$$\hat{H}_{d,X} = \sum_{\mu=1}^{N_c} \varepsilon_{d\mu,X} \hat{n}_{d\mu} + \sum_{\mu=1}^{N_c} \sqrt{\frac{2\Gamma_\mu}{\pi}} [\hat{d}_\mu^\dagger \hat{f}_{0\mu} + \mathrm{H.c.}]. \quad (10)$$

Here, $\varepsilon_{d\mu,X}$ is the energy of dot level μ in the initial or final configuration, and $\hat{n}_{d\mu} = \hat{d}_\mu^\dagger \hat{d}_\mu$ is its electron number. $\Gamma_\mu \equiv \pi \rho_\mu V_\mu^2$ is the effective width of level μ induced by its hybridization with channel μ of the Fermi sea, with V_μ the μ-conserving matrix element connecting the d-level with the bath states $\hat{c}_{\varepsilon\mu}$, taken independent of energy, for simplicity.

Finally, the interacting third term is given in the case of the single-impurity Anderson model (SIAM) by the uniform Coulomb interaction U at the impurity

$$\hat{H}_{\mathrm{int}}^{\mathrm{SIAM}} = \tfrac{1}{2} U \hat{n}_d (\hat{n}_d - 1), \qquad (11)$$

with $\hat{n}_d = \sum_\mu \hat{n}_{d\mu}$, while in the case of the interacting resonant-level model (IRLM), the interacting part is given by

$$\hat{H}_{\mathrm{int}}^{\mathrm{IRLM}} = U' \hat{n}_d \hat{n}_0, \qquad (12)$$

with $\hat{n}_0 = \sum_\mu f_{0,\mu}^\dagger f_{0,\mu} \equiv \sum_\mu \hat{n}_{0,\mu}$. In particular, most of our results are for the one- or two-lead versions of the SIAM for spinful or spinless electrons

$$\hat{H}_X^{\mathrm{SIAM}} = \hat{H}_b + \hat{H}_{d,X} + \hat{H}_{\mathrm{int}}^{\mathrm{SIAM}}. \qquad (13)$$

We consider either a single dot level coupled to a single lead (spinful, $N_c = 2 : \mu \in \{\uparrow, \downarrow\}$), or a dot with two levels

coupled separately to two leads (spinless, $N_c = 2 : \mu \in \{1,2\}$; spinful, $N_c = 4 : \mu \in \{1\uparrow, 1\downarrow, 2\uparrow, 2\downarrow\}$). A splitting of the energies $\varepsilon_{d\mu,X}$ in the spin label (if any) will be referred to as magnetic field B. We also present some results for the IRLM, for a single channel of spinless electrons ($N_c = 1$):

$$\hat{H}_X^{\mathrm{IRLM}} = \hat{H}_b + \hat{H}_{d,X} + \hat{H}_{\mathrm{int}}^{\mathrm{IRLM}}. \qquad (14)$$

In this paper, we focus on the case that $\hat{H}_I$ and $\hat{H}_F$ differ only in the local level positions ($\varepsilon_{d\mu,I} \neq \varepsilon_{d\mu,F}$). It is emphasized, however, that our methods are equally applicable for differences between initial and final values of any other parameters, including the case that the interactions are channel specific, e.g., $\sum_{\mu\mu'} U_{\mu\mu'} \hat{n}_{d\mu} \hat{n}_{d\mu'}$ or $\sum_{\mu\mu'} U'_{\mu\mu'} \hat{n}_{d\mu} \hat{n}_{0\mu'}$.

B. AO on Wilson chains

Wilson discretized the spectrum of $\hat{H}_b$ on a logarithmic grid of energies $\pm D\Lambda^{-k}$ (with $\Lambda > 1$, $k = 0,1,2,\ldots$), thereby obtaining exponentially high resolution of low-energy excitations. He then mapped the impurity model onto a semi-infinite "Wilson tight-binding chain" of sites $k = 0$ to ∞, with the impurity degrees of freedom coupled only to site 0. To this end, he made a basis transformation from the set of sea operators $\{\hat{c}_{\varepsilon\mu}\}$ to a new set $\{\hat{f}_{k\mu}\}$, chosen such that they bring $\hat{H}_b$ into the tridiagonal form

$$\hat{H}_b \simeq \sum_{\mu=1}^{N_c} \sum_{k=1}^{\infty} t_k (\hat{f}_{k\mu}^\dagger \hat{f}_{k-1,\mu} + \mathrm{H.c.}). \qquad (15)$$

The hopping matrix elements $t_k \propto D\Lambda^{-k/2}$ decrease exponentially with site index k along the chain. Because of this separation of energy scales for sufficiently large Λ, typically $\Lambda \gtrsim 1.7$, the Hamiltonian can be diagonalized iteratively by solving a Wilson chain of length k [restricting the sum in Eq. (15) to the first k terms] and increasing k one site at a time: Starting with a short Wilson chain, a new *shell* of many-body eigenstates for a Wilson chain of length k, say $|s\rangle_k$, is constructed from the states of site k and the M_K lowest-lying eigenstates of shell $k - 1$. The latter are the so-called *kept* states $|s\rangle_{k-1}^K$ of shell $k - 1$, while the remaining higher-lying states $|s\rangle_{k-1}^D$ from that shell are *discarded*.

The typical spacing between the few lowest-lying states of shell k, i.e., the energy scale dE_k, is set by the hopping matrix element t_k to the previous site, hence,

$$dE_k \simeq t_k \propto D\Lambda^{-k/2}. \qquad (16)$$

Now, for a noninteracting Fermi sea with N particles, the mean single-particle level spacing at the Fermi energy scales as $dE \propto D/N$. This also sets the energy scale for the mean level spacing of the few lowest-lying many-body excitations of the Fermi sea. Equating this to Eq. (16), we conclude that a Wilson chain of length k represents a Fermi sea with an actual size $L \propto N$, i.e., an *effective* number of electrons N, that grows *exponentially* with k,

$$N \propto \Lambda^{k/2}. \qquad (17)$$

Now consider two impurity models that differ only in their local terms $\hat{H}_{d,X}$, and let $|G_X\rangle_k$ be the ground states of their respective Wilson chains of length k, obtained via

two separate NRG runs.[9] Combining Anderson's prediction (1) and Eq. (17), the ground-state overlap is expected to decay exponentially with k as

$$|_k\langle G_I|G_F\rangle_k| \propto \Lambda^{-k\Delta_{AO}^2/4} \equiv e^{-\alpha k} \qquad (18)$$

with

$$\Delta_{AO}^2 = \frac{4\alpha}{\log \Lambda} \ . \qquad (19)$$

Thus, the AO exponent can be determined by using NRG to *directly* calculate the left-hand side of Eq. (18) as a function of chain length k, and extracting Δ_{AO} from the exponent α characterizing its exponential decay with k.

For *noninteracting* impurity models ($U = U' = 0$), a finite Wilson chain represents a single-particle Hamiltonian for a finite number of degrees of freedom that can readily be diagonalized numerically, without the need for implementing NRG truncation. The ground state is a Slater determinant of those single-particle eigenstates that are occupied in the Fermi sea. The overlap $\langle G_I|G_F\rangle$ is then given simply by the determinant of a matrix whose elements are overlaps between the I and F versions of the occupied single-particle states. It is easy to confirm numerically in this manner that $\langle G_I|G_F\rangle \sim e^{-\alpha k}$, leading to the expected AO in the limit $k \to \infty$. We will thus focus on interacting models henceforth, which require the use of NRG.

In the following three sections, we discuss several technical aspects needed for calculating AO with NRG on Wilson chains.

C. Ground-state overlaps

The calculation of state space overlaps within the NRG is straightforward, in principle,[9,22] especially considering its underlying matrix product state structure.[23–25] Now, the overlap in Eq. (18), which needs to be calculated in this paper, is with respect to ground states as a function of Wilson chain length k. As such, two complications can arise. (i) For a given k, the system can have several degenerate ground states $\{|s\rangle_k^X : s \in G\}$, with the degeneracy $d_{X,k}$ typically different for even and odd k. (ii) The symmetry of the ground-state space may actually differ with alternating k between certain initial and final configurations $X \in \{I,F\}$, leading to strictly zero overlap there. A natural way to deal with (i) is to essentially average over the degenerate ground-state spaces, while (ii) can be ameliorated by partially extending the ground-state space to the full kept space $\{|s\rangle_k^X : s \in K\}$, as will be outlined in the following.

The $d_{X,k}$-fold degenerate ground-state subspace is described by its projector, written in terms of the fully mixed density matrix

$$\hat\rho_{G,k}^X \equiv \frac{1}{d_{X,k}} \sum_{s\in G}^{d_{X,k}} |s\rangle_k^{X\ X}{}_k\langle s|. \qquad (20)$$

It is then convenient to calculate the overlap of the ground-state space as

$$z_{GK}^2(k) \equiv \mathrm{tr}_{K,k}^F\big(\hat\rho_{G,k}^I\big)$$
$$= \frac{1}{d_{I,k}} \sum_{s\in G}\sum_{s'\in K} \big|_k^I\langle s|s'\rangle_k^F\big|^2, \qquad (21)$$

where $\mathrm{tr}_{K,k}^F(\cdot)$ refers to the trace over the kept space at iteration k of the final system. The final expression can be interpreted, up to the prefactor, as the square of the Frobenius norm of the overlap matrix $_k^I\langle s|s'\rangle_k^F$ between the NRG states $s \in G$ and $s' \in K$ at iteration k for the initial and final Hamiltonians, respectively.

Note that the specific overlap in Eq. (21), as used throughout later in this paper, not only includes the ground space of the final system at iteration k, but rather includes the *full kept space* of that system. Yet, each such overlap scales as $e^{-\alpha k}$, with the same exponent α for all combinations of s and s', because (i) the states $|s\rangle_k^I$ with $s \in G$ are taken from the initial ground-state space, and (ii) the states $|s'\rangle_k^F$ with $s' \in K$ from the final kept shell differ from a final ground state only by a small number of excitations. Therefore, Eq. (21) is essentially equivalent, up to an irrelevant prefactor, to strictly taking the overlap of ground-state spaces as in $z_{GG}^2(k) \equiv \mathrm{tr}_{G,k}^F(\hat\rho_{G,k}^I)$. This will be shown in more detail in the following. In particular, the overlap in Eq. (21) can be easily generalized to

$$z_{PP'}^2(k) \equiv \mathrm{tr}_{P',k}^F\big(\hat\rho_{P,k}^I\big), \quad 0 \leqslant z_{PP'}^2(k) \leqslant 1 \qquad (22)$$

where $P^{(')}\in\{G,K,\infty\}$ represents the ground-state space, the full kept space, or the ground state taken at $k \to \infty$ with respect to either the initial or final system, respectively. The overlap $z_{PP'}^2(k)$ in Eq. (22) then represents the fully mixed density matrix in space P of the initial system traced over space P' of the final system, all evaluated at iteration k.

A detailed comparison for several different choices of $z_{PP'}^2(k)$, including $z_{GG}^2(k)$, is provided in Fig. 1 for the standard SIAM with $\mu \in \{\uparrow, \downarrow\}$. The topmost line (identified with legend by heavy round dot) shows the overlap Eq. (2) used as default for calculating the overlap in the rest of the paper. This measure is most convenient, as it reliably provides data with a smooth k-dependence for large k, insensitive to alternating k-dependent changes of the symmetry sector and degeneracy of the ground-state sector of $\hat H_{X,k}$ (note that the exact ground-state symmetry is somewhat relative within the NRG framework, given an essentially gapless continuum of states of the full system). The overlap z_{GG} (data marked by triangle) gives the overlap of the initial and final *ground-state* spaces, but is sensitive to changes in symmetry sector; in particular, for $k \lesssim 28$, it is nonzero for odd iterations only. The reason as to why it can be vanishingly small for certain iterations is, in the present case, that the initial and final occupancies of the local level differ significantly, as seen from the values for $\langle n_{dot}^I\rangle$ and $\langle n_{dot}^F\rangle$ specified in the panel. Therefore, initial and final ground states can be essentially orthogonal, in the worst case throughout the entire NRG run. Nonetheless, the AO exponent is expected to be well defined and finite, as reflected in z_{GK}.

The AO measure z_{KK} (data marked by star) is smooth throughout, and although it is not strictly constrained to the ground-state space at a given iteration, in either the initial or final system, it gives the correct AO exponent, the reason being the underlying energy scale separation of the NRG. Finally, $z_{\infty,K} = \mathrm{Tr}_{F,k}^K\{\hat\rho_{I,\infty}^G\}$ (data marked by squares) refers to an AO measure that calculates the overlap of the ground-state space of an essentially infinite initial system (i.e., $k \to \infty$, or in practice, the last site of the Wilson chain), with the kept space
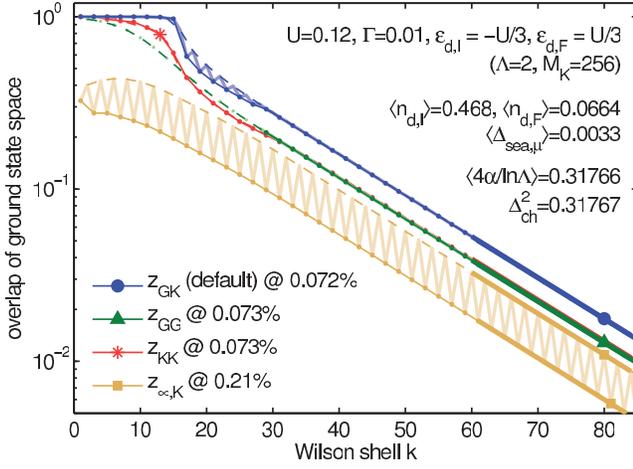
FIG. 1. (Color online) Anderson orthogonality for the spin-degenerate standard SIAM for a single lead [Eq. (10), $\mu \in \{\uparrow , \downarrow\}$], with μ-independent parameters ε_d and Γ for $\hat{H}_I$ and $\hat{H}_F$ as specified in the panel (the full ε_d^F dependence of Δ_{AO} for fixed ε_d^I is analyzed in more detail in Fig. 5). Several alternative measures for calculating the AO overlap are shown, using $z_{PP'}(k)$ in Eq. (22) with $P^{(')} \in \{G, K, \infty\}$, as defined in the text. All overlaps are plotted for even and odd iterations separately to account for possible even-odd behavior within the Wilson chain (thin solid lines with dots, and dashed lines, respectively, while heavy symbols identify lines with corresponding legends). If even and odd data from the same $z_{PP'}(k)$ do not lie on the same smooth line, the combined data are also plotted (light zigzag lines) as guides to the eye. For large k, all AO overlaps exhibit exponential decay of equal strength. Separate fits of $e^{\lambda-\alpha k}$ to even and odd sectors are shown as thick solid lines, the lengths of which indicate the fitting range used. The values for Δ_{AO}^2 extracted from these fits using Eq. (19) are in excellent agreement with the displaced charge Δ_{ch}^2, as expected from Eq. (8). The relative error is less than 1% throughout, with the detailed values specified in the legend, and $\langle 4\alpha / \ln \Lambda \rangle$ representing the averaged value with regard to the four measures considered.

at iteration k of the final system. Since the latter experiences k-dependent even-odd differences, whereas the initial density matrix $\rho_{I,\infty}^G$ is independent of k, $z_{\infty, K}$ exhibits rather strong k-dependent oscillations. Nevertheless, their envelopes for even and odd iterations separately decay with the same exponent α as the other AO measures.

In summary, Fig. 1 demonstrates that all AO measures decay asymptotically as $e^{\lambda-\alpha k}$, as expected from Eq. (18), with the *same* exponent α, independent of the details of the construction. These details only affect the constant prefactor λ, which is irrelevant for the determination of Δ_{AO}.

D. Channel-specific exponents from chains of different lengths

Equation (6) expresses the exponent Δ_{AO} of the full system in terms of the AO exponents $\Delta_{AO,\mu}$ of the individual channels. This equation is based on the assumption (the validity of which, for the models studied here, is borne out by the results presented below) that for distances sufficiently far from the dot, the asymptotic tail of the ground-state wave function factorizes, in effect, into independent products, one for each channel μ. This can be exploited to calculate, in a straightforward fashion, the individual exponent $\Delta_{AO,\mu}$ for a

given channel μ: one simply constructs a modified Wilson chain, which, in effect, is much longer for channel μ than for all others. The overlap decay for large k is then dominated by that channel.

To be explicit, the strategy is as follows. First we need to determine when a Wilson chain is "sufficiently long" to capture the aforementioned factorization of ground-state tails. This will be the case beyond that chain length, say k_0, for which the NRG energy flow diagrams for the kept space excitation spectra of the original Hamiltonians $\hat{H}_I$ and $\hat{H}_F$ are well converged to their $T = 0$ fixed point values. To calculate $\Delta_{AO,\mu}$, the AO exponent of channel μ, we then add an artificial term to the Hamiltonian that in effect depletes the Wilson chain beyond site k_0 for all other channels $\nu \neq \mu$ by drastically raising the energy cost for occupying these sites. This term has the form

$$H_{art}^\mu = C \sum_{\nu \neq \mu} \sum_{k > k_0} t_k \hat{f}_{k\nu}^\dagger \hat{f}_{k\nu}, \tag{23}$$

with $C \gg 1$. It ensures that occupied sites in the channels $\nu \neq \mu$ have much larger energy than the original energy scale t_k, so that they do not contribute to the low-energy states of the Hamiltonian. We then calculate a suitable AO measure (such as z_{GK}) using only k values in the range $k > k_0$. From the exponential decay found in this range, say $\sim e^{-\alpha_\mu k}$, the channel-specific AO exponent can be extracted [cf. Eq. (19)]:

$$\Delta_{AO,\mu}^2 = \frac{4\alpha_\mu}{\log \Lambda} . \tag{24}$$

This procedure works remarkably well, as illustrated in Fig. 2, for the spin-asymmetric single-lead SIAM of Eq. (13) (with $N_c = 2$, $\mu \in \{\uparrow , \downarrow\}$). Indeed, the values for $\Delta_{AO,\mu}$ and Δ_{AO} displayed in Fig. 2 fulfill the addition rule for squared exponents [Eq. (6)] with a relative error of less than 1%.
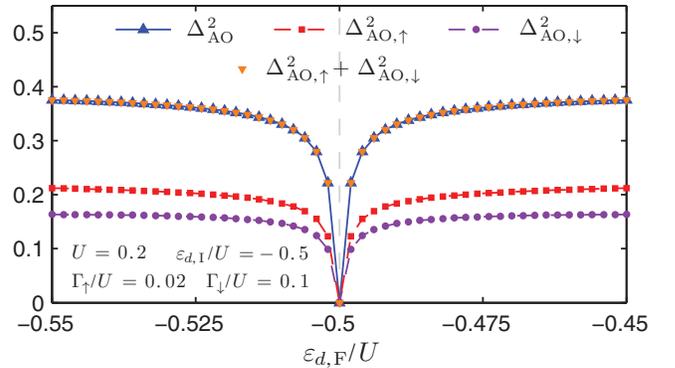

FIG. 2. (Color online) AO exponents for the standard spin-degenerate SIAM with spin-asymmetric hybridization [Eq. (13), with $\mu \in \{\uparrow , \downarrow\}$] as functions of $\varepsilon_{d,F}$ (all other parameters are fixed as specified in the panel). The vertical dashed line indicates $\varepsilon_{d,I}/U = -0.5$; at this line, the initial and final Hamiltonians are identical, hence all exponents vanish. The squared AO exponents for the individual channels $\Delta_{AO,\uparrow}^2$ (squares) and $\Delta_{AO,\downarrow}^2$ (dots) were calculated from Eq. (24). Their sum agrees (with a relative error of less than 1%) with Δ_{AO}^2 calculated from Eq. (19) (downward- and upward-pointing triangles coincide), confirming the validity of the addition rule for squared exponents in Eq. (6).

E. Displaced charge

The displaced charge $\Delta_{\text{ch},\mu}$ defined in Eq. (7) can be calculated directly within NRG. However, to properly account for the contribution from the Fermi sea $\Delta_{\text{sea},\mu}$, a technical difficulty has to be overcome: the Hamiltonians considered usually obey particle conservation and thus every eigenstate of $\hat{H}$ is an eigenstate of the total number operator, with an integer eigenvalue. Consequently, evaluating Eq. (4) over the *full* Wilson chain *always* yields an integer value for the total $\Delta_{\text{ch},\mu}$. This integer, however, does not correspond to the charge within the large but finite volume V_{large} that is evoked in the definition of the displaced charge.

To obtain the latter, we must consider subchains of shorter length. Let

$$\hat{n}_{\text{sea},\mu}^{(k)} = \sum_{k'=0}^{k} \hat{f}_{k'\mu}^{\dagger} \hat{f}_{k'\mu} \qquad (25)$$

count the charge from channel μ sitting on sites 0 to k. These sites represent, loosely speaking, a volume $V_{\text{large}}^{(k)}$ centered on the dot, the size of which grows exponentially with increasing k. The contribution from channel μ of the Fermi sea to the displaced charge within $V_{\text{large}}^{(k)}$ is

$$\Delta_{\text{sea},\mu}^{(k)} \equiv \langle G_F | \hat{n}_{\text{sea},\mu}^{(k)} | G_F \rangle - \langle G_I | \hat{n}_{\text{sea},\mu}^{(k)} | G_I \rangle, \qquad (26)$$

where $|G_I\rangle$ and $|G_F\rangle$ are the initial and final ground states of the *full-length* Wilson chain of length N ($\geqslant k$).

Figure 3 shows $\Delta_{\text{sea}}^{(k)}$ for the spinless IRLM of Eq. (14), where we dropped the index μ, since $N_c = 1$. $\Delta_{\text{sea}}^{(k)}$ exhibits even-odd oscillations between two values, say $\Delta_{\text{sea}}^{\text{even}}$ and $\Delta_{\text{sea}}^{\text{odd}}$, but these quickly assume essentially constant values over a large intermediate range of k values. Near the very end of the chain, they change again rather rapidly, in such a way that the total displaced charge associated with the full Wilson chain of length N, $\Delta_{\text{ch}}^{(N)} = \Delta_{\text{sea}}^{(N)} + \Delta_{\text{dot}}$, is an integer (see Fig. 3) because the overall ground state has well-defined particle number. Averaging the even-odd oscillations in the intermediate regime yields the desired contribution of the Fermi sea to the displaced charge $\Delta_{\text{sea}} = \frac{1}{2}(\Delta_{\text{sea}}^{\text{even}} + \Delta_{\text{sea}}^{\text{odd}})$. The corresponding result for $\Delta_{\text{ch}} = \Delta_{\text{sea}} + \Delta_{\text{dot}}$ is illustrated by the black dashed line in Fig. 3.
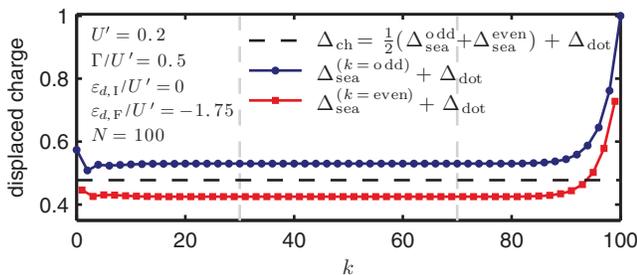

FIG. 3. (Color online) Determination of Δ_{ch}, for the interacting resonant-level model of Eq. (14), for a single specific set of parameters for $\hat{H}_I$ and $\hat{H}_F$, specified in the figure legend (the $\varepsilon_{d,F}$ dependence of Δ_{AO} for fixed $\varepsilon_{d,I}$ is analyzed in more detail in Fig. 4). We obtain Δ_{ch} (dashed line) by calculating $\Delta_{\text{sea}}^{(k)} + \Delta_{\text{dot}}$ and averaging the results for even and odd k. To reduce the influence of chain's boundary regions, we take the average over the region between the vertical dashed lines.

IV. RESULTS

In this section, we present results for the single-channel interacting resonant-level model [Eq. (14)], and for single-lead and two-lead Anderson impurity models [Eq. (13)]. These examples were chosen to illustrate that the various ways of calculating AO exponents by NRG, via Δ_{AO}, Δ_{ph}, or Δ_{ch}, are mutually consistent with high accuracy, even for rather complex (multilevel, multilead) models with local interactions. In all cases, the initial and final Hamiltonians $\hat{H}_I$ and $\hat{H}_F$ differ only in the level position: $\varepsilon_{d,I}$ is kept fixed, while $\varepsilon_{d,F}$ is swept over a range of values. This implies different initial and final dot occupations $n_{d\mu,X} = \langle G_X | \hat{n}_{d\mu} | G_X \rangle$, and hence different local scattering potentials, causing AO.

AO exponents are obtained as described in the previous sections: We calculate the AO measure $z_{GK}(k)$ using Eq. (2), obtaining exponentially decaying behavior (as in Fig. 1). We then extract α by fitting to $e^{-\alpha k}$ and determine Δ_{AO} via Eq. (19). In the figures below, the resulting Δ_{AO}^2 is shown as function of $\varepsilon_{d\mu,F}$, together with Δ_{ch}^2, and also Δ_{ph}^2 in Fig. 4. The initial dot level position $\varepsilon_{d\mu,I}$ is indicated by a vertical dashed line. When $\varepsilon_{d\mu,F}$ crosses this line, the initial and final Hamiltonians are identical, so that all AO exponents vanish. To illustrate how the changes in $\varepsilon_{d\mu,F}$ affect the dot, we also plot the occupancies $n_{d\mu,F}$ of the dot levels.

A. Interacting resonant-level model

We begin with a model for which the contribution of the Fermi sea to the displaced charge is rather important, namely, the spinless fermionic interacting resonant-level model [Eq. (14), $N_c = 1$]. The initial and final Hamiltonians $\hat{H}_I^{\text{IRLM}}$ and $\hat{H}_F^{\text{IRLM}}$ differ only in the level position: the initial one is kept fixed at $\varepsilon_{d,I} = 0$, while the final one is swept over a range of values, $\varepsilon_{d,F} \in [-1,1]$. The results
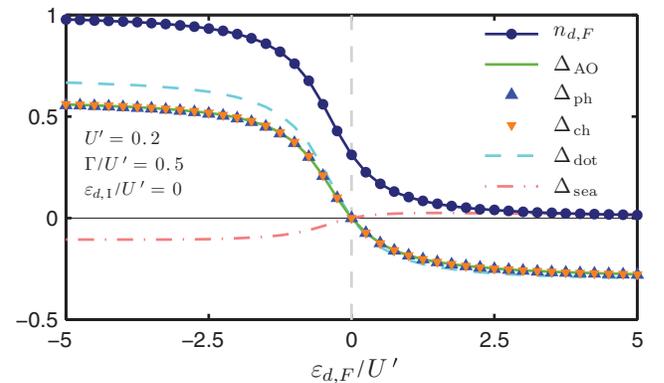

FIG. 4. (Color online) Verification that $\Delta_{\text{AO}} = \Delta_{\text{ph}} = \Delta_{\text{ch}}$ [Eq. (5)] for the spinless fermionic interacting resonant-level model [Eq. (14)]. All quantities are plotted as functions of $\varepsilon_{d,F}$, with all other parameters fixed (as specified in the panel). The vertical dashed line indicates $\varepsilon_{d,I}/U' = 0$. Heavy dots indicate the final occupation of the dot n_d. The exponent Δ_{AO} (light solid line) agrees well with Δ_{ph} and Δ_{ch} (triangles), with relative errors of less than 1%. The local and Fermi-sea contributions to the displaced charge Δ_{ch} are plotted separately, namely, Δ_{dot} (dashed line) and Δ_{sea} (dashed-dotted line). The latter is determined according to the procedure illustrated, for $\varepsilon_{d,F}/U' = -1.75$, in Fig. 3.

are shown in Fig. 4. The final dot occupancy $n_{d,F}$ (heavy dots) varies from $\simeq 1$ to 0, and $\Delta_{dot} = n_{d,F} - n_{d,I}$ (dashed line) decreases accordingly, too. The total displaced charge $\Delta_{ch} = \Delta_{dot} + \Delta_{sea}$ (downward-pointing triangles) decreases by a smaller amount since the depletion of the dot implies a reduction in the strength of the local Coulomb repulsion felt by the Fermi sea, and hence an increase in Δ_{sea} (dashed-dotted line). Throughout these changes, Δ_{AO}, Δ_{ph}, and Δ_{ch} mutually agree with errors of less than 1%, confirming that NRG results comply with Eq. (5) to high accuracy.

B. Single-impurity Anderson model

Next we consider the standard spin-degenerate SIAM for a single lead [Eq. (13), $\mu \in \{\uparrow, \downarrow\}$] with $\varepsilon_{d,\mu} = \varepsilon_d$ and $\Gamma_\mu = \Gamma$. This model exhibits well-known Kondo physics, with a strongly correlated many-body ground state.

In this model, the dot and Fermi sea affect each other only by hopping, and there is no direct Coulomb interaction between them ($U' = 0$). Hence, the contribution of the Fermi sea to the displaced charge is nearly zero, $\Delta_{sea} \simeq 0$. Apart from very small even-odd variations for the first ~ 35 bath sites corresponding to the Kondo scale, the sites of the Wilson chain are half-filled on average to a good approximation. Therefore, $\Delta_{sea} \ll \Delta_{dot}$ (explicit numbers are specified in the figure panels; see also Fig. 1), so that $\Delta_{ch,\mu}$ in Eq. (7) is dominated by the change of dot occupation only,[21]

$$\Delta_{ch}^2 \simeq \Delta_{dot}^2 \equiv \sum_\mu (n_{d\mu,F} - n_{d\mu,I})^2. \tag{27}$$

As a consequence, despite the neglect of Δ_{sea} in some previous works involving Anderson impurity models, the Friedel sum rule ($\Delta_{ph} = \Delta_{ch}$) was nevertheless satisfied with rather good accuracy (typically with errors of a few percent). However, despite being small, Δ_{sea} in practice is on the order of $|\Delta_{sea}| \lesssim \Gamma/D$ and thus *finite*. Therefore, the contribution of Δ_{sea} to Δ_{ch} will be included throughout, while also indicating the overall smallness of Δ_{sea}. In general, this clearly improves the accuracy of the consistency checks in Eq. (5), reducing the relative errors to well below 1%.

The Anderson orthogonality is analyzed for the SIAM in detail in Fig. 5. The initial system is kept fixed at the particle-hole symmetric point $\varepsilon_{d,I} = -U/2$ [indicated also by vertical dashed line in Fig. 5(a)], where the initial ground state is a Kondo singlet. The final system is swept from double to zero occupancy by varying $\varepsilon_{d,F}/U$ from -2 to 1. The final ground state is a Kondo singlet in the regime $n_{d\mu,F} \simeq 1/2$, corresponding to the intermediate shoulder in Fig. 5(a). Figure 5(b) shows the AO measure $z_{GK}(k)$ as function of k, for a range of different values of $\varepsilon_{d,F}$. Each curve exhibits clear exponential decay for large k (as in Fig. 1) of the form $e^{\lambda - \alpha k}$. The prefactor, parametrized by λ, carries little physical significance, as it also depends on the specific choice of $z_{PP'}$; its dependence on $\varepsilon_{d,F}$ is shown as a thick gray dashed line in Fig. 5(a), but it will not be discussed any further. In contrast, the decay exponent α directly yields the quantity of physical interest, namely, the AO exponent Δ_{AO}^2 via Eq. (19). Figure 5(a) compares the dependence on $\varepsilon_{d,F}$ of Δ_{AO}^2 (dashed line) with that of the displaced charge Δ_{ch}^2 (light thick line), that was calculated independently from Eqs. (7)
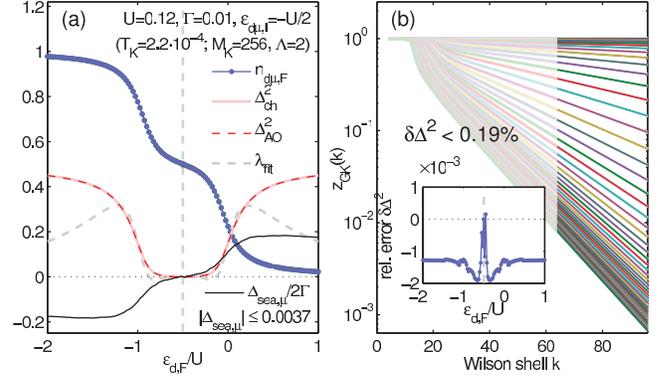

FIG. 5. (Color online) Anderson orthogonality for the single-lead, spin-symmetric SIAM [Eq. (13), with parameters as specified in the legend]. The energy of the d-level of the final system $\varepsilon_{d,F}$ is swept past the Fermi energy of the bath, while that of the initial reference system is kept fixed in the Kondo regime at $\varepsilon_{d,I} = -U/2$, indicated by vertical dashed line in panel (a) and in the inset to panel (b). Panel (a) shows, as function of $\varepsilon_{d,F}$, the dot occupation per spin $n_{d\mu}$ (dotted solid line), the contribution to the displaced charge by the Fermi sea $\Delta_{sea\mu}^2$ (thin black line), the displaced charge Δ_{ch}^2 (light solid line), and the parameters of the large-k exponential decay $e^{\lambda - \alpha k}$ of $z_{GK}(k)$ as extracted from panel (b), namely, λ (thick dashed line) and Δ_{AO} (dark dashed line), derived from α via Eq. (19). Panel (b) shows the AO measure $z_{GK}(k)$ in Eq. (2) (light lines) for the range of $\varepsilon_{d,F}$ values used in panel (a). The heavy lines shown on top for $k \geqslant 64$ are exponential fits, the results of which are summarized in panel (a). The inset shows the relative error in the AO exponents $\delta\Delta^2 \equiv (\Delta_{AO}^2 - \Delta_{ch}^2)/\Delta_{ch}^2$, i.e., the deviation between the light solid and dark dashed curves in panel (a); this error is clearly less than 1% over the full range of ε_d analyzed.

and (8). As expected from Eq. (5), they agree very well: the relative difference between the two exponents Δ_{AO}^2 and Δ_{ch}^2 is clearly below 1% throughout the entire parameter sweep, as shown in the inset of Fig. 5(b).

The contribution of the Fermi sea to the displaced charge is close to negligible, yet finite throughout [black line in Fig. 5(a)]. Overall, $\Delta_{sea} \lesssim 0.0037$, as indicated in Eq. (27). Nevertheless, by including it when calculating Δ_{ch}, the relative error $\delta\Delta^2$ is systematically reduced from a few percent to well below 1% throughout, thus underlining its importance.

C. Multiple channels and population switching

Figure 6 analyzes AO for lead-asymmetric two-level, two-lead SIAM models, with Hamiltonians of the form Eq. (13) (explicit model parameters are specified in the panels). Figure 6(a) considers a spinless case ($N_c = 2$, $\mu = j \in \{1,2\}$), the dot levels of which have mean energy ε_d at fixed splitting δ,

$$\varepsilon_{d1} = \varepsilon_d - \delta/2, \quad \varepsilon_{d2} = \varepsilon_d + \delta/2. \tag{28a}$$

Figure 6(b) considers a spinful case [$N_c = 4$, $\mu = (j\sigma)$ with $j \in \{1,2\}$, $\sigma \in \{\uparrow, \downarrow\}$], where both the lower and upper levels have an additional (small) spin splitting $B \ll \delta$,

$$\varepsilon_{dj\uparrow} = \varepsilon_{dj} + B/2, \quad \varepsilon_{dj\downarrow} = \varepsilon_{dj} - B/2. \tag{28b}$$

Charge is conserved in each of the N_c channels since these only interact through the interaction on the dot. In both models,
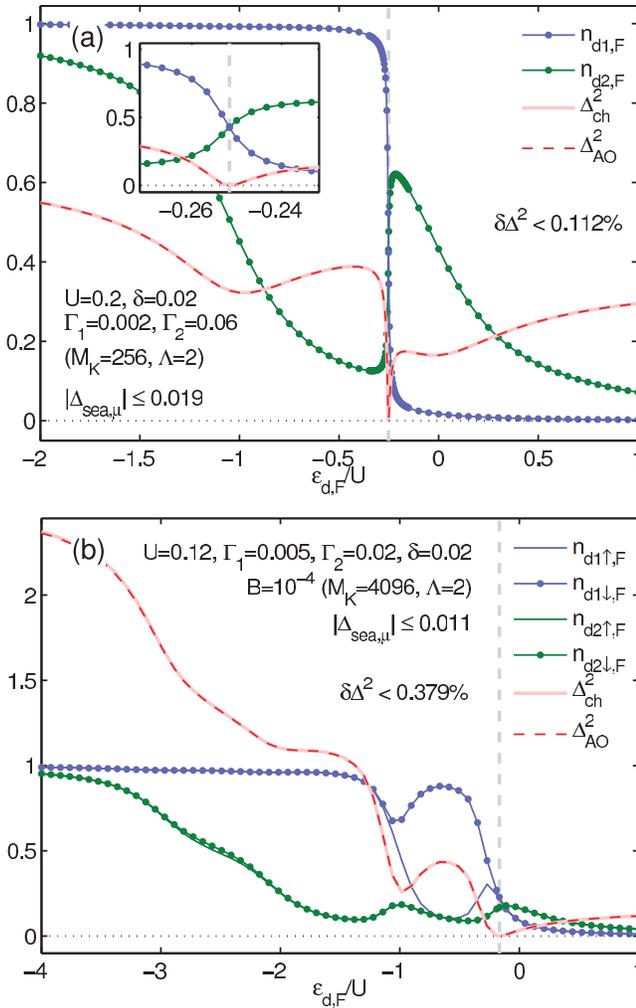
117

P6.

FIG. 6. (Color online) Anderson orthogonality for a spinless (a) and spinful (b) two-lead SIAM, with dot levels of unequal width and a split level structure as defined in Eq. (28) (all relevant model parameters are specified in the legends). In both cases, the higher level 2 is broader than the lower level 1 ($\Gamma_2 > \Gamma_1$), leading to population switching as function of the average final level energy $\varepsilon_{d,F}$. The fixed value of $\varepsilon_{d,I}$ is indicated by the vertical dashed line. The inset to panel (a) shows a zoom into the switching region, clearly demonstrating that population switching occurs smoothly. For panel (b), a finite magnetic field B causes a splitting between spin-up and spin-down levels, resulting in a more complex switching pattern. In both panels, Δ_{AO}^2 and Δ_{ch} agree very well throughout the sweep, with a relative error $\delta\Delta^2$ well below 1%.

the upper level 2 is taken to be broader than the lower level 1, $\Gamma_2 > \Gamma_1$ (for detailed parameters, see figure legends). As a consequence,[15–19] these models exhibit population switching: When $\varepsilon_{d,F}$ is lowered (while all other parameters are kept fixed), the final state occupancies of upper and lower levels cross, as seen in both panels of Fig. 6.

Consider first the spinless case in Figure 6(a). The broader level 2 shows larger occupancy for large positive $\varepsilon_{d,F}$. However, once the narrower level 1 drops sufficiently far below the Fermi energy of the bath as $\varepsilon_{d,F}$ is lowered, it becomes energetically favorable to fill level 1, while the

Coulomb interaction will cause the level 2 to be emptied. At the switching point, occupations can change extremely fast, yet they do so smoothly, as shown in the zoom in the inset to Fig. 6(a).

Similar behavior is seen for the spinful case in Fig. 6(b), although the filling pattern is more complex, due to the nonzero applied finite magnetic field B (parameters are listed in the legend). The occupations $n_{d1\sigma}$ of the narrower level 1 show a strong spin asymmetry since the magnetic field is comparable, in order of magnitude, to the level width ($B = \Gamma_1/2$). This asymmetry affects the broader level 2, which fills more slowly as ε_d is lowered. Due to the larger width of level 2, the asymmetry in its spin-dependent occupancies is significantly weaker. As in Fig. 6(a), population switching between the two levels occurs: as the narrower level 1 becomes filled, the broader level 2 gets depleted.

The details of population switching, complicated as they are [extremely rapid in Fig. 6(a) and involving four channels in Fig. 6(b)] are not the main point of Fig. 6. Instead, its central message is that despite the complexity of the switching pattern, the relation $\Delta_{AO}^2 = \Delta_{ch}^2$ is satisfied with great accuracy throughout the sweep (compare light thick and dark dashed lines). Moreover, since Δ_{ch} was calculated by adding the contributions from separate channels according to Eq. (8), this also confirms the additive character of AO exponents for separate channels.

As was the case for the single-channel SIAM discussed in Sec. IV B above, a direct interaction between dot and Fermi sea is not present in either of the models considered here ($U' = 0$). Consequently, the displaced charge Δ_{ch} is again dominated by Δ_{dot}, with $\Delta_{sea} \ll \Delta_{dot}$ [cf. Eq. (27)]. Specifically, for the spinless or spinful models, we find $\Delta_{sea} < 0.019$ or 0.011, respectively, for the entire sweep.

V. SUMMARY AND OUTLOOK

In summary, we have shown that NRG offers a straightforward, systematic, and self-contained way for studying Anderson orthogonality, and illustrated this for several interacting quantum impurity models. The central idea of our work is to exploit the fact that NRG allows the size dependence of an impurity model to be studied, in the thermodynamic limit of $N \to \infty$, by simply studying the dependence on Wilson chain length k. Three different ways of calculating AO exponents have been explored, using wave-function overlaps (Δ_{AO}), changes in phase shift at the Fermi surface (Δ_{ph}), and changes in displaced charge (Δ_{ch}). The main novelty in this paper lies in the first of these, involving a direct calculation of the overlap of the initial and final ground states themselves. This offers a straightforward and convenient way for extracting the overall exponent Δ_{AO}. Moreover, if desired, it can also be used to calculate the exponents $\Delta_{AO,\mu}$ associated with individual channels, by constructing a Wilson chain that is longer for channel μ than for the others. We have also refined the calculation of Δ_{ch} by showing how the contribution Δ_{sea} of the Fermi sea to the displaced charge can be taken into account in a systematic fashion.

The resulting exponents Δ_{AO}, Δ_{ph}, and Δ_{ch} agree extraordinarily well, with relative errors of less than 1% for a wide range of Λ. In particular, we have checked in the context of Fig. 1

that the resulting relative errors remain this small for a range of Λ values between 1.7 and 8.0. Moreover, this accuracy can be achieved using a remarkably small number of kept states M_K. For example, for the spinful SIAM analyzed above, for $\Lambda = 2$, a better than 5% agreement can be obtained already for $M_K \geqslant 32$. (For comparison, typically $M_K = 250$ is required to obtain an accurate description of the Kondo resonance of the d-level spectral function in the local moment regime of this model.)

Our analysis has been performed on models exhibiting Fermi liquid statistics at low temperatures. As an outlook, it would be interesting to explore to what extent the non-Fermi liquid nature of a model would change AO scaling properties, an example being the symmetric spinful two-channel Kondo model.

Finally, we note that nonequilibrium simulations of quantum impurity models in the time domain in response to quantum quenches are a highly interesting topic for studying AO physics in the time domain. The tools to do so using NRG have become accessible only rather recently.[10,22,23,26] One considers a sudden change in some local term in the Hamiltonian and studies the subsequent time evolution, characterized, for example, by the quantity $\langle G_I | e^{-i\hat{H}_F t} | G_I \rangle$. Its numerical evaluation requires the calculation of overlaps of eigenstates of $\hat{H}_I$ and $\hat{H}_F$. The quantity of present interest $|\langle G_I | G_F \rangle|$ is simply a particular example of such an overlap. As a consequence, the long-time decay of $\langle G_I | e^{-i\hat{H}_F t} | G_I \rangle$ is often governed by Δ_{AO}, too,[3,5] showing power-law decay in time with an exponent depending on Δ_{AO}. This will be elaborated in a separate publication.[12]

[1] P. W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).

[2] G. D. Mahan, Phys. Rev. **153**, 882 (1967).

[3] K. D. Schotte and U. Schotte, Phys. Rev. **182**, 479 (1969).

[4] K. D. Schotte and U. Schotte, Phys. Rev. **185**, 509 (1969).

[5] P. Nozières, J. Gavoret, and B. Roulet, Phys. Rev. **178**, 1084 (1969).

[6] G. Yuval and P. W. Anderson, Phys. Rev. B **1**, 1522 (1970).

[7] J. Kondo, Prog. Theor. Phys. **32**, 37 (1964).

[8] D. Goldhaber-Gordon, J. Göres, M. A. Kastner, H. Shtrikman, D. Mahalu, and U. Meirav, Phys. Rev. Lett. **81**, 5225 (1998).

[9] R. W. Helmes, M. Sindel, L. Borda, and J. von Delft, Phys. Rev. B **72**, 125301 (2005).

[10] H. E. Türeci, M. Hanl, M. Claassen, A. Weichselbaum, T. Hecht, B. Braunecker, A. Govorov, L. Glazman, A. Imamoglu, and J. von Delft, Phys. Rev. Lett. **106**, 107402 (2011).

[11] C. Latta, F. Haupt, M. Hanl, A. Weichselbaum, M. Claassen, W. Wuester, P. Fallahi, S. Faelt, L. Glazman, J. von Delft, H. E. Türeci, and A. Imamoglu, e-print arXiv:1102.3982v1.

[12] W. Münder, A. Weichselbaum, M. Goldstein, Y. Gefen, and J. von Delft (unpublished).

[13] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[14] R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

[15] G. Hackenbroich, W. D. Heiss, and H. A. Weidenmüller, Phys. Rev. Lett. **79**, 127 (1997); R. Baltin, Y. Gefen, G. Hackenbroich, and H. A. Weidenmüller, Eur. Phys. J. B **10**, 119 (1999).

[16] P. G. Silvestrov and Y. Imry, Phys. Rev. Lett. **85**, 2565 (2000).

[17] D. I. Golosov and Y. Gefen, Phys. Rev. B **74**, 205316 (2006).

[18] C. Karrasch, T. Hecht, A. Weichselbaum, Y. Oreg, J. von Delft, and V. Meden, Phys. Rev. Lett. **98**, 186802 (2007).

[19] M. Goldstein, R. Berkovits, and Y. Gefen, Phys. Rev. Lett. **104**, 226805 (2010).

[20] J. Friedel, Can. J. Phys. **34**, 1190 (1956).

[21] D. C. Langreth, Phys. Rev. **150**, 516 (1966).

[22] F. B. Anders and A. Schiller, Phys. Rev. B **74**, 245113 (2006).

[23] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[24] A. Weichselbaum, F. Verstraete, U. Schollwock, J. I. Cirac, and J. von Delft, Phys. Rev. B **80**, 165117 (2009).

[25] U. Schollwöck, Ann. Phys. (NY) **326**, 96 (2011).

[26] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005).

P6.

P7. General framework for non-abelian symmetries in tensor networks

▶ Andreas Weichselbaum (cond-mat/1202.5664v1; submitted to Annals of Physics).

Symmetries allow to make numerical calculations efficient. Given the symmetries of a system, this usually implies that many matrix elements of its corresponding Hamiltonian are strictly zero, hence the Hamiltonian becomes sparse. Given orthonormal (effective) state spaces then, abelian symmetries such as $U(1)$ symmetries like charge conservation, allow to efficiently collect non-zero matrix elements into dense blocks. As a consequence, the original sparse representation of a Hamiltonian becomes block-diagonal. By dealing with the set of these blocks sequentially rather than the full original matrix, this captures the effect of abelian symmetries on numerical efficiency.

Moreover, given non-abelian symmetries, many matrix elements despite being non-zero, are actually *not independent* of each other. That is, by knowing one non-zero matrix elements, typically a range of other matrix elements can be predicted [*cf.* Wigner Eckart theorem]. Therefore by shifting from a state space representation to multiplet representation, the *effective* dimension of the Hamiltonian can be significantly reduced. The corresponding Clebsch-Gordan algebra is split off in terms of a tensor product, which again can be dealt with separately. This captures the essence of non-abelian symmetries in numerical calculations.

These observations are basic and simple. Nevertheless, implementing these symmetries in a generic fashion, in practice, generates a significant amount of coding overhead through book-keeping. It is probably fair to say, that 99% of the coding is spent on the treatment of symmetries themselves. The payoff, of course, is an orders of magnitude faster code. Given the significant overhead in book-keeping, however, a transparent framework and setup is absolutely crucial.

Nevertheless, the non-abelian symmetry $SU(2)$ has already been widely used in the past. In particular, it was quintessential for the NRG from its very beginning,[72] dictated by numerical resources. Within the NRG, however, the setup is still relatively simple: given multiplets in some large effective basis for the Wilson chain up to some site n, a new small local state space σ_{n+1} is added. The resulting multiplet structure needs to be determined. In contrast, within the DMRG, there is usually always *another large effective state space* (say the environment), which implies that locally within a DMRG step, three state spaces must be combined into a proper total symmetry. For non-abelian symmetries this makes things significantly more complicated already as compared to the NRG. As a consequence and also considering the exceedingly abundant numerical resources, significantly fewer numerical applications exist within the DMRG that use the non-abelian symmetry $SU(2)$.[39]

Along that spirit, my first implementation of symmetries was based on abelian symmetries only. This, however, in a completely generic setting: a tensor library of arbitrary rank that can deal with an arbitrary number of abelian $U(1)$ symmetries. These codes had been very successful applied in a multitude of projects since they were developed. Never-

theless, challenging models[16] had been on the horizon, where abelian symmetries alone (1) were very close to insufficient, and (2) required excessive amount of numerical resources. This served as the motivation to think about non-abelian symmetries more carefully. The symmetry $SU(2)$, however, is special in that respect, since one does not have to worry about *inner and outer multiplicity*, which are present all the time in a more generic non-abelian setting, such as $SU(N)$ for $N > 2$. Hence a transparent framework for *arbitrary* non-abelian symmetries was highly desirable.[3]

An important turning point was the simple realization by Singh et al. (2010) that Clebsch-Gordan coefficients (CGC) factorize in general. This has been the basis of the generalization of my original QSpace library to arbitrary non-abelian symmetries. The CGC spaces are computed and incorporated explicitly. By starting from the Lie algebra with its defining matrix representation, the CGCs are determined by the commutator relations of the Lie algebra itself. This allows to easily extend the setup to new symmetries, including symmetries where analytic formulas in its description are few. CGCs are considered here generically as the unitary transformation that decomposes a tensor product of irreducible multiplets into proper global symmetry multiplets. As such the concept of CGCs can be applied to any symmetry, including abelian or point symmetries.

The QSpace library is well-equipped to deal with a large range of typical operations in tensor networks. The global Hamiltonian, however, as well as reduced density matrices, by construction are scalar operators. Therefore one may argue that by constraining oneself to a very specific set of contractions, actually the $6j$ symbols for a given symmetry should suffice for a large range of applications. In this context, the explicit treatment of CGCs can be replaced by so-called fusion rules, indeed.[34,60] Nevertheless, this (1) appears more constrictive, while (2) the overhead for the Clebsch-Gordan coefficient spaces for symmetries up to rank-2 [such as $SU(3)$ or $Sp(2)$] is negligible. Rank-3 symmetries, such as $Sp(6)$ are numerically still feasible as demonstrated in the attached paper, but there the multiplet dimensions in the CGCs already quickly grows to several thousands, thus becoming comparable to the actual dimension of the reduced multiplet spaces. Hence the overhead of CGCs is no longer negligible. Therefore, only for large-rank symmetries the $6j$ symbols really become of interest. There, however, it is unclear whether the $6j$ symbols can be readily computed from known analytical formulas without explicit reference to CGCs.

---

[3]Since NRG as well as DMRG are built from sites, *i.e.* small local state spaces, the arbitrary non-abelian symmetries discussed here are limited to compact symmetries, *i.e.* with finite-dimensional irreducible representations.

General framework for non-abelian symmetries in tensor networks

A. Weichselbaum

Physics Department, Arnold Sommerfeld Center for Theoretical Physics,
and Center for NanoScience, Ludwig-Maximilians-Universität, 80333 Munich, Germany
(Dated: February 28, 2012)

A general framework for non-abelian symmetries is presented for matrix-product and tensor-network states in the presence of local orthonormal basis sets. The two crucial ingredients, the Clebsch-Gordan algebra for multiplet spaces as well as the Wigner-Eckart theorem for operators, are accounted for in a natural, well-organized, and computationally straightforward way. The unifying tensor-representation for quantum symmetry spaces, dubbed QSpace, is particularly suitable to deal with standard renormalization group algorithms such as the numerical renormalization group (NRG), the density matrix renormalization group (DMRG), or also more general tensor networks such as the multi-scale entanglement renormalization ansatz (MERA). In this paper, the focus is on the application of the non-abelian framework within the NRG. A detailed analysis is given for a fully screened spin-3/2 three-channel Anderson impurity model in the presence of conservation of total spin, particle-hole symmetry, and SU(3) channel symmetry. The same system is analyzed using several alternative symmetry scenarios. This includes the more traditional symmetry setting $\mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)^{\otimes 3}_{charge}}$, the larger symmetry $\mathrm{SU(2)_{spin}} \otimes \mathrm{U(1)_{charge}} \otimes \mathrm{SU(3)_{channel}}$, and their much larger enveloping symplectic symmetry $\mathrm{SU(2)_{spin}} \otimes \mathrm{Sp(6)}$. These three symmetry settings are compared in detail, including their respective dramatic gain in numerical efficiency. In the appendix, finally, an extensive introduction to non-abelian symmetries is given for practical applications, together with simple self-contained numerical procedures to obtain Clebsch-Gordan coefficients and irreducible operators sets. The symmetries considered can consist of abelian symmetries together with arbitrary non-abelian symmetries with compact, *i.e.* finite-dimensional, semi-simple Lie algebras.

PACS numbers: 02.70.-c, 05.10.Cc, 75.20.Hr, 78.20.Bh

P7.

TABLE OF CONTENTS

I. INTRODUCTION

Numerical methods for strongly correlated quantum-many-body systems are confronted with exponentially large Hilbert spaces. With a limited number of exact analytical solution at hand and with perturbative treatments for low-energy or ground-state physics often insufficient, a certain systematic treatment with respect the Hilbert space is required. Besides quantum Monte Carlo approaches, that explore quantum systems stochastically,[1] a systematic state space decimation is provided by renormalization group (RG) techniques such as the density matrix renormalization group (DMRG)[2]

or the numerical renormalization group (NRG),[3] both highly efficient for quasi-one-dimensional systems, and since non-perturbative, considered essentially exact.

Both methods, DMRG as well as NRG, are based on the same algebraic structure of matrix product states (MPS).[4,5] Initially introduced for one-dimensional systems with MPS owing its name to this case, a wide range of activity has emerged within recent years to generalize MPS to tensor-networks for two- or higher-dimensional systems.[6–9] While clearly appealing from the point of area laws for entanglement-entropy,[10–12] tensor network states (TNS) often share the same disadvantage as linear systems with periodic boundary conditions within the

DMRG, namely that state spaces become intrinsically non-orthogonal. Therefore also the unique association of symmetry labels with each index in a tensor is compromised. This, however, can be circumvented by introducing an emerging extra-dimension, which is at the basis of the recently developed multi-scale entanglement renormalization ansatz (MERA).[13,14] Nevertheless, the *traditional* DMRG approach applied to 2D systems[15] with open or cylindrical boundary conditions yet with long-range interactions has continued to provide a highly competitive, extremely well-controlled, even though numerically expensive approach.

Within both, traditional DMRG as well as NRG, state spaces of entire blocks are built iteratively by adding and merging one site at a time. Clearly, the single index describing an effective basis for the entire block or site can be chosen orthogonal. Moreover, the basis states can be labeled in terms of the symmetries of the underlying Hamiltonian. Operators written as matrix elements in this very same basis therefore also share the same well-defined partitioning in terms of symmetry sectors. By grouping symmetry state spaces together, the Hamiltonian becomes block-diagonal, while general operators usually obey well-defined selection rules between symmetry sectors. Consequently, the sparsity of these operators due to symmetry can be efficiently and exactly included in the numerical description, such that usually only a few dense data blocks with non-zero matrix elements remain, given the symmetry constraints. While this well represents the advantage of implementing generic abelian or point symmetries in a calculation, the presence of non-abelian symmetries offers yet another strong simplification: many of the non-zero matrix elements are actually *not* independent of each other, bearing in mind, for example, the Wigner-Eckart theorem. Therefore going beyond abelian symmetries, non-abelian symmetries allow to significantly *compress* the non-zero blocks in terms of multiplet spaces,[16,17] while also reducing their number. With the Clebsch-Gordan coefficient spaces factorizing,[9] they can be split off systematically in terms of a tensor-product and dealt with separately.

MPS is optimal for one-dimensional systems. When exploring systems that are not strictly one dimensional but acquire *width*, such as ladders of several rungs in DMRG or multi-channel models in NRG, the price to be paid for orthonormal state spaces is that one must represent the system as a one-dimensional MPS nevertheless. This introduces longer-range interactions to the mapped 1D system, with the effect that the typically required dimensions of the state spaces to be kept in a calculation, grow roughly exponentially with system width. The number of symmetries then that (i) are available and (ii) are also be exploited in practice, decides whether or not a calculation is feasible. Abelian symmetries such as particle (charge Q) or spin (S_z) conservation are usually implemented in DMRG calculations. However, only very few groups have implemented non-abelian symmetries, and these are also constrained to SU(2) symmetries only,[16] due to its complexity in the actual implementation. General treatment of non-abelian symmetries within the MERA, on the other hand, is currently under development.[9,18] NRG, in contrast, had been set up including non-abelian SU(2) spin symmetry from its very beginning,[3] dictated by limited numerical resources. So far, however, only a very few isolated attempts including more complex non-abelian settings exist within the NRG,[19] while to our knowledge there exists no general realization yet of arbitrary non-abelian symmetries in either method.

This paper focuses on the systematic description and implementation of non-abelian symmetries of a given Hamiltonian within the generalized MPS framework. This naturally also does include the description of abelian symmetries where necessary, as they can be trivially written in terms of Clebsch-Gordan coefficients. While the focus within non-abelian symmetries belongs to SU(N) and the symplectic group Sp($2n$), the generalization to other non-abelian symmetries or also point groups is straightforward once their particular Clebsch-Gordan coefficients are worked out. In contrast to the well-known SU(2) then, general non-abelian symmetries, such as SU($N \geq 3$), represent a significant increase in algorithmic complexity, in that they can and routinely do exhibit inner and outer multiplicity. The latter, for example, implies that in the decomposition of the tensor-product of two irreducible representations (IREPs) into a direct sum of IREPs, the same IREP may occur *multiple* times. Nevertheless, this can be dealt with properly on the algorithmic level, as will be shown in detail in this paper.

While the presented non-abelian framework for general tensors is straightforwardly applicable to traditional DMRG as well as NRG, the paper focuses on the application within the NRG. Detailed results are presented for a fully screened spin-3/2 Anderson impurity model with SU(3) channel-symmetry [*i.e.* see Hamiltonian in Eq. (25)]. This model has been suggested as the effective microscopic Kondo model for iron impurities in gold or silver[20], historically the first system where Kondo physics was observed experimentally.[21,22] Being a true three-channel system, this cannot be trivially rotated into a simpler configuration of fewer relevant channels. The result is an extremely challenging calculation within the NRG that requires non-abelian symmetries for fully converged numerical results for reasonable coarse-graining of the continuous bath. The non-abelian symmetries present in the model considered are (i) particle-hole symmetry in each of the three channels, SU(2)$_\text{charge}^{\otimes 3}$, (ii) total spin symmetry, SU(2)$_\text{spin}$, and (iii) channel symmetry, SU(3)$_\text{channel}$. The non-abelian particle-hole SU(2) symmetry, however, does *not* commute with the channel SU(3) symmetry, while the plain abelian charge U(1) symmetry does commute. Overall, this suggests a larger enveloping symmetry, which turns out to be the symplectic symmetry Sp(6) [for an introduction, see App. A 10]. With this, the following symmetry scenarios are consid-

P7.

ered and compared in detail,

$$\text{SU}(2)_\text{spin} \otimes \text{SU}(2)_\text{charge}^{\otimes 3} ,$$
$$\text{SU}(2)_\text{spin} \otimes \text{U}(1)_\text{charge} \otimes \text{SU}(3)_\text{channel} , \text{ and}$$
$$\text{SU}(2)_\text{spin} \otimes \text{Sp}(6).$$

While the first setting represents a more traditional setup based on multiple sets of plain SU(2) symmetries only, the second setting already includes the larger channel SU(3) symmetry. Both of these symmetries do not capture the full symmetry of the model, which finally is achieved by using the enveloping Sp(6) symmetry.

Due to the internal two-dimensional structure of the SU(3) symmetry based on the fact that SU(3) has two commuting generators, *i.e.* is of rank 2, its multiplets have significantly larger internal dimension, in practice, up to over a hundred. Therefore despite the reduction of the particle-hole symmetry to a plain abelian symmetry, the second setting with the SU(3) channel symmetry allows to outperform the more traditional setup based on SU(2) symmetries only. Similarly, with Sp(6) a rank-3 symmetry, multiplets then easily reach dimensions of several thousands there, which allows to reduce multiplet spaces significantly further still. A detailed analysis of this is provided in this paper, with a more general self-contained introduction to non-abelian symmetries considered given in the appendix [*cf.* App. C 3].

From an NRG point of view,[23] a few essential steps are required. These are (i) the evaluation of relevant operator matrix elements required to construct the Hamiltonian, (ii) the generic setup of an iteration, adding one site to the so-called Wilson chain, and finally, for thermodynamical properties (iii) also the treatment of the full thermal density matrix.[23] All of these steps are simple in principle, yet come with the essential challenge to have a flexible transparent framework for the treatment of non-abelian symmetries in practice. In this paper, such a framework is presented in terms of generalized contractions of tensors in the presence of symmetry spaces, introduced as QSpaces below.

The paper is thus organized as follows. Section II describes the MPS implementation of non-abelian symmetries in terms of QSpaces. Section III describes the implications for calculating correlation functions in the presence of irreducible operator sets. Section IV gives a short review of the NRG together with specialties related to non-abelian symmetries, such as calculating reduced density matrices. This section also introduces the model Hamiltonian of a fully symmetric 3-channel Anderson model. Section V then presents explicit NRG results, followed by summary and outlook. Finally, also an extended Appendix has been added to the paper. The latter is intended to provide a more general pedagogical self-contained introduction to non-abelian symmetries as they occur in fermionic lattice models, together with their actual implementation in practice in terms of QSpaces.

II. MPS IMPLEMENTATION OF NON-ABELIAN SYMMETRIES

Consider some Hamiltonian $\hat{H}$ that is invariant under a set of n_S symmetries,

$$\mathcal{S} \equiv \bigotimes_{\lambda=1}^{n_\text{S}} \mathcal{S}^\lambda, \tag{1}$$

that is, $[\hat{H}, \hat{S}_\alpha^\lambda] = 0$, where α identifies the generator $\hat{S}_\alpha^\lambda$ for the simple (non-abelian) symmetry $\mathcal{S}^\lambda$. To be specific, for example, $\mathcal{S} = \text{SU}(2)_\text{spin} \otimes \text{SU}(2)_\text{charge} \equiv \mathcal{S}^1 \otimes \mathcal{S}^2$ with $\lambda \in \{1, 2\}$ would stand for the combination of spin and charge SU(2) symmetry, respectively. The tensor-product notation in Eq. (1) indicates that the symmetries act independently of each other, that is $[\hat{S}_\alpha^\lambda, \hat{S}_{\alpha'}^{\lambda'}] = 0$ for $\lambda \neq \lambda'$.

Given the symmetries as in Eq. (1), this allows to organize the complete basis of eigenstates of $\hat{H}$ in terms of the symmetry eigenbasis. Every state then belongs to a well-defined irreducible multiplet q^λ for each symmetry $\mathcal{S}^\lambda$. The multiplet itself has an internal state space structure that is described by the additional quantum labels q_z^λ. For example, in the case of $\mathcal{S}^\lambda = \text{SU}(2)$, q^λ (q_z^λ) corresponds to the spin multiplet S (the S_z label), respectively.

Thus all states in a given vector space can be categorized using the hierarchical label structure

$$|qn; q_z\rangle \quad \text{(state-space label structure)}, \tag{2}$$

where

(i) $q \equiv (q^1, q^2, \ldots, q^{n_\text{S}})$, to be referred to as *q-labels* (quantum labels), references the irreducible representations (IREPs) for each symmetry $\mathcal{S}^\lambda$, $\lambda = 1, \ldots, n_\text{S}$. All states in given Hilbert space with the same q-labels are *blocked* together, to be referred to as *symmetry block q*.

(ii) Given a symmetry block q then, the multiplet index $n_{(q)}$ identifies a specific multiplet within this space. It is therefore a plain index associated with given symmetry space q. Together with the q-labels, this forms the *multiplet level* which is considered the topmost conceptual level. Using the composite notation (qn) to identify an arbitrary multiplet, the subscript q to the multiplet index $n_{(q)}$ is considered implicit and hence is dropped, for simplicity.

(iii) Finally, the set of labels $q_z \equiv (q_z^1, q_z^2, \ldots, q_z^{n_\text{S}})$, to be referred to as *z-labels*, resolves the internal structure of each multiplet in q. That is, for each IREP q^λ, referring to the symmetry $\mathcal{S}^\lambda$ in q, q_z^λ labels its internal IREP space. As such, the z-labels are entirely defined by the symmetries considered. By construction, the eigenstates of the Hamiltonian $\hat{H}$ are fully degenerate in the z-labels.

Here the symmetry labels q and q_z describe the combined record of labels derived from all symmetries considered. In practice, states can mostly be treated on the higher multiplet level, while the lower level in terms of the z-labels is split off and taken care of by Clebsch-Gordan algebra and the coefficient spaces derived from it.

When non-abelian symmetries are broken, they are often reduced to their abelian subalgebra. This can be easily implemented, nevertheless, consistent with the presented framework. In particular, in the abelian case, the non-abelian multiplet labels q are absent, while the abelian q_z quantum numbers remain. Therefore the q_z labels can be promoted to the status of q-labels, $q := q_z$. As a consequence, the concept of the actual q_z labels becomes irrelevant (therefore subsequently, the q_z label space may simply be set to zero, $q_z := 0$). The corresponding Clebsch Gordan coefficients are all trivial scalars, *i.e.* equal to 1. Yet these "Clebsch Gordan coefficients for abelian symmetries" do maintain an important role, in that they take care of the proper addition rules that come with abelian symmetries, resulting in $\langle q_1; q_2 | q \rangle = 1 \cdot \delta_{q,q_1+q_2}$.

Given the MPS background of NRG or DMRG, states spaces are generated iteratively, in terms of a product-space of a given effective state space with a newly added local site. Operators, on the other hand, are typically represented in local state spaces, and starting from there, they can be written in terms of matrix elements in the effective global state spaces. With this in mind, the implementation of non-abelian symmetries within the MPS framework therefore is based on the following two basic observations with respect to state space and operator representations, respectively.

(1) State spaces: consider two distinct state spaces, $|Qn; Q_z\rangle$ and $|ql; q_z\rangle$ that, for example, represent a large effective state space and a small new local state space, respectively. Assuming that both state spaces all well-categorized in terms of IREPs, then their tensor-product space can also be decomposed into a direct sum of new combined IREPs $|\tilde{Q}\tilde{n}; \tilde{Q}_z\rangle$ using Clebsch-Gordan coefficients (CGCs),

$$|\tilde{Q}\tilde{n}; \tilde{Q}_z\rangle = \sum_{Qn;Q_z} \sum_{ql;q_z} \left(A_{Q\tilde{Q}}^{[q]}\right)_{n\tilde{n}}^{[l]} \cdot C_{Q_z\tilde{Q}_z}^{[q_z]} \times |Qn; Q_z\rangle |ql; q_z\rangle. \quad (3)$$

Note that the Clebsch Gordan coefficients given by $C_{Q_z\tilde{Q}_z}^{[q_z]} \equiv \langle QQ_z; qq_z | \tilde{Q}\tilde{Q}_z\rangle$ (i) *fully define* the internal multiplet space as specified by the Lie algebra, and (ii) determine the splitting, *i.e.* which output multiplets $\tilde{Q}$ occur for given multiplets Q and q. On the multiplet level, on the other hand, where $(A_{Q\tilde{Q}}^{[q]})_{n\tilde{n}}^{[l]}$ combines the multiplets Qn and ql into the multiplet $\tilde{Q}\tilde{n}$ consistent with the splitting provided by the CGCs, the coefficients $(A_{Q\tilde{Q}}^{[q]})_{n\tilde{n}}^{[l]}$ may encode an arbitrary unitary transformation within

the $\tilde{n}$ output space for each $\tilde{Q}$. The *r.h.s.* of Eq. (3) demonstrates, that the CGC spaces clearly factorize from the multiplet space $A_{Q\tilde{Q}}^{[q]}$ as a tensor product.

(2) Operators: the matrix elements of a specific irreducible operator set (IROP) $\hat{F}^q$, *i.e.* an IROP that transforms according to multiplets q for given symmetries [*cf.* App. Eq. (A3b), or also Sec. A 7] within some symmetry space $|Qn; Q_z\rangle$ can be written using the *Wigner-Eckart theorem* as

$$\langle Q'n'; Q'_z | \hat{F}_{q_z}^q | Qn; Q_z\rangle = \left(F_{QQ'}^{[q]}\right)_{nn'}^{[1]} \cdot C_{Q_z Q'_z}^{[q_z]}, \quad (4)$$

with $C_{Q_z Q'_z}^{[q_z]}$ again the Clebsch-Gordan coefficients as in Eq. (3). On the multiplet level, the *reduced matrix elements* $(F_{QQ'}^{[q]})_{nn'}^{[1]} \equiv \langle Q'n' \| \hat{F}^q \| Qn\rangle$ refer to the *single* irreducible operator set labeled by q, which is indicated by the superscript [1]. The Wigner-Eckart theorem thus allows to *compactify* the operator matrix elements on the *l.h.s.* of Eq. (4) as the tensor-product of reduced matrix elements and CGCs, as shown on the *r.h.s.* of Eq. (4).

P7.

Therefore in both cases above, i.e. *in all tensor objects relevant for a numerical calculation*, the CGC spaces factorize. This allows to *strongly* compress their size, and thus to drastically improve on overall numerical performance. Moreover, note that in both cases, Eq. (3) as well as Eq. (4) the underlying structure comprises tensors of rank-3 throughout. This rank-3 structure holds for both, the reduced multiplet space as well as the CGC spaces. Therefore, in either case, the final data structure of either state space decomposition as well as reduced operator sets is *exactly* the same. It is implemented, in practice, in terms of what will be referred to as QSpace for general tensors of arbitrary rank.

A. General quantum space representation (QSpaces)

The generic representation, used in practice to describe all symmetry related tensors B, is given by a listing of the following type,

$$B \equiv \left\{ \begin{array}{c|c|c} \text{q-labels} & \text{reduced space } \|B\| & \text{CGC spaces} \\ \hline \{Q\}_1 & B_1 & \{C\}_1 \\ \{Q\}_2 & B_2 & \{C\}_2 \\ \cdots & \cdots & \cdots \end{array} \right\}. \quad (5)$$

By notational convention, an actual operator $\hat{B}$ will be written with a hat, while its representation in terms of matrix elements in a specific basis will be written without the hat, hence the corresponding QSpace is referred to as QSpace B. Many explicit examples of QSpaces are introduced and discussed in detail in the appendix

[Sec. C]. As an up-front illustration, consider, for example, the general Hamiltonian of a single spinful fermionic site in the presence of SU(2) symmetry in the spin (S) and charge sector (C), which can be written as the QSpace [see Eq. (C8)]

$$
H \equiv \left\{
\begin{array}{cc|c|cc}
(S;C) & (S';C') & \|H\| & \multicolumn{2}{c}{\text{CGC spaces}} \\
\hline
\frac{1}{2};0 & \frac{1}{2};0 & h_{\frac{1}{2},0} & \mathbf{1}^{(2)} & 1. \\
0;\frac{1}{2} & 0;\frac{1}{2} & h_{0,\frac{1}{2}} & 1. & \mathbf{1}^{(2)}
\end{array}
\right\}. \quad (6)
$$

With every non-zero block listed as an individual row, one can see that the only two reduced matrix elements $\|H\|$ free to choose without compromising the $\mathrm{SU}(2)_{\mathrm{SC}}^{\otimes 2} \equiv \mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)_{\mathrm{charge}}$ symmetry are the parameters (numbers) $h_{1/2,0}$ and $h_{0,1/2}$. By definition, the Hamiltonian is a scalar operator, therefore it is the only operator within its IROP, hence can be written as plain rank-2 QSpace (the third dimension for this IROP would be a singleton dimension, hence can be dropped). Being a scalar operator, the Hamiltonian is block diagonal, which is reflected in equal symmetry sectors $(C;S)$ and $(C';S')$ in each row for first and second dimension, respectively. Moreover, in given case, the corresponding Clebsch-Gordan coefficient (CGC) spaces also result in trivial identities, with $\mathbf{1}^{(2)}$ the two-dimensional identity. Note that the full set of CGC spaces in each row needs to be interpreted as appearing in a tensor product with the multiplet space, here the reduced matrix elements $h_{1/2,0}$ or $h_{0,1/2}$ [e.g. see Eq. (8) below].

In general, the representation of a tensor B of arbitrary rank-r in the QSpace in Eq. (5) [with Eq. (6) an example for a rank-2 QSpace], only lists the non-zero, i.e. relevant symmetry combinations. Having r tensor dimensions, each of its r indices refers to its specific state space $|qn;q_z\rangle_i \equiv |(q)_i n_i;(q_z)_i\rangle$ with $i = 1, \ldots, r$, and hence carries its own label structure as in Eq. (2). The q-labels $(q)_i \equiv \{q^\lambda\}_i$ already represent the *combined* set of n_{S} IREP labels from all symmetries $\mathcal{S}^\lambda$ for the state space at tensor dimension i. In general, by convention, the internal order of the q-labels $(q)_i$ *w.r.t.* λ is fixed and follows the order of symmetries used in Eq. (1).

For a certain row ν of the QSpace listing in Eq. (5) then, the set of r q-labels are grouped into

$$
\{Q\}_\nu \equiv \{(q)_1, \ldots, (q)_r\}_\nu. \quad (7a)
$$

The reduced matrix elements are stored in the dense rank-r tensor B_ν indexed by n_i with $i = 1, \ldots, r$. This is a plain tensor, with the multiplet spaces possibly already rotated by arbitrary unitary transformations and truncated. This is also reflected in the fact that the indices n_i are plain indices, *i.e.* carry no further internal structure. Finally, for every one of the $\lambda = 1 \ldots n_{\mathrm{S}}$ symmetries included, the corresponding CGC space is stored in the sparse tensors $\mathrm{C}_{\lambda,\nu}$, *each of which is also of rank r*. These CGC spaces are grouped into $\{\mathrm{C}\}_\nu$ in the last column,

$$
\{\mathrm{C}\}_\nu \equiv \{\mathrm{C}_{1;\nu}, \ldots, \mathrm{C}_{n_{\mathrm{S}};\nu}\}. \quad (7b)
$$

As the q-labels $\{Q\}_\nu$ also define the z-labels, there is no explicit need to store the z-labels $(q_z)_{i,\nu}$. The internal running indices in $\mathrm{C}_{\lambda,\nu}$, however, are uniquely associated with the z-labels. Note also the different index setting: in contrast to Eq. (7a), which contains a set of r q-labels, *i.e.* one for every dimension of the rank-r tensor B, Eq. (7b) contains a set of n_{S} rank-r CGC spaces, *i.e.* one for every symmetry.

In addition to the QSpace listing in Eq. (5), also the type and order of symmetries considered is stored with a QSpace, *cf.* Eq. (1), even though this is usually the same throughout an entire calculation. Moreover, note that the row or record index ν in Eq. (5) is purely for convenience without any specific meaning, as the order of records in a QSpace can be chosen arbitrarily. Nevertheless, it is required to refer to a specific entry in a QSpace.

For a given record ν in the QSpace in Eq. (5) then, the reduced space and the CGC spaces are to be interpreted as an overall tensor-product,

$$
B_\nu \otimes \{\mathrm{C}\}_\nu \equiv B_\nu \otimes \left(\bigotimes_{s=1}^{n_{\mathrm{S}}} \mathrm{C}_{s;\nu}\right), \quad (8)
$$

while, of course, this is never explicitly done in practice. Yet Eq. (8) demonstrates the single most important motivation to implement non-abelian symmetries in a numerical computation. By splitting off the CGC spaces in terms of a tensor product, block dimensions can be *strongly* reduced for larger calculations with several symmetries present. For the models analyzed in this paper, for example, this was typically an average dimensional reduction from plain abelian symmetries by a factor of 10 up to several hundreds. Considering that both NRG and DMRG scale like $\mathcal{O}(D^3)$ with D the typical dimension of data blocks, this is an enormous gain in efficiency. The factorized CGC spaces, on the other hand, can be dealt with independently, as will be explained in detail later. Assuming that usually the dimensions of the reduced states spaces B_ν still exceed by far the typical dimensions encountered for the CGC spaces, the latter bear little numerical overhead. Only for larger-rank symmetries, such as the symmetry Sp(6) discussed later, multiplet dimensions can become large themselves such that one needs to pay more attention to an efficient treatment of their corresponding CGC spaces [see App. C 3 b].

For QSpaces where the CGC spaces in Eq. (5) exactly correspond to the standard Clebsch-Gordan coefficients for each symmetry, one may argue that actually similar to the z-labels, it is not explicitly necessary to store the CGC spaces altogether, since these are known. This is true, indeed, for these particular cases, and CGC spaces may simply be referenced then. Nevertheless, the explicit storage of the CGC spaces with a QSpace as in Eq. (5) has practical value. When combining QSpaces through contractions, *i.e.* sum over shared indices, for example, quite frequently *intermediate* objects can arise that do have rank different, in particular also larger than 3 [*e.g.* see

P7.

the intermediate objects indicated by the dashed boxes marked by X in Fig. 4]. These then elude a description in terms of standard rank-3 CGCs. In this case, the actual CGC spaces for intermediate QSpace are important, and even though they do not necessarily resemble the interpretation of the original standard rank-3 spaces of standard Clebsch-Gordan algebra anymore, these spaces will be referred to as CGC spaces nevertheless, owing to their origin.

For specific algorithms such as NRG and DMRG then, one typically ends up with simple scalar operators such as the Hamiltonian or a density matrix, apart from intermediate steps where complex CGC structures can arise. Therefore the *full combined* contractions on the CGC level [*e.g.* see Fig. 4] could be replaced by analytical expressions or sum rules for Clebsch-Gordan coefficients. Then, again, the explicit storage of the CGS spaces with the QSpaces would not be strictly necessary and could be avoided altogether. Note, however, that these aforementioned contractions must be implemented within the code dependent on the specific contraction and relying on explicit analytical expressions which, depending on the symmetry, may or may not be known. For the QSpace as outlined in this paper, on the other hand, no special treatment is required for specific contractions, and no explicit knowledge of possibly symmetry dependent CGC sum rules is required. The QSpace approach solely relies on the correct construction of the standard CGC spaces to start with, with the subsequent sums over CGC spaces performed explicitly numerically and not analytically through *exactly* the same contraction as on the reduced multiplet level, as discussed in more detail later.

Finally, the explicit inclusion of the CGC spaces allows to build in strong consistency checks in the actual numerical implementation. Imagine that the Hamiltonian is built by a sequence of complex contractions. The Hamiltonian eventually must be a scalar operator, *i.e.* it is block diagonal in the symmetries and the CGC spaces reduce to plain identities. This can simply be checked at the end of the calculation, which thus provides a strong check of whether the symmetries have been implemented correctly or not. At the stage of intermediate contraction, however, the CGC spaces guarantee the correct splitting and weight distribution between different emerging symmetry sectors.

B. *A*-tensors and Operators

Consider the prototypical MPS scenario as in Eq. (3) that takes some previously constructed state space $|i\rangle \equiv |Qn; Q_z\rangle$ and adds a new local state space $|\sigma\rangle \equiv |ql; q_z\rangle$, *e.g.* a new physical site. The state spaces are thus combined in a product-space described in terms of the IREPs $|j\rangle \equiv |\tilde{Q}\tilde{n}; \tilde{Q}_z\rangle$. Here the states i, σ, and j are introduced as notational shorthand for better readability. The product space then is spanned by $|i\sigma\rangle \equiv |\sigma\rangle|i\rangle$. The order of states in the latter product emphasizes that state $|\sigma\rangle$ is

typically added *after* and thus onto the existing state $|i\rangle$, which is of particular importance for fermionic systems. In general, the combined states[24]

$$|j\rangle = \sum_{l\sigma} |i\sigma\rangle \underbrace{\langle i\sigma|j\rangle}_{\equiv A_{ij}^{[\sigma]}},\qquad(9)$$

are described in terms of linear superpositions of the product space $|i\sigma\rangle$ given by the coefficients $A_{ij}^{[\sigma]}$, henceforth called *A*-tensor (rank-3) or $A^{[\sigma]}$-matrices (rank-2). Without truncation, $A_{ij}^{[\sigma]}$ denotes a full unitary matrix $U_{(i\sigma),j}$ where the round bracket indicates that the indices i and σ have been *fused*, *i.e.* combined into an effective single index. The presence of symmetry and the proper categorization of state spaces, however, imposes certain constraints on this unitary matrix, as pointed out already with Eq. (3). In particular, the fully determined CGC spaces $C_{Q_z\tilde{Q}_z}^{[q_z]}$ factorize from the *A*-tensor, allowing an arbitrary rotation in the reduced multiplet space $A_{Q\tilde{Q}}^{[q]}$ only. For the specific case then, that the reduced multiplet spaces are identical to partitions of identity matrices with a clear one-to-one correspondence still of input and output multiplets, the corresponding *A*-tensor will be referred to as the *identity A-tensor* [see Fig. 2 later; for explicit examples, see App. Eq. (C4) or Eq. (C6)]. An identity *A*-tensor therefore represents the full state space still without any state space truncation, and is unique up to permutations in the combined output space. Its explicit construction is a convenient starting point, in practice, when merging new local state spaces with existing effective state spaces.

The entire construction of an *A*-tensor can be encoded compactly in terms of a rank-3 QSpace. Both coefficient spaces in Eq. (3), $C_{Q_z\tilde{Q}_z}^{[q_z]}$ as well as $A_{Q\tilde{Q}}^{[q]}$, directly enter the QSpace description in Eq. (5). A schematic pictorial representation of an *A*-tensor is given in Fig. 1. There the states i (j) represent the open composite index to the left (right), respectively, while σ refers to the open composite index at the bottom.

As already argued with Eq. (4), an irreducible operator shares exactly the same underlying CGC structure as an *A*-tensor. Thus also its representation in terms of a QSpace is *completely analogous*. Consider an IROP set $\hat{F}^q \equiv \{\hat{F}_{q_z}^q\}$, which transforms according to IREP q. Here, the composite index $\sigma \equiv (ql; q_z)$, for short, identifies the specific operators in the IROP set. As already indicated by the superscript [1] in Eq. (4), its associated multiplet index l has the trivial range $l = 1$, since, by definition, the IROP represents a single IREP on the operator level. With the states $|i\rangle \equiv |Q'n'; Q'_z\rangle$ and $|j\rangle \equiv |Qn; Q_z\rangle$ now representing the same state space within which the operator acts, with usually many multiplets and different symmetries, the operator representation of the IROP $\hat{F}^q$ in the states $\langle i|$ and $|j\rangle$ is evaluated using the Wigner-Eckart theorem in Eq. (4). Similar to the *A*-tensor earlier, the resulting factorization of the
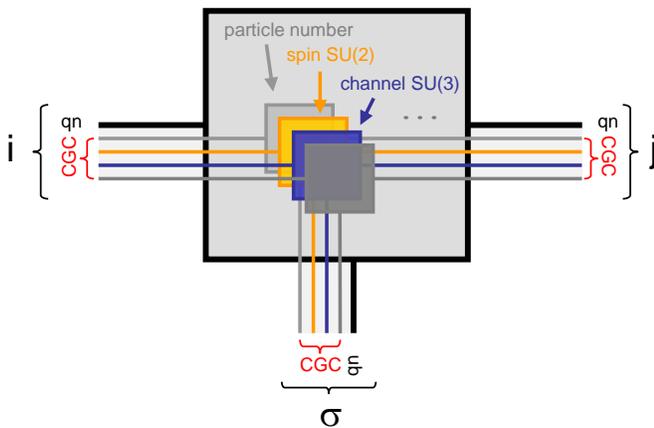
P7.

Figure 1. (Color online) Schematic depiction of a rank-3 QSpace as an example for a basic building block for an MPS or a tensor network, where lines (boxes) represent indices (data spaces), respectively. Every index is assumed to refer to a state space with similar physical background, hence refers to the same global symmetries as in Eq. (1), and has the generic composite structure $|qn; q_z\rangle$ as in Eq. (2), where q_z specifies the states within the CGC spaces. The rank-3 QSpace depicted can be interpreted in two entirely different ways while sharing exactly the same underlying algebraic structure. These are (i) the state space decomposition into IREPs and (ii) operator representation for a given IROP in a given basis (see text). For the general interpretation of the QSpace depicted, consider for simplicity, a single row ν in Eq. (5). The set $\{Q\}_\nu$ defines the q-labels for all tensor dimensions (here a total of three). With the q-labels fixed, the corresponding multiplet index n indexes the typically large reduced multiplet space, indicated by the thick black lines for each tensor dimension. The corresponding reduced rank-3 multiplet space A_ν is depicted by the large gray box in the background. Moreover, with the q-labels fixed, this fixes the IREPs for every tensor dimension and every symmetry. The resulting sparse CGC spaces are indicated by the small boxes around the center, with one box for every symmetry, such as, for example, abelian particle conservation, non-abelian spin SU(2), non-abelian channel SU(3), or other. By construction, all CGC spaces share the same rank as the underlying QSpace. Therefore each CGC space also has three lines attached, one for every tensor dimension. In general, the CGC spaces refer to finite multiplet dimensions for non-abelian symmetries, while for simpler symmetries, such as abelian symmetries, the CGC spaces actually become trivial, *i.e.* scalars. These, nevertheless, are also interpreted as having the same rank as the QSpace using singleton dimensions throughout.

CGC spaces $C_{Q_z Q'_z}^{[q_z]}$ together with the remaining multiplet space $F_{QQ'}^{[q]}$ of reduced matrix elements directly enter the QSpace description in Eq. (5).

So even though an operator is usually considered a rank-2 object, the fact that an IROP consists of an operator *set* indexed by σ, adds a *third* index to the QSpace. In contrast to the state interpretation of σ for the A-tensor above, however, here the "index" σ has a different interpretation in that it points to a specific operator in

the IROP set. By convention, the operator index σ will always be listed as third tensor dimension in its QSpace representation. Given the three-dimensional representation of a general IROP, therefore its entire construction mimics the construction of an A-tensor in terms of a QSpace. As a consequence, Fig. 1 exactly also resembles the QSpace structure of an IROP. The states i (j) used for the calculation of the matrix element represent the open index to the left (right), respectively, while the operator index σ refers to the open index at the bottom.

Scalar operators, finally, such as the Hamiltonian of the system or density matrices, represent a special case, since there the IROP set contains just a single operator. Therefore the third index, *i.e.* the *operator index*, becomes a singleton and hence can simply be dropped [*e.g.* see Eq. (6)]. Scalar operators therefore are represented by rank-2 QSpaces. They are block-diagonal in their symmetries, and their CGC spaces are all equal to identity matrices, with an example already given in Eq. (6).

C. Multiplicity

For general non-abelian symmetries, frequently inner and outer multiplicity occur.[25,26] Both are absent in SU(2), yet do occur on a regular basis for SU($N \geq 3$). Inner multiplicity describes the situation where for a given IREP, several states may share exactly the *same* z-labels. Let m_z^q denote the number of times a specific z-label occurs within IREP q. Then the presence of inner multiplicity implies $m_z^q > 1$ for at least one z-label. Within such degenerate subspaces an arbitrary rotation is allowed in principle. For global consistency, therefore the CGC spaces must adopt a well-defined internal convention on how to deal with inner multiplicity. This issue, however, is entirely contained within the CGC algebra, which is explored in more detail in the App. A [*e.g.* see discussion following Eq. (A19), and App. B 1]. On the level of a QSpace, it is of no further importance otherwise. Essentially, the only implication of inner multiplicity is $q_z \to (q_z, \alpha_z)$ with $\alpha_z = 1, \ldots, m_z^q$ [*cf.* Eq. (A19)], where m_z^q depends on the multiplet q. With this minor adjustment, it is assumed throughout that the z-labels fully identify the internal multiplet space. Note that, in practice, the extra label α_z is never included explicitly. What is important, however, is a *consistent internal multiplet ordering* that respects multiplicity [see App. B 1].

Outer multiplicity, on the other hand, describes the situation where in the state space decomposition of a product-space of two IREPs, q_1 and q_2, the *same* output IREP q may appear multiple times, the number of which is specified by $M_q^{[q_1,q_2]}$ [*cf.* App. Eqs. (A35-A38) and discussion]. Therefore outer multiplicity primarily also enters at the level of Clebsch-Gordan coefficients, as it is based on pure symmetry considerations. In contrast to inner multiplicity, however, outer multiplicity also affects the reduced multiplet space, as will be elaborated upon in what follows.

In the absence of outer multiplicity [*i.e.* $M_q^{[q_1,q_2]} \leq 1$ for all q_1, q_2, and q of the symmetry, an example being SU($N \leq 2$)], all rows in the QSpace in Eq. (5) must have *unique* $\{Q\}_\nu$. If this is not the case, then the rows can be made unique by combining the rows with the same $\{Q\}$. Assume, for example, $\{Q\}_\nu = \{Q\}_{\nu'}$ with $\nu \neq \nu'$: clearly, the $\{Q\}$'s are already the same. Having the *same* symmetry labels, this refers to the same set of IREPs, hence also the CGC spaces of these records must be identical, up to a possible global normalization factor which can be associated with the multiplet space, instead. Furthermore, given $\{Q\}_\nu = \{Q\}_{\nu'}$, the A_ν and $A_{\nu'}$ data blocks do live in *exactly* the same vector spaces for each individual tensor dimension! Therefore A_ν and $A_{\nu'}$ can be simply added up [here multiple contributions with the same $\{Q\}$ are considered additive, consistent with general conventions regarding sparse tensors; otherwise, say having given the same matrix element twice with different values, would immediately lead to contradictions].

In the presence of outer multiplicity, on the other hand, the uniqueness of the q-labels $\{Q\}_\nu$ in the QSpace in Eq. (5) has to be relaxed. The reason for this is as follows. Since outer multiplicity derives from the Clebsch-Gordan algebra as in Eq. (A38), the CGC spaces

$$C_{Q_z \tilde{Q}_z}^{[q_z]} \rightarrow C_{Q_z \tilde{Q}_z, \alpha}^{[q_z]} \equiv \langle Q Q_z; q q_z | \alpha \tilde{Q}, \tilde{Q}_z \rangle \qquad (10)$$

acquire an additional label $\alpha = 1, \ldots, M_{\tilde{Q}}^{[Q,q]}$ [different from the α_z used with inner multiplicity], where $M_{\tilde{Q}}^{[Q,q]}$ indicates the outer multiplicity in $\tilde{Q}$, given the product space of the IREPs Q and q. In terms of a QSpace object, one may therefore be tempted to enlarge the CGC space from rank-3 to rank-4, with the dimension of the 4th index being equal to $M_{\tilde{Q}}^{[Q,q]}$. This strategy alone, however, does not capture the full picture since outer multiplicity also *enlarges* and thus effects the multiplet space A_ν of an A-tensor. By definition, outer multiplicity means that *different* multiplets with the same q can emerge. The only way they can be distinguished is through their Clebsch-Gordan coefficients. Therefore rather than enlarging the CGC space in a QSpace, $M_{\tilde{Q}}^{[Q,q]}$ records with the *same* $\{Q\}_\nu$ are allowed, instead. These records have CGC spaces of the same rank-3 dimensions, which, however, are clearly distinguishable, as they are orthogonal to each other [*cf.* appendix Eq. (A39)]. The $M_{\tilde{Q}}^{[Q,q]}$ sets of Clebsch-Gordan coefficients arising from outer multiplicity are thus spread over $M_{\tilde{Q}}^{[Q,q]}$ records within a QSpace object.

The situation in the multiplet space for an identity A-tensor is depicted schematically in Fig. 2. In the absence of outer multiplicity, each symmetry combination $(q, q')_i$ can only contribute at most once to a given symmetry space q'' and gets its space allocated, as depicted, for example, for $(q, q')_4$ in Fig. 2(a), having only one non-zero block (shaded block) within the q'' output multiplet. The symmetry combinations $(q, q')_1$ and $(q, q')_2$, on the other
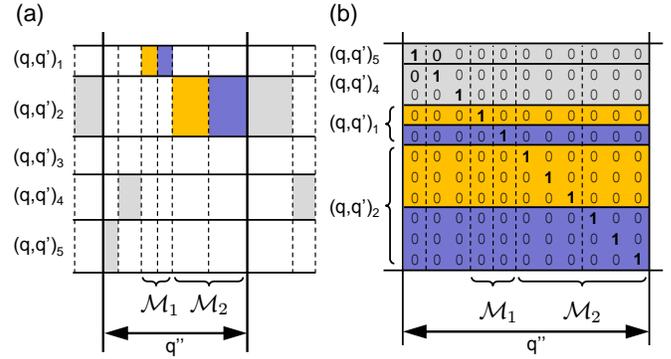

Figure 2. Effect of outer multiplicity on multiplet space (A_ν) in terms of an identity A-tensor– Panel (a) Schematic depiction of the state space decomposition of two input multiplet spaces with unique symmetry combinations (q, q') into combined multiplets q'' (rows and columns, respectively). State spaces of the same symmetry are grouped into blocks separated by solid lines (horizontally and vertically). For simplicity, an identity A-tensor is depicted, for which the individual sectors in q'' can be uniquely associated with the (q, q') they originate from. Hence each column, separated by solid lines, has exactly one shaded block considered non-zero, with all-zero blocks shown in white. Here vertical thin lines indicate sub-blocks that originate from different (q, q'), yet are eventually combined in the same block as they belong to the same symmetry q'' (separated by thick lines). Now, in the presence of outer multiplicity a specific (q, q') can contribute to the same q'' several times, as depicted schematically by the spaces $\mathcal{M}_1$ and $\mathcal{M}_2$ for the rows $(q, q')_1$ and $(q, q')_2$, respectively, both showing a multiplicity of $M_{q''} = 2$. Panel (b) depicts the *enlarged* multiplet space for the output multiplet q'' of panel (a) in order to accommodate the additional multiplets arising from outer multiplicity. Being an identity A-tensor, the entire block shown in panel (b) represents an identity matrix (in contrast to an arbitrary A-tensor, which may have an arbitrary unitary matrix in its place). The vertical lineup of (q, q') sectors is arbitrary, making the identity A-tensor unique up to permutations. The identity matrix shown in the panel is sliced into horizontal blocks as indicated, each of which is associated with its own unique CGC space [not shown] as derived from the Lie algebra of the symmetry under consideration. Each of these slices then directly enters as a reduced multiplet space A_ν in a separate row in the QSpace as in Eq. (5).

hand, show outer multiplicity, in that they result twice in the same multiplet q'', *i.e.* $M_{q''}^{[q,q']_1} = M_{q''}^{[q,q']_2} = 2$.

For simplicity, in the absence of truncation and without any further unitary rotation, the tensor-product on the multiplet level can be represented as an identity A-tensor with a clear one-to-one correspondence of input to output multiplets. This is depicted in Fig. 2(b) in terms of an identity matrix in the reduced multiplet space. The identity matrix in panel (b) then is sliced horizontally into blocks for each (q, q') that contributes to q''. In the presence of outer multiplicity, the state space for q'' needs to be *enlarged* to accommodate the additional multiplets.

The slicing (horizontal solid lines) then also proceeds for every output multiplet resulting from outer multiplicity, as indicated in panel (b). As a result, $M_{q''}^{[q,q']}$ slices are associated with exactly the same $Q \equiv \{q, q', q''\}$, distinguishable only through their Clebsch-Gordan coefficients. These slices directly enter as A_ν in separate rows in a QSpace as in Eq. (5).

In summary, outer multiplicity requires an adaptation of the multiplet space, which is naturally incorporated into a QSpace by allowing multiplet entries with the same $\{Q\}_\nu$ labels yet with clearly distinguishable CGC spaces. That is, specific records are also considered to refer to *different* state spaces if their CGC spaces are not exact copies (up to a global factor that can be incorporated into the multiplet data) but rather orthogonal to each other [see App. Eq. (A39)]. In practice, this is checked within a small numerical threshold ($\sim 10^{-12}$) accounting for numerical double precision noise. The great advantage of this prescription is that then multiplicities fall completely in line with the rest of the MPS algorithm without any specific further treatment.

Finally, it is important to notice that the same concept of relaxing the uniqueness of the $\{Q\}_\nu$ labels actually also can become relevant for symmetries that *do not* have intrinsic outer multiplicity in its actual sense. Yet, in fact, through contractions *intermediate objects* can arise of rank larger than three [*e.g.* see the QSpaces indicated by the dashed boxes marked by X in Fig. 4], where records in a QSpace with the same $\{Q\}_\nu$ labels can also have *incompatible* CGC spaces, in the sense that they are *not* the same up to overall factors. In this case, also the uniqueness of the $\{Q\}_\nu$ must be relaxed temporarily. For simplicity, this will also be referred to as outer multiplicity.

D. Contractions

The contraction of QSpaces will be introduced in the following in terms of a simple example, namely the orthonormalization condition on the combined state space in a tensor-product space. Putting symmetry labels aside for the sake of the argument, the A-tensor $A_{ij}^{[\sigma]} \equiv \langle i\sigma | j \rangle$ in Eq. (9) combines the state spaces $|i\sigma\rangle$ into a combined (possibly truncated) orthonormal state space $|j\rangle$. This directly leads to the standard orthogonality relation for an A-tensor,

$$\sum_{i\sigma} A_{ij}^{[\sigma]*} A_{ij'}^{[\sigma]} = \delta_{jj'}, \qquad (11)$$

which is a simple example for the simultaneous contraction of two tensors *w.r.t.* to two common indices, here i and σ. By construction, it is completely analogous in structure to the orthogonality condition of CGCs as in App. Sec. A39. Including symmetries, the contraction in Eq. (11) is depicted in terms of QSpaces in Fig. 3. Overall, indices are represented by lines, and lines connecting

two blocks such as the indices i and σ are summed over, *i.e.* contracted. In practice, contraction of QSpaces as defined in Eq. (5) happens at several levels, since state indices are always composite indices that refer to a symmetry basis of the type $|qn; q_z\rangle$. This implies for a contraction $\sum_{i=i'}$ of two QSpace objects with respect to some common state space i and i', that (i) the q-labels q_i and $q_{i'}$ of the QSpaces as in Eq. (5) must be matched for the indices i and i', respectively. For a given specific match of rows ν and ν' then, (ii) this is followed by the contraction of the corresponding reduced multiplet spaces, and (iii) by exactly the same contraction of the CGC spaces, one for each symmetry. Note that the rank of a QSpace and its index order are always shared by the multiplet space and CGC spaces for consistency. Hence the overall contraction of the QSpaces is directly reflected in the elementary contraction of the plain numerical tensors A_ν and $\{C\}_\nu$. That is, the *contraction pattern* depicted schematically in Fig. 3, drawn in terms of boxes with connecting lines, is exactly the same on all levels of the contraction. By collecting the remaining *non*-contracted q-labels, this forms a new entry ν'' in the resulting QSpace, with the (tensor) index order of the resulting tensor dimensions being the same for all $\{Q\}_{\nu''}$, $A_{\nu''}$, and $\{C\}_{\nu''}$ for consistency.

Finally, the resulting QSpace is made unique in the $\{Q\}_{\nu''}$ labels as far as outer multiplicity permits. Records can only be combined, *i.e.* summed over, *iff* the CGC spaces for given records are all the same up to global factors which can be absorbed into the multiplet data, instead (see Sec. II C). Outer multiplicity plays no special role with contractions otherwise. Note that independent of whether or not outer multiplicity is present, when specifying a *subset* of tensor dimensions within $\{Q\}_\nu$ for contraction, the resulting QSpace will, in general, always have *many* contributions to the same $\{Q\}_{\nu''}$. For comparison, consider the completely analogous case of regular square matrices of dimension $D > 1$: a matrix element $(M)_{ij}$ is uniquely identified in the overall index (i, j), while for example, the index i is not sufficient as it refers to an entire row of matrix elements. Moreover, when two matrices M_1 and M_2 are multiplied together,

$$(B)_{jj'} = (M_1 M_2)_{jj'} = \sum_{i=i'} (M_1)_{ji} (M_2)_{i'j'}, \qquad (12)$$

the common index space (second index of M_1 and first index of M_2) is summed over, *i.e.* contracted. *Every* match $i = i'$ results in a contribution. In particular, for some given j and j', all D matches $i = i'$ contribute and are summed up to the same output space (j, j'). In the case of QSpaces the situation is *exactly* analogous. All matches $i = i'$ in the q-labels q_i and $q_{i'}$ for the contracted index must be included. The only real consequence of outer multiplicity is that in the resulting QSpace B in Eq. (12) not necessarily all records with the same $\{Q\}_\nu$ labels can be merged by adding them together. In the specific case of the contraction in Eq. (11), however, the resulting QSpace is simply the identity, and as such a
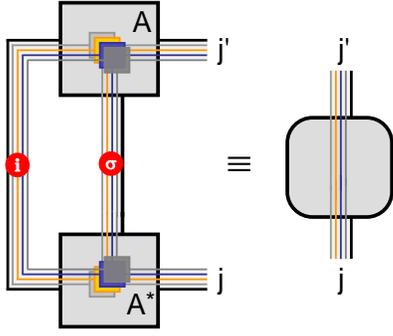
Figure 3. (Color online) Contraction of (i) an A-tensor or (ii) an irreducible operator into a scalar. All indices specified are composite indices of the type $|qn; q_z\rangle$. An A-tensor describes a (truncated) basis transformation of the product-space of the new local space $|\sigma\rangle$ with an effective previously constructed basis $|i\rangle$, resulting in the combined state space $|j'\rangle \equiv \sum_{i'\sigma'} A_{i'j}^{[\sigma']}|\sigma'\rangle|i'\rangle$, with the corresponding bra-space $\langle j| \equiv \sum_{i\sigma} A_{ij}^{[\sigma]*}\langle i|\langle \sigma|$ depicted in the lower part of the figure. The result is the scalar identity operator, reflecting the orthonormality condition Eq. (11). An entirely different interpretation of the same contraction pattern can be given when the A-tensor is replaced by an IROP F^σ. The contraction then describes Eq. (13b) and yields a scalar operator, with its generic QSpace representation schematically depicted to the right.

scalar operator with unique $\{Q\}_\nu$.

E. Scalar operators

Given the definition of an A-tensor in Eq. (9), the contraction of the two QSpaces A and A^* in Fig. 3 leads to the identity operator $\hat{\mathbf{1}}^{(C)} \equiv \sum_j |j\rangle\langle j|$ in the possibly truncated combined space C [cf. Eq. (11)]. Clearly, this also provides a strong check on the numerical implementation of the symmetries. In particular, $\hat{\mathbf{1}}^{(C)}$ represents a (trivial) example of a scalar operator, that can be described as rank-2 QSpace. The CGC spaces are all identity matrices (up to overall factors that can be associated with the multiplet space), and therefore the lines, that usually connect to the CGC spaces within a QSpace, can be directly connected through from j to j' on the r.h.s. of Fig. 3, with the CGC spaces themselves no longer shown. In given case, due the orthonormality condition in Eq. (11), also the reduced multiplet space is given by identity matrices. This actually also would allow to connect through the thick black line on the r.h.s. of Fig. 3, and thus also to skip the large remaining block on the r.h.s. for the reduced multiplet space altogether.

Figure 3, however, allows yet an entirely different interpretation. Remember that an irreducible operator set $\hat{F}^q$ has a completely analogous structure and interpretation in terms of its internal CGC spaces when compared to an A-tensor (cf. Fig. 1). Therefore it must hold that

the scalar-product-like contraction,

$$\hat{F}^2 \equiv \hat{F} \cdot \hat{F}^\dagger \equiv \sum_{q_z} \hat{F}_{q_z}^q (\hat{F}_{q_z}^q)^\dagger \tag{13a}$$

also results in a scalar operator (note that through the Wigner-Eckart theorem, by convention, the state space associated with the *right* index of the operator $\hat{F}^q$ is combined with the multiplet space q; *cf.* App. A 7). With $\sigma \equiv (q1; q_z)$ and the further sum through the operator (matrix) multiplication, Eq. (13a) shares exactly the same contraction pattern as discussed in Fig. 3 in the context of the orthonormality of A-tensors earlier. Here the resulting scalar operator, however, can have arbitrary positive hermitian matrices in its multiplet space still, represented by the large gray box on the r.h.s. of Fig. 3. The reduction of Eq. (13a) to a scalar operator is also intuitively clear, given that the Hamiltonian itself is typically constructed in terms of scalar operators of exactly this type [see, for example, App. Eq. (A48) or Eq. (A60) given the Hamiltonian in Eq. (A46)]. The notation in Eq. (13a) emphasizes that in the scalar product the *same* irreducible operator set $\hat{F}^q$ must be taken, considering that the IROP $\hat{F}^q$ is *different* from the IROP $(\hat{F}^\dagger)^q$. Nevertheless, since $(\hat{F}_{q_z}^q)^\dagger \sim (\hat{F}^\dagger)_{-q_z}^q$, up to possible signs originating from the definition of the CGC algebra [*e.g.* compare the QSpaces in App. Tbls. C3a and C3b and accompanying discussion], these signs are irrelevant in the scalar contraction. Hence it follows that also

$$\tilde{F}^2 \equiv \hat{F}^\dagger \cdot \hat{F} \equiv \sum_{q_z} (\hat{F}_{q_z}^q)^\dagger \hat{F}_{q_z}^q \tag{13b}$$

is a scalar operator, yet different from Eq. (13a), as indicated by the tilde on $\tilde{F}^2$. Similarly, note that if the A-tensor had been contracted on the right instead of the left index in Fig. 3, this also would have yielded a scalar operator, namely a reduced density matrix up to normalization (*e.g.* Fig. 5 below using $\rho_k \equiv \mathbf{1}$).

F. Operator matrix elements

The typical calculation of matrix elements of operators for iterative methods such as NRG or DMRG is depicted schematically in Fig. 4. While the complex many body states are generated iteratively and described by A-tensors [*cf.* Eq. (9)], an elementary irreducible operator set $\hat{f}^q$, on the other hand, usually operates locally within the state space $|\sigma\rangle$ of a specific site. Therefore, the operator is described initially in terms of the matrix elements $f_{\sigma\sigma'}^{[q]} \equiv \langle \sigma|\hat{f}^q|\sigma'\rangle$. The link to the many body states is given through the A-tensor that connects given site to a generated effective state space $|i\rangle$, $|j\rangle = \sum_{i\sigma} A_{ij}^{[\sigma]}|\sigma\rangle|i\rangle$. The matrix elements of an IROP in the combined space
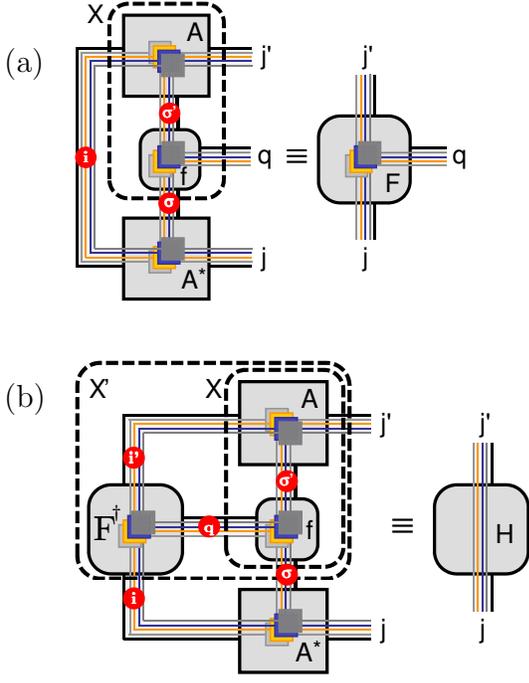
(a)

(b)

P7.

Figure 4. (Color online) Typical evaluation of matrix elements given an A-tensor. The nested dashed boxes $X^{(\prime)}$ indicate the sequential order of contractions prior to the final contraction. In panel (a), the local IROP set $\hat{f}^q$ acts within the state space $|\sigma\rangle$ of a given site. Its local matrix elements, $\langle\sigma|\hat{f}^q|\sigma'\rangle$, are assumed to be known and described in terms of the local rank-3 QSpace f. The local IROP set is mapped into the larger effective space linked through the A-tensor, $|j\rangle = \sum_{i\sigma} A_{ij}^{[\sigma]}|\sigma\rangle|i\rangle$. The overall result is the rank-3 QSpace F on the *r.h.s.*, *i.e.* the desired matrix elements $F_{jj'}^{[q]} \equiv \langle j|\hat{f}^q|j'\rangle$. Panel (b) depicts a typical scalar nearest-neighbor contribution to a Hamiltonian $\hat{H} \equiv [\hat{f}^q]_k^\dagger \cdot [\hat{f}^q]_{k+1}$ of two consecutive sites, say k and $k+1$ using their respective A-tensors. This contraction already uses an effective description of the local operator $\hat{f}_k^\dagger$ at site k in terms of the QSpace $F^\dagger$, obtained from the A-tensor $A_{(k)}$ at site k as in panel (a) form the prior iteration. Using the A-tensor $A_{(k+1)}$ of site $k+1$, the overall contraction can be completed as indicated.

$|j\rangle$ then become,

$$F_{jj'}^{[q]} \equiv \langle j|\hat{f}^q|j'\rangle = \sum_{i\sigma,i'\sigma'} A_{ij}^{[\sigma]*} A_{i'j'}^{[\sigma']} \underbrace{\langle i|\langle\sigma|\hat{f}^q|\sigma'\rangle|i'\rangle}_{\equiv \delta_{ii'} f_{\sigma\sigma'}^{[q]}}$$

$$= \left[\sum_\sigma A^{[\sigma]\dagger}\left(\sum_{\sigma'} f_{\sigma\sigma'}^{[q]} A^{[\sigma']}\right)\right]_{jj'}, \qquad (14)$$

It is exactly this procedure that is depicted in Fig. 4(a). The matrix elements are calculated in a two-stage process. The sum in the round brackets of Eq. (14) (contraction of σ') is carried out first, leading to the temporary rank-4 tensor with open indices (i, j', σ, q) [box X in Fig. 4(a)]. This rank-4 tensor then is contracted simultaneously in the indices i and σ with the A^* tensor, provid-
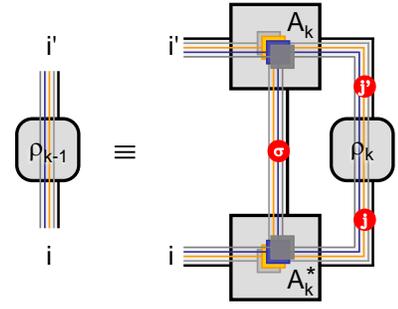
Figure 5. (Color online) Backward update of density matrix ρ_k given in the effective basis $|j\rangle$ of a system up to and including site k (right index) by tracing out the local state space $|\sigma_k\rangle$ (middle index) given the basis transformation A_k that introduced site k. The result is the reduced density matrix ρ_{k-1} in the effective basis $|i\rangle$ of the system up to and including site $k-1$.

ing the final result shown on the *r.h.s.* of Fig. 4(a). Quite generally, for contractions including several blocks as in Fig. 4, these are always done sequentially, adding one block at a time. This is explicitly indicated in Fig. 4 by the (nested) dashed boxes, with the final contraction connecting the remaining tensor to the outer-most dashed box. Every individual contraction then follows the multistage process over composite indices as described earlier in Sec. II D.

The so obtained effective description $F_k^{[q]}$ of an operator $\hat{f}^q$ acting on site k using A_k can be used then to describe, for example, the typical scalar nearest-neighbor contribution $[\hat{f}^q]_k^\dagger \cdot [\hat{f}^q]_{k+1}$ to the Hamiltonian including site $k+1$. This operation is shown in Fig. 4(b). In particular, one may use the identity A-tensor A_{k+1}^{Id} for site $k+1$, such that the resulting Hamiltonian is constructed in the full tensor-product space $|\sigma\rangle_{k+1}|i\rangle_k$ of the system up to and including site $k+1$. Here $|i\rangle_k$ describes the effective space up to and including site k, whereas $|\sigma\rangle_{k+1}$ describes the new local state space of site $k+1$. This exactly corresponds the two-stage prescription used within the NRG (and similarly also for the DMRG) to build the Hamiltonian for the next iteration: (i) the tensor-product space including the newly added site must be mapped into proper symmetry spaces. This is taken care of by the construction of the identity A-tensor A_{k+1}^{Id}. (ii) The new Hamiltonian is built using this identity A-tensor through contractions as shown in Fig. 4(b) [note that while the presence of outer multiplicity in QSpace f is typically inherited by QSpace F through the basis transformation as in Fig. 4(a), the internal contraction over the IROP set index q in Fig. 4(b) eventually leads to a scalar contribution to the Hamiltonian, as discussed with Eq. (13b)]. After diagonalization and state space truncation in the combined state space, the part of the resulting unitary matrix describing the kept states can be contracted onto A_{k+1}^{Id}, yielding the actual final A_{k+1}.

G. Density matrix and backward update

Consider the density matrix $\hat{\rho}_k \equiv \sum_{jj'} (\rho_k)_{jj'} |j\rangle\langle j'|$ given in the basis $|j\rangle_{(k)}$, which is assumed to include all sites of a system up to and including site k. With the local state space of the last site k described by $|\sigma_k\rangle$, tracing out this last site from the density matrix ρ_k corresponds to contracting the A_k-tensor that connected site k to the system's MPS,

$$
\begin{aligned}
\hat{\rho}_{k-1} &= \operatorname*{tr}_{\sigma_k}\left(\hat{\rho}_k\right) \\
&= \sum_{ij,i'j',\sigma} A_{ij}^{[\sigma_k]*} A_{i'j'}^{[\sigma_k]} (\rho_k)_{jj'} |i\rangle\langle i'| \\
&\equiv \sum_{ii'} \left(A^{[\sigma_k]\dagger} \rho_k A^{[\sigma_k]}\right)_{ii'} |i\rangle\langle i'|.
\end{aligned}
\tag{15}
$$

Equation (15) leads to a density matrix $\hat{\rho}_{k-1}$, which now is written in the many-body basis $|i\rangle_{(k-1)}$ which includes all sites up to and including site $k-1$. This *backward* update is a well-known operation within the NRG.[17,23,27,28] Its graphical depiction is given in Fig. 5 [note that the sum over i and i' in Eq. (15) connects to state spaces that are not yet contracted; hence these correspond to open indices in Fig. 5].

The backward update of the density matrix in Eq. (15) preserves its properties as a density matrix and as a scalar operator. The former directly follows from the realization that the orthonormality condition Eq. (11) with the A-tensor in the last line of Eq. (15) is exactly equivalent to a complete positive map. Moreover, by tracing out part of a system such as a site that has been added through a tensor product space and that itself can be fully categorized using given symmetries, this procedure cannot break symmetries by itself. This is to say, that the partial trace in Eq. (15) preserves the property of a scalar operator. However, the trace over CGC spaces adds important weight factors to the reduced multiplet spaces, which are crucial, for example, to preserve the overall trace of the density matrix during back-propagation. While the contraction in Fig. 5 can be easily performed, in practice, without the explicit knowledge of these weights, their determination is straightforward and instructive, nevertheless, as will be shown in the following.

The contraction in Fig. 5 clearly also holds for the CGC spaces of every symmetry individually. Therefore it is sufficient to focus on one specific symmetry. Let i contain several multiplets q_i, and consider, for simplicity, the special case where the local state space σ contains one specific multiplet q_σ only. In addition, also the reduced density matrix $\hat{\rho}_k$ is chosen such that it only picks one very specific multiplet q_j. Focusing on the Clebsch Gordan coefficients $C_{q_{iz}q_{jz}}^{[q_{\sigma z}]} \equiv \langle q_i q_{iz}; q_\sigma q_{\sigma z}|q_j q_{jz}\rangle$ for chosen symmetry then, which properly combine the irreducible multiplets q_i and q_σ into the multiplet q_j, the contraction

in Eq. (15) with respect to the fixed q_j is given by

$$
\begin{aligned}
&\sum_{q_{\sigma z} q_{jz} q_{j'z}} \langle q_{i'}q_{i'z}; q_\sigma q_{\sigma z}|q_j q_{j'z}\rangle\langle q_i q_{iz}; q_\sigma q_{\sigma z}|q_j q_{jz}\rangle^* \cdot \delta_{q_{jz}q_{j'z}} \\
&= \sum_{q_{\sigma z}} \langle q_{i'}q_{i'z}; q_\sigma q_{\sigma z}|\left(\sum_{q_{jz}} |q_j q_{jz}\rangle\langle q_j q_{jz}|\right)|q_i q_{iz}; q_\sigma q_{\sigma z}\rangle \\
&= f_{q_i q_j} \cdot \delta_{q_i q_{i'}} \delta_{q_{iz} q_{i'z}},
\end{aligned}
\tag{16}
$$

where the $\delta_{q_{jz}q_{j'z}}$ in the first line comes from the assumption that the initial $\hat{\rho}_k$ is a scalar. The last identity follows from the fact that also $\hat{\rho}_{k-1}$ shall be a scalar operator. Alternatively, the last equality can also be understood as a general intrinsic completeness property of Clebsch-Gordan coefficients. Either way, the remaining factor $f_{q_i q_j}$ in the last line must be independent of the z-labels. The factor $f_{q_i q_j}$ then, in a sense, reflects the weight of how the IREP q_i together with the traced over IREPs q_σ contributes to the final total q_j. If, for example, for fixed q_i and the known set of q_σ some final total q_j cannot be reached, then it holds $f_{q_i q_j} = 0$ for this case.

From the scalar property of $\hat{\rho}_{k-1}$, Eq. (16) can be further constrained to some specific $q_i = q_{i'}$. Also summing over $q_{iz} = q_{i'z}$ then, the second line in Eq. (16) becomes equal to $\operatorname{tr}\left(\sum_{q_{jz}} |q_j q_{jz}\rangle\langle q_j q_{jz}|\right) = d_{q_j}$, *i.e.* the internal multiplet dimension of the IREP q_j. Together with the last line in Eq. (16), it follows,

$$
f_{q_i q_j} = \frac{d_{q_j}}{d_{q_i}}.
\tag{17}
$$

as demonstrated, for example, for SU(2) in [17]. Note that Eq. (17) holds in general for arbitrary symmetries, and also in the presence of outer multiplicity. This follows by recalling that one of the main assumptions that entered Eq. (16) was to pick one specific multiplet q_j. This single IREP, however, may equally well also have been any of the multiplets resulting from outer multiplicity, say multiplet $q_j \to q_{j,\alpha}$, which nevertheless again leads to Eq. (17).

H. Generalized A-tensor for tensor networks

The prototypical A-tensor as defined in Eq. (9) combines two physically distinct state spaces in terms of their tensor-product space. One may be interested, however, in the case where three or more state spaces need to be combined into the description of a single combined state space, while also respecting symmetries. This situation, for example, occurs regularly in the context of tensor network states.[6–9] Let m be the number of states spaces to be combined. Then this requires the generalized Clebsch-Gordan coefficients $\langle q_1 q_{1z}; \ldots; q_m q_{m,z}|q q_z\rangle$. Once known, in principle, they can be combined compactly into a generalized A-tensor of rank $m+1$. The question is, how to obtain this generalized A-tensor in a simple manner in practice.

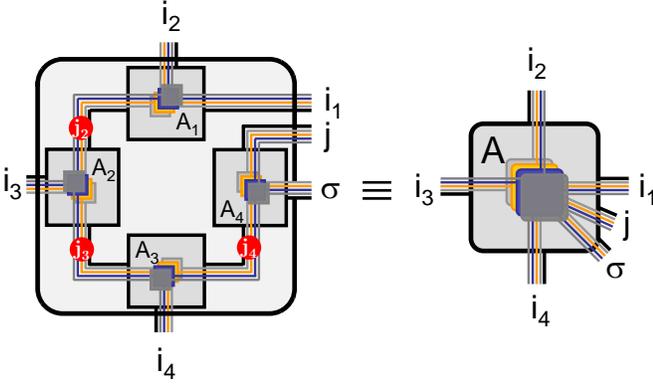For this, the QSpace structure introduced in this paper proves very useful. In particular, a generalized A-tensor

P7.

Figure 6. (Color online) Generalized A-tensor that combines multiple state spaces, *i.e.* four effective state spaces $|i_1\rangle, \ldots, |i_4\rangle$ together with one local degree of freedom $|\sigma\rangle$. Here it is assumed that all input state spaces describe proper orthonormal state spaces that act in different spaces, such that they can be combined into a simple product space. The index j, finally, represents the common global state space. In particular, it can be used to truncate the global Hilbert space to the state space of interest (within the DMRG, this may simply be the ground state, where the index j, being a singleton dimension, simply may be skipped then). While the general Clebsch-Gordan coefficients for the entire object may not be easily available (object to the right), the overall A-tensor can be built iteratively by adding one state space at a time (object to the left), starting, say, from A_1 which links the two state spaces i_1 and i_2 into the combined state space j_2 and hence allows to employ Clebsch-Gordan coefficients in the usual manner. The state space j_2 can then be combined with state spaces i_3, and so forth. Contraction of the intermediate indices $j_2, \ldots, j_4$, finally, leads to the generalized A-tensor to the right.

can be obtained based on the iterative pairwise addition of individual state spaces, which is a well-established procedure at every step. The situation is depicted schematically in Fig. 6. To be specific, Fig. 6 considers four effective state spaces $|i_\alpha\rangle \equiv |q_\alpha q_{\alpha,z}\rangle$ with $\alpha = 1, \ldots, 4$, together with a local state space $|\sigma\rangle \equiv |q_5 q_{5,z}\rangle$, thus having $m = 5$. This specific setting may correspond, for example, to the situation in a tensor network state that describes a two-dimensional system which, from the point of view of a specific site with state space σ, has four effective states spaces to the top, bottom, left, and right, respectively. Note, however, that here the state spaces $|i_\alpha\rangle$ with $\alpha = 1, \ldots, m$ are assumed to be *physically different*, in principle, such that the tensor product space is a well-defined concept.

Starting with state spaces $|i_1\rangle$ and $|i_2\rangle$ in Fig. 6, their state space can be combined in terms of and identity A-tensor A_1^{Id} in the usual fashion using standard Clebsch-Gordon coefficients. The resulting combined state space $|j_2\rangle$ then can be combined with state space i_2 in terms of another identity A-tensor A_2^{Id}, thus yielding $|j_3\rangle$, again using standard Clebsch-Gordan algebra. The procedure

is repeated, until at the last step, for example, the local state space σ is added, resulting in the full combined state space $|j\rangle$ properly categorized in terms of symmetries. The iteratively generated $m - 1$ identity A-tensors A_k^{Id}, on the other hand, can be contracted into a single tensor of rank $m + 1$ by contracting the intermediate indices $j_2, \ldots, j_{m-1}$. This then results in the desired generalized A-tensor, shown at the *r.h.s.* of Fig. 6. Furthermore, in the context of DMRG or tensor network states, one is typically interested in a single state such as the ground state of the system. In this case, the full combined state space $|j\rangle$ is truncated to a single state, such that the extra index $|j\rangle$ becomes a singleton and can be dropped, for simplicity.

The construction of a generalized A-tensors, as outlined above, emphasizes the versatility of the QSpaces introduced in this paper. In particular, by explicitly including the CGC spaces in Eq. (5), generalized Clebsch-Gordan coefficients can be easily obtained in terms of a generalized A-tensor, which itself is constructed through an iterative transparent procedure.

III. CORRELATIONS FUNCTIONS

Correlations functions are usually calculated with respect to operators whose transformation under given symmetries is known. This is specifically so, as these operators often naturally derive from the same fundamental building blocks, that also enter the Hamiltonian. In the presence of non-abelian symmetries, a single specific operator then that is not a scalar operator, is usually part of a larger irreducible operator set. In practice, thus also its correlation function is calculated *w.r.t.* the *full* IROP, for simplicity, as will be explained in the following.

Consider, for example, the retarded Greens function

$$G_\sigma(\omega) \equiv \langle d_\sigma \| d_\sigma^\dagger \rangle_\omega \qquad (18)$$

that, in the time domain, creates a particle of preserved flavor σ, and destroys it some time later. Clearly, a particle with the same flavor must be destroyed later, otherwise the Greens function is zero, *i.e.* the Greens function is diagonal with respect to symmetries. Now in the presence of symmetries, it must be possible to write the operators $d_\sigma^\dagger$ as part of an irreducible operator set, *e.g.* some spinor (IROP) $\hat{\psi}^q$ that transforms according to IREP q with internal dimension d_q (in the case of plain abelian symmetries, it typically holds $d_q = 1$, *i.e.* the operator $d_\sigma^\dagger$ is the only member of the IROP). Thus the calculation of the very specific correlation function with respect to specific elements d_σ and $d_\sigma^\dagger$ above can replaced by the Greens function

$$G_{\psi^q}(\omega) \equiv \langle (\hat{\psi}^q)^\dagger \| \hat{\psi}^q \rangle_\omega. \qquad (19)$$

To be clear, if $d_q > 1$, this includes the scalar product of the spinor components, and thus one is actually calculat-

ing the *same* Greens functions as in Eq. (18) d_q times,

$$G_{\psi^q}(\omega) = \sum_{q_z=1}^{d_q} \langle (\hat{\psi}_{q_z}^q)^\dagger \| \hat{\psi}_{q_z}^q \rangle_\omega = d_q \cdot \langle d_{(\sigma)} \| d_{(\sigma)}^\dagger \rangle_\omega$$

$$\Rightarrow G_{(\sigma)}(\omega) = \tfrac{1}{d_q} G_{\psi^q}(\omega). \tag{20}$$

with $G_{(\sigma)}$ independent of σ within its multiplet, as implied by the round brackets. This apparent overhead, however, only affects the CGC space, so this is negligible numerical overhead, yet makes the calculation conceptually simple. Specifically, when calculating matrix elements and their contribution to the Greens function, eventually all indices can be fully contracted, so there is no need for a special treatment of a specific z-label that represents a peculiar $d_\sigma^\dagger$. Moreover, given the discussion of scalar operators in Sec. II E earlier, one realizes that the scalar product $\hat{\psi}^\dagger \cdot \hat{\psi}$ of the IROP $\hat{\psi}^q$ yields a scalar operator.

In the following, two explicit prototypical examples for correlation functions are given that are used explicitly for the numerical results presented in this paper. The first example is the spin-spin correlation function or magnetic susceptibility $\chi_d(\omega)$ defined at some site d. In the presence of spin SU(2) symmetry,

$$\chi_d(\omega) = \langle S_{x,d} \| S_{x,d} \rangle_\omega = \langle S_{y,d} \| S_{y,d} \rangle_\omega = \langle S_{z,d} \| S_{z,d} \rangle_\omega$$

$$\equiv \tfrac{1}{3} \langle \hat{\mathbf{S}}_d \| \hat{\mathbf{S}}_d \rangle_\omega. \tag{21}$$

Clearly, the local operator $\hat{S}_d^2 \equiv \hat{\mathbf{S}}_d \cdot \hat{\mathbf{S}}_d$ is a scalar operator, with the corresponding spinor $\hat{\mathbf{S}}_{(d)}$ given by [*cf.* App. Eq. (A11)]

$$\hat{\mathbf{S}} \equiv \begin{pmatrix} -\frac{1}{\sqrt{2}} \hat{S}_+ \\ \hat{S}_z \\ +\frac{1}{\sqrt{2}} \hat{S}_- \end{pmatrix}.$$

The second example is the spectral function for a single spinful channel in the presence of spin and particle-hole SU(2) symmetry. The spinor is given by [*cf.* App. Eq. (A59)],

$$\hat{\psi} \equiv \begin{pmatrix} s\hat{c}_\uparrow^\dagger \\ \hat{c}_\downarrow \\ s\hat{c}_\downarrow^\dagger \\ -\hat{c}_\uparrow \end{pmatrix}.$$

In the evaluation of the correlation function,

$$\langle \hat{\psi}^\dagger \| \hat{\psi} \rangle_\omega = \sum_{q_z=1}^{d_q=4} \langle \hat{\psi}_{q_z}^\dagger \| \hat{\psi}_{q_z} \rangle_\omega$$

$$= \langle \hat{c}_\uparrow \| \hat{c}_\uparrow^\dagger \rangle_\omega + \langle \hat{c}_\downarrow^\dagger \| \hat{c}_\downarrow \rangle_\omega + \langle \hat{c}_\downarrow \| \hat{c}_\downarrow^\dagger \rangle_\omega + \langle \hat{c}_\uparrow^\dagger \| \hat{c}_\uparrow \rangle_\omega$$

the signs, including $s \equiv \pm 1$, of the individual components are irrelevant. Given the spin symmetry and the

fact, that in the presence of particle-hole symmetry spectral functions are symmetric with respect to $\omega = 0$, and in general $G_{B,B^\dagger}(\omega) \equiv \langle B \| B^\dagger \rangle_\omega = G_{B^\dagger,B}(-\omega)$, it follows that all four contributions above describe exactly the same function, indeed, and therefore

$$G_\sigma(\omega) \equiv \langle \hat{c}_\sigma \| \hat{c}_\sigma^\dagger \rangle_\omega = \tfrac{1}{4} \langle \hat{\psi}^\dagger \| \hat{\psi} \rangle_\omega.$$

IV. THE NUMERICAL RENORMALIZATION GROUP

The non-abelian setup described above is straightforwardly applicable to the NRG.[3,29] Before doing so in detail, here a brief reminder of the essentials of NRG is given, followed by the introduction of the model Hamiltonian to be analyzed. By construction, the NRG deals with so-called quantum impurity models – an arbitrary small quantum system (the *impurity*) that is in contact with a macroscopic non-interacting typically fermionic bath. Each part is simple to solve exactly on its own. The combination of both, specifically in presence of interactions at the location of the impurity, however, gives rise to strongly-correlated quantum-many-body effects.

The systematic approach introduced by Wilson[3] was a logarithmic discretization in energy space of the continuum of the bath (coarse graining), followed by an exact mapping onto a semi-infinite so-called Wilson-chain, with the intact impurity space coupled only to the very first site of this chain. Given the half-bandwidth $W := 1$ of the bath, the discretization parameter $\Lambda > 1$, typically $\Lambda \gtrsim 1.7$, defines the logarithmic discretization in terms of the intervals $\pm [\Lambda^{-m}, \Lambda^{-(m+1)}]$ with $m \geq 0$ an integer, and energies taken relative to the Fermi energy $\varepsilon_f \equiv 0$. Each of these intervals is then described by a single effective fermionic state, with its coupling and exact energy position chosen consistently *w.r.t.* the hybridization of the original continuum model.[30,31] The resulting discretized model is then mapped onto the semi-infinite Wilson-chain (Lanczos tridiagonalization).[32] Hereby, the logarithmic discretization of the non-interacting bath translates to an effective tight-binding chain, with the hopping $t_k \sim \Lambda^{-k/2}$ between sites k and $k+1$, decaying *exponentially* in the discretization parameter Λ. The latter then justifies the essential renormalization group ansatz of the NRG in terms of *energy scale separation* – large energies are considered first, with approximate eigenstates at large energies *discarded* and considered unimportant for the description of the still following lower energy scales. Thus each site of the Wilson chain corresponds to an energy shell with a characteristic energy scale $\omega_k \equiv \frac{a}{2}(\Lambda+1)\Lambda^{-k/2}$. Here the constant a of order 1. is chosen such, $t_k/\omega_k \to 1$ for large k.[28]

In practice, when considering the system up to site k, the Hamiltonian of the rest of the system is ignored, equivalent to assuming degeneracy in the state space of the remainder of the system. With $\hat{H}_k$ the full Hamiltonian $\hat{H}$ including the Wilson chain up to site k, its

eigenstates $|s\rangle_k$, $\hat{H}_k|s\rangle_k = E_s^k|s\rangle_k$, and with $|e\rangle_k$ an arbitrary state of the rest of the system following site k, then the essential spirit of NRG after coarse graining of the bath can be condensed in the following approximation,[33]

$$H|se\rangle_k \simeq E_s^k|se\rangle_k, \tag{22}$$

expressing energy scale separation, with $|se\rangle_k \equiv |s\rangle_k \otimes |e\rangle_k$. The energies E_s^k are usually taken relative to the ground state energy E_0^k of iteration k, and rescaled by the energy scale ω_k. All of this will be referred to as rescaled energies, and has the advantage that independent of the Wilson shell k, energies are always of order 1.

In this paper, the state space truncation at a given NRG iteration is energy-based, *i.e.* all states with $E_s^k \leq E_K$ are kept, typically with $E_K \simeq 5\ldots7$ in rescaled energies. The number of kept states N_K thus changes dynamically.[28,29,31]

A. Full density matrix

Within the NRG,[3] a complete many-body basis set can be formulated from the state space discarded at every iteration.[33] Initially introduced for explicit time-dependence of quantum quenches, they actually can also be used to improve on existing calculations for thermodynamical quantities and expectation values,[34] with a clean extension to arbitrary temperatures using the full density matrix (FDM).[23] The density matrix $\hat{\rho} \equiv e^{-\beta\hat{H}}/Z$ with $\beta = 1/k_BT$, k_B the Boltzmann constant and T the temperature, obviously commutes with the Hamiltonian and is a scalar operator in itself. Within the FDM-NRG approach,[23] the density matrix

$$\hat{\rho} = \frac{1}{Z} \sum_{k;qn;q_z;e} e^{-\beta E_{qn}^k} |qn; q_z; e\rangle_k\ {}_k\langle qn; q_z; e|, \tag{23}$$

can be constructed straightforwardly in terms of a QSpace for every Wilson shell k. Here $s \equiv (qn)$ stands for the multiplet label for a given shell k. Note that the symmetry of the states e is irrelevant here, as this space is fully traced over. Given the usual practice of NRG to rescale and shift energies at every iteration, all of this, of course, must be undone before entering Eq. (23) [given a general thermal density matrix, of course, *all* energies in Eq. (23) must be (i) at the same energy scale, *i.e.* *non-rescaled*, and (ii) specified with respect to a *common* energy reference, *e.g.* the overall ground state energy of the Wilson chain].[23]

By construction, all eigenenergies E_{qn}^k are degenerate, *i.e.* do not depend on the z-labels. With the reduced density matrix being a scalar operator, therefore the CGC spaces in the QSpaces describing Eq. (23) are all proportional to identity matrices, leading to the overall normalization

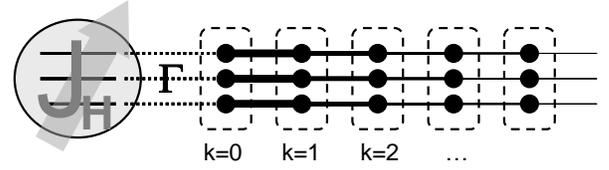$$Z = \sum_{k;qn} d_q d_w^{N-k} e^{-\beta E_{qn}^k}, \tag{24}$$

Figure 7. Schematic depiction of the fully screened Kondo-Anderson hybrid model [Eq. (25) with $m = 3$] in the NRG setup of a Wilson chain. Three d-levels with onsite Hund's interaction of strength J_H couple uniformly to their respective channel with hybridization Γ. The semi-infinite Wilson chain for each channel represents a tight-binding chain with exponentially decaying couplings, that interacts with the other channels through the impurity only. For a given NRG iteration, all terms in the Hamiltonian of the same energy scale must be included simultaneously, leading to an extended Wilson site [dashed boxes] of three spinful fermionic levels with a state space of $3^4 = 64$ states each.

where d_q is the internal dimension of multiplet q, and d_s^{N-k} reflects the degeneracy *w.r.t.* the rest of the Wilson chain of final length N, with d_w the state space dimension of a Wilson site.[23]

B. Model: symmetric three-channel system

The historically first physical system where Kondo physics was observed was that of Fe impurities in Au.[21,22] The effective microscopic model for this material, however, is far from trivial. It was argued only very recently in an extended study[20] that the physics of the five d-orbitals of substitutional Fe in Ag or Au is dominated by 3-fold degenerate triplet space t_{2g}, with the doublet space e_g split-off by crystal fields and thus playing a minor role. Together with the effective spin $3/2$ of the iron impurity, this then results in an SU(3) symmetric fully screened 3-channel Kondo model.

The actual model analyzed[20] is depicted schematically in Fig. 7. It consists of $m = 3$ spinful d-levels comprising the impurity, that are interacting through the Hund's coupling of strength J_H. Each of these impurity levels is coupled to its own spinful bath channel with uniform hybridization Γ. This leads to the Kondo-Anderson hybrid Hamiltonian,

$$\hat{H} \equiv \hat{H}_d + \sum_{i=1}^{m=3} \sum_{p\sigma}\left[\sqrt{\tfrac{2\Gamma}{\pi}}\big(\hat{d}_{i\sigma}^\dagger \hat{c}_{ip\sigma} + \text{H.c.}\big) + \varepsilon_p \hat{c}_{ip\sigma}^\dagger \hat{c}_{ip\sigma}\right], \tag{25a}$$

where all energies will be given in context in units of the half-bandwidth $W := 1$. The impurity is described by

$$\hat{H}_d \equiv -J_H \hat{\mathbf{S}} \cdot \hat{\mathbf{S}}, \tag{25b}$$

with the impurity spin

$$\hat{\mathbf{S}}_\alpha \equiv \sum_i^m \hat{\mathbf{S}}_{i,\alpha} \equiv \sum_{i=1}^m \sum_{\sigma\sigma'} (\tfrac{1}{2}\tau_\alpha)_{\sigma\sigma'} \hat{d}_{i\sigma}^\dagger \hat{d}_{i\sigma'} \tag{25c}$$

given in terms of the Pauli matrices τ_α with $\alpha \in \{x, y, z\}$. Here $\hat{d}_{i\sigma}^\dagger$ $[\hat{c}_{ip\sigma}^\dagger]$ creates a particle with spin $\sigma \in \{\uparrow, \downarrow\}$ on d-level i at energy $\varepsilon_d = 0$ [in bath channel i at energy ε_p], respectively. For $J_H \gtrsim \Gamma$, an effective spin-3/2 forms at the impurity, leading to a symmetric fully-screened spin-3/2 system. The resulting Kondo temperature T_K decays exponentially with J_H/Γ, with T_K quickly becoming the smallest energy scale in the system. In practice, choosing $J_H = 2\Gamma/(m + \frac{1}{2})$ leads to comparable Kondo temperatures T_K for different m. Compared to the standard Kondo Hamiltonian with $\mathbf{S} \cdot \mathbf{s}$ coupling of the dot spin $\mathbf{S}$ with the lead spin $\mathbf{s}$, the Hamiltonian in Eq. (25) in terms of Γ and J_H also allows for charge-fluctuations, while the model maintains particle-hole symmetry.

In particular, the Anderson-like model in Eq. (25) has the advantage that the impurity self-energy $\Sigma(\omega)$ can be evaluated within the NRG in a simple fashion. From a more technical point of view, this allows the straightforward calculation of an improved spectral function from the self-energy.[35] The impurity Greens function [cf. Eq. (20)]

$$
\begin{aligned}
G_{(i\sigma)}(\omega) &\equiv \langle \hat{d}_{(i\sigma)} \| \hat{d}_{(i\sigma)}^\dagger \rangle_\omega \\
&\equiv G'_{(i\sigma)}(\omega) - i\pi G''_{(i\sigma)}(\omega),
\end{aligned} \tag{26}
$$

consisting of real and imaginary part, respectively, is constructed within the NRG framework, as usual, from the spectral function $A_{(i\sigma)}(\omega) \equiv -\frac{1}{\pi} \mathrm{Im} G_{(i\sigma)}(\omega) \equiv G''_{(i\sigma)}(\omega)$. Subsequently, the real part $G'_{(i\sigma)}(\omega)$ is obtained through Kramers-Kronig transform of $A_{(i\sigma)}(\omega)$.[29] The calculation of the additional correlation function $F_{(i\sigma)}(\omega)$ then,

$$
F_{(i\sigma)}(\omega) \equiv \langle [\hat{d}_{(i\sigma)}, \hat{H}_d] \| \hat{d}_{(i\sigma)}^\dagger \rangle_\omega, \tag{27}
$$

obtained similarly from its spectral part $F''_{(i\sigma)}(\omega) \equiv -\frac{1}{\pi} \mathrm{Im} F(\omega)$, allows to evaluate the self-energy $\Sigma(\omega)$ at the impurity[35]

$$
\Sigma_{(i\sigma), J_H} \equiv \frac{F_{(i\sigma)}}{G_{(i\sigma)}}, \tag{28}
$$

Note that, the commutator of the IROP $\hat{d}_{(i\sigma)}$ with the scalar Hamiltonian in Eq. (27) again leads to an IROP *w.r.t.* the same IREP q. Moreover, by symmetry, both $G_{(i\sigma)}$ and $F_{(i\sigma)}$ are independent of $(i\sigma)$, as indicated by the subscript bracket, and hence will be skipped altogether in the following, for simplicity.

1. Kondo limit from numerical perspective

While the procedure to obtain the self-energy is straightforward for an Anderson-like model, there is no simple way to do so for the plain Kondo-like model with $\mathbf{S} \cdot \mathbf{s}$ interaction.[35] However, from the NRG point of view, the transition from one to the other is straightforward. That is, knowing that the Kondo temperature T_K decays

exponentially with J_H/Γ, both, J_H as well as Γ can be taken *much* larger than the bandwidth $W := 1$ of the model, while keeping their ratio constant,

$$
J_H, \ \Gamma \gg 1, \qquad \frac{J_H}{\Gamma} = \text{const.} \tag{29}
$$

This is a well-known procedure in the analytical Schrieffer-Wolff transformation for the Anderson model into a Kondo model.[36] But, of course, exactly the same strategy can also be pursued here within the NRG [see Fig. 8 later]. For the local density of states at the impurity this leads to a well-separated nearly discrete contribution to the spectral function at $|\omega| \gg 1$ far outside the bandwidth. For the spectral range within the bandwidth, the actual spectral function for the Kondo-model emerges. In particular, this procedure allows to fully eliminate the free-orbital (FO) regime with strong charge-fluctuations in the Anderson-like model right within the first truncation step. From a numerical point of view, this is desirable as the FO regime is typically the most expensive one. For example, for the model discussed here using the symmetries below, using energy-based truncation indicates that about a factor of $5 \ldots 10$ more multiplets are required for the FO regime as compared to the local moment (LM) or strong coupling (SC) regime at later NRG iterations [*cf.* Fig. 9]. Nevertheless, by maintaining an Anderson-like description, the impurity self-energy remains easily accessible numerically within the NRG, even though essentially the correlation functions for the Kondo model are calculated.

V. NRG RESULTS

The model in Eq. (25) is a true three-channel system, in that it is not possible to simply decouple a certain unitary superposition of bath channels. Furthermore, within an NRG iteration, a site from each channel must be included as they have the same coupling strength, *i.e.* energy scale, as schematically depicted in Fig. 7 [dashed boxes].

The non-abelian symmetries present in the system are,

- total spin symmetry: $\mathrm{SU}(2)_\mathrm{spin}$,

- particle-hole symmetry in each of the three channels: $\mathrm{SU}(2)_\mathrm{charge}^{\otimes 3}$, and

- channel symmetry: $\mathrm{SU}(3)_\mathrm{channel}$.

The latter symmetry $\mathrm{SU}(3)_\mathrm{channel}$, however, *does not* commute with particle-hole symmetry, while it does commute with the total charge $\mathrm{U}(1)_\mathrm{charge}$, *i.e.* the abelian subalgebra of particle-hole symmetry [*cf.* App. Eq. (A72), and subsequent discussion]. Having non-commutative symmetries, however, directly suggests a larger enveloping symmetry, which in the present case is the symplectic symmetry $\mathrm{Sp}(6)$ [*i.e.* $\mathrm{Sp}(2m)$ with $m = 3$, *cf.* App. A 10].

P7.

This allows to consider the following symmetry settings,

$$SU(2)_{\text{spin}} \otimes SU(2)_{\text{charge}}^{\otimes 3} \equiv SU(2)_{\text{SC}}^{\otimes 4}, \qquad (30a)$$

$$SU(2)_{\text{spin}} \otimes U(1)_{\text{charge}} \otimes SU(3)_{\text{channel}}, \text{ and} \qquad (30b)$$

$$SU(2)_{\text{spin}} \otimes Sp(6). \qquad (30c)$$

All of these symmetry settings have been implemented, in practice, and applied within the NRG framework, with results presented in the following. The first setting in (30a) represents a more traditional NRG scenario based on a set of plain SU(2) symmetries. The second setting (30b) includes $SU(3)_{\text{channel}}$ together with the simple abelian symmetry $U(1)_{\text{charge}}$ for total charge, while the last setting (30c) represents the actual full symmetry of the model.

Even though the second setting in (30b) actually includes an abelian component in terms of charge, it nevertheless represents a stronger symmetry as compared to the first setting (30a). Since $SU(3)_{\text{channel}}$ is a rank-2 symmetry with two commuting z-operators, *i.e.* generators of the Cartan subalgebra, it possesses a *two*-dimensional multiplet representation. This results in much larger multiplets for setting (30b), with the bare SU(3) multiplet dimensions easily reaching up to 100 (*e.g.* $d_q \leq 125$ for the NRG run underlying Fig. 12, *cf.* App. C 3). As a consequence, this allows, on average, smaller multiplet spaces and thus better numerical performance.

The first two symmetry settings in (30) emphasize different symmetry aspects, yet allow to break certain symmetries which, nevertheless, are present in the model Hamiltonian in Eq. (25). The first symmetry setup (30a) strongly emphasizes particle-hole symmetry, while it does not use the symmetric coupling of the levels to their respective channels. The channel symmetry can thus be broken without reducing the symmetry setting (30a). The second symmetry setting (30b), on the other hand, emphasizes the channel symmetry (uniform Γ), while it allows to break the particle-hole symmetry. Hence, in principle, a uniform level-shift could be applied to the d-levels within this setting. Only the third symmetry (30c) captures the full symmetry of the model, as it combines channel symmetry with particle-hole symmetry into the enveloping symmetry Sp(6). This is a rank-3 symmetry with multiplet dimensions now easily reaching up to a several thousands (*e.g.* see Tbl. VIII for actual multiplets generated in a full NRG run). A more detailed general discussion and comparison of all of above symmetry setups in terms of their overall multiplet structure and representation of a site with three spinful levels (*i.e.* a Wilson site) is given in App. C 3.

A. Spectral functions

The spinor $\hat{\psi}^q$ to be used for fermionic hopping term as well as for the calculation of spectral functions can be represented for all symmetry settings by IROPs with a well-defined multiplet label q [*cf.* Eq. (20)]. For the first symmetry setting in (30), $SU(2)_{\text{SC}}^{\otimes 4}$, the IROP for the calculation of spectral functions involves three 4-component spinors, $\hat{\psi}_i^{[4]}$ for short, one for every channel $i = 1, \ldots, (m = 3)$. The corresponding IROP labels are $q_1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$, $q_2 = (\frac{1}{2}, 0, \frac{1}{2}, 0)$, and $q_3 = (\frac{1}{2}, 0, 0, \frac{1}{2})$, respectively. The number of components in the spinor derives from the two participating SU(2) multiplets $(S, C) = \frac{1}{2}$ for spin and one specific channel, thus having $2 \times 2 = 4$ operators in one specific IROP $\hat{\psi}_i^{[4]}$, indeed. With this, the hopping in the Hamiltonian, for example, is given by $\hat{h}_{k,k+1} = \sum_{i=1}^m \hat{\psi}_{k,i}^{[4]\dagger} \cdot \hat{\psi}_{k+1,i}^{[4]}$. Note that this *excludes* the hermitian conjugate part, as this is already fully incorporated through the particle-hole symmetry [see App. Eq. (A73b)]. Furthermore, note that particle-hole symmetry gives rise to intrinsic even-odd alternations for the spinors along a chain [see App. A 9 b for a detailed discussion].

In contrast, the second symmetry setting in (30), $SU(2)_{\text{spin}} \otimes U(1)_{\text{charge}} \otimes SU(3)_{\text{channel}}$, leads to a single 6-component spinor, $\hat{\psi}^{[6]}$ for short. Its IROP multiplet label is given by $q = (\frac{1}{2}, -\frac{1}{2}, 0\ 1)$. This combines a 2-dimensional $SU(2)_{\text{spin}}$ multiplet $S = \frac{1}{2}$ and an abelian 1-dimensional multiplet $C_z = -\frac{1}{2}$ with the 3-dimensional $SU(3)_{\text{channel}}$ multiplet $T = (0\ 1)$, resulting in the $2 \times 1 \times 3 = 6$ operators in the multiplet. For comparison, here the hopping term in the Hamiltonian in Eq. (25) is reduced to a total of two terms, $\hat{h}_{k,k+1} = (\hat{\psi}_k^{[6]\dagger} \cdot \hat{\psi}_k^{[6]} + \text{H.c.})$ [see App. Eq. (A73a)].

Finally, for the third symmetry setting in (30), $SU(2)_{\text{spin}} \otimes Sp(6)$, again a single spinor $\hat{\psi}^q$ is obtained, but now with 12 components, written as $\hat{\psi}^{[12]}$ for short. Its IROP label is given by $q = (\frac{1}{2}, 1\ 0\ 0)$, which combines the 2-dimensional $SU(2)_{\text{spin}}$ multiplet $S = \frac{1}{2}$ with the 6-dimensional Sp(6) multiplet $(1\ 0\ 0)$, *i.e.* defining representation of Sp(6). Overall, this again recovers the $2 \times 6 = 12$ components of the spinor, indeed. For comparison, now the hopping term in the Hamiltonian in Eq. (25) is reduced to the *single* term $\hat{h}_{k,k+1} = (\hat{\psi}_k^{[12]\dagger} \cdot \hat{\psi}_{k+1}^{[12]})$. The scalar contraction of the spinor $\hat{\psi}^{[12]}$ with itself recovers the original 12 terms in the fermionic hopping structure between two sites in the Hamiltonian in Eq. (25). Since particle-hole symmetry is part of Sp(6), this again implies that (i) the hermitian conjugate is already taken care of in the hopping term, and (ii) that Sp(6) again gives rise to the same intrinsic even-odd alternations for the spinors along a chain, exactly analogous to what has already been encountered for standard particle-hole symmetry.

The correlation functions calculated for the model in Eq. (25) are presented in Fig. 8, with the model parameters indicated at the bottom left of panel (a). Panel (a) shows the spectral data on a linear scale, while panel (b) shows the same data vs. $\log(|\omega|)$ which therefore allows a logarithmic zoom into the low energy regime. The legend shown with panel (b) also applies to panel (a).
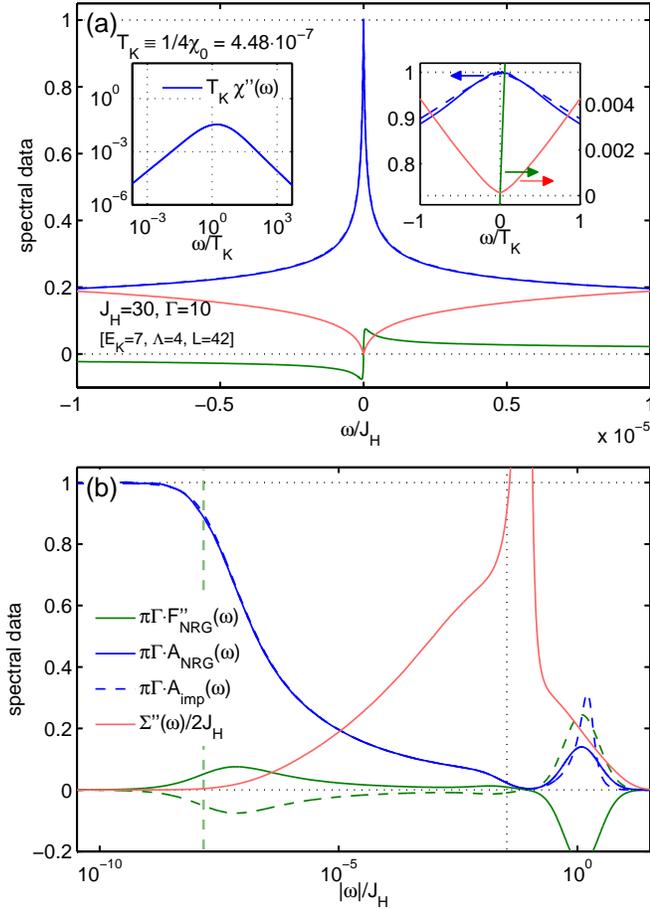
P7.

Figure 8. (Color online) $SU(2)_{\text{spin}} \otimes U(1)_{\text{charge}} \otimes SU(3)_{\text{channel}}$ analysis of the symmetric three-channel Anderson model [Eq. (25) with model parameters specified in the lower left of panel (a)]; the same data as in panel (a) is shown vs. $\log(|\omega|)$ in panel (b) to zoom into the Kondo peak at small frequencies with the legend for both panels shown in panel (b). The spectral data A_{NRG} and the auxiliary F''_{NRG} are shown together with the derived self-energy $\Sigma''(\omega)$ and the improved spectral function A_{imp} (see text). A zoom around $\omega = 0$ is show in the right inset of panel (a), with the left (right) axis belonging to $A(\omega)$ [$\Sigma''(\omega)$ and $F''(\omega)$], respectively. The spectral data for $A(\omega)$ and $\Sigma''(\omega)$ is symmetric around $\omega = 0$ and strictly positive, while $F''(\omega)$ is antisymmetric. In panel (b) therefore the $\omega < 0$ branch of $F''(\omega)$ has been plotted in dashed lines, same color otherwise. The left inset to panel (a) shows the spin-spin spectral data $\chi''(\omega)$, with the resulting $T_K \equiv 1/4\chi_0$ indicated in panel (b) and the left insets of both panels by the vertical dashed line.

The data in Fig. 8 was obtained using the symmetry setting in (30b) including $SU(3)_{\text{channel}}$. Note that having chosen an energy-based NRG truncation with $E_K = 7$, the spectral data for the other two symmetry settings is identical, hence not shown. While the calculation is somewhat more involved for the more traditional setup (30a), it becomes significantly more compact still when finally including $Sp(6)$ as in (30c). Their individual nu-

merical efficiency will be discussed with Fig. 9 below.

In Fig. 8, the spectral function obtained from the NRG is plotted as $A_{\text{NRG}}(\omega) \equiv G''_{\text{NRG}}(\omega)$. The spectral data satisfies the Friedel sum-rule to an excellent approximation, in that $\lim_{\omega \to 0}(\pi \Gamma A_{\text{NRG}}(\omega)) = 1$ [see right inset to panel (a) for a zoom around $\omega = 0$]. The self-energy $\Sigma(\omega)$ was obtained by calculating the additional correlation function $F_{\text{NRG}}(\omega)$ [Eq. (27), to be used in Eq. (28)]. The imaginary part $\Sigma''(\omega) \equiv -\frac{1}{\pi}\text{Im}\Sigma(\omega)$, plotted in Fig. 8, clearly approaches zero in a smooth parabolic fashion at the Fermi energy, i.e. $\lim_{\omega \to 0} \Sigma(\omega) = \lim_{\omega \to 0} \frac{d}{d\omega}\Sigma(\omega) = 0$, as expected for a system who's low-energy behavior corresponds to that of a Fermi liquid. This is seen more clearly still in the zoom around $\omega = 0$ in right inset of Fig. 8(a), with the self-energy data associated with the right axis. The self-energy $\Sigma(\omega)/J_H$ sharply drops within $|\omega| \lesssim T_K$ from order 1. accurately down to about 10^{-4} which is considered the NRG resolution limit.

The *improved* spectral function $A_{\text{imp}}(\omega)$ derived from the self-energy[35] is also shown in Fig. 8 [dashed red (black) line]. Within the Kondo regime, the result closely follows the original $A_{\text{NRG}}(\omega)$, as demonstrated in the zoom in the right inset of Fig. 8(a) or also in panel (b). As expected from the self-energy treatment,[35] the improved spectral function $A_{\text{imp}}(\omega)$ allows clearly sharper resolution for structures at finite frequencies, specifically so for larger Λ. This can be observed, for example, for the hybridization side peaks in Fig. 8(b) at the energy of the Hund's coupling J_H. Having chosen J_H much larger than the bandwidth [with the bandwidth indicated by the vertical dotted line in panel (b)], these hybridization side peaks essentially correspond to very narrow, nearly discrete peaks that are much overbroadened through the standard log-Gauss broadening of the NRG.[23,29] In principle, these side peaks could be narrowed significantly further by an adaptive broadening scheme.[37] For the purposes of this paper, however, this was irrelevant.

The dynamically generated exponentially small Kondo temperature T_K for the system can be determined by taking the full-with-at-half-maximum (FWHM) of the Kondo peak in the spectral function. However, with NRG somewhat sensitive to broadening of the underlying discrete data[29] (see also supplementary material in Ref. 23), T_K is simply determined therefore through the static magnetic susceptibility $\chi_0 =: 1/4T_K$,[35] where χ_0 is obtained from the impurity spin-spin correlation function $\chi(\omega) \equiv \langle S_{(z),d} \| S_{(z),d} \rangle_\omega \equiv \chi'(\omega) - i\pi\chi''(\omega)$ evaluated at $\omega = 0$, with $S_{(z),d}$ the total spin at the impurity [cf. Eq. (21)]. The resulting spin-spin spectral function $\chi''(\omega)$ is shown in the left inset to Fig. 8(a), together with the resulting $T_K = 4.4 \cdot 10^{-7}$ (in units of bandwidth). As expected, $\chi''(\omega)$ shows a pronounced maximum around T_K. The value for T_K is also indicated by the vertical dashed line in panel (b).

The NRG data presented in Fig. 8 clearly suggests converged data, even without necessarily having to resort to self-energy to get the low energy physics correct.[20] The convergence is also supported by the analysis of the *dis-*

P7.

carded weight[28] which, inspired by DMRG, analyzes the decay of the eigenspectrum of site-specific reduced density matrices built from the ground state space a few iterations later. For given NRG run, the discarded weight is estimated as $\varepsilon^D_{\chi=5\%} = 3 \cdot 10^{-11}$. This suggests good convergence, in agreement with Ref. 28. If, for example, an energy truncation of $E_K = 5, \ldots, 6$ had been used, instead, NRG intrinsic parameter dependent deviations of up to ten percents can still be seen *w.r.t.* to the Friedel sum-rule or the agreement of $A_{\mathrm{NRG}}(\omega)$ with $A_{\mathrm{imp}}(\omega)$.

B. Detailed comparison of symmetry settings

An NRG specific technical comparison of the symmetry settings in Eq. (30) for the calculation in Fig. 8 is presented in Fig. 9. The underlying truncation had been energy-based in all calculations ($E_K = 7$), thus leading to a fair comparison in terms of accuracy. With this, the physical properties, and in particular the energy flow diagram[3,35] in Fig. 9(a), show perfect agreement using either symmetry setting. Having sufficiently many states implies that for symmetries, that are not explicitly and thus exactly included in the QSpace setup, their unintended breaking due to numerical double precision noise does not play role.

Figure 9(b) shows the number of kept multiplets for each iteration k. Having chosen J_H and Γ much larger than the bandwidth [*cf.* Fig. 8(a)], the free-orbital regime is absent, with the transition from the local moment to the strong coupling regime given by the energy scale of T_K [vertical dashed line at $k \simeq 22$]. As expected from physical grounds, also the local moment regime ($k < 22$) requires a larger state space (multiplet) dimension still for the same accuracy, *i.e.* the same E_K, as compared to the strong coupling regime ($k > 22$).

With the state space truncation based on the energy cutoff $E_K = 7$, the actual Hilbert state space dimension, *i.e.* when including the internal CGC space dimensions, is exactly the same for all symmetry settings. In particular, the maximum total Hilbert state space dimension per iteration that was diagonalized exactly for either symmetry setting was $N^*_{\mathrm{tot}} \leq 4,369,024$ or $N^*_K \leq 68,266$ *w.r.t.* kept space only. These state spaces could be strongly reduced to the effective and manageable multiplet dimension as indicated at the top of Fig. 9(b), with Wilson shell specific multiplet dimensions plotted in the panel.

Figure 9(c) analyzes the actual reduction in multiplet space due to presence of the CGS spaces in terms of the ratio of the actual Hilbert space dimension N^*_{tot} relative to the total multiplet dimension N_{tot} for each site along the Wilson chain. Depending on the symmetry setting, on average, the treatment of non-abelian symmetries allows to reduce the Hilbert space dimension by at least a factor of 16, 20, or 300 for the symmetries in (30), respectively. This demonstrates an enormous numerical gain, considering that the numerical cost of NRG roughly scales like $\mathcal{O}(N^3_{\mathrm{tot}})$. Note that it is exactly through the di-
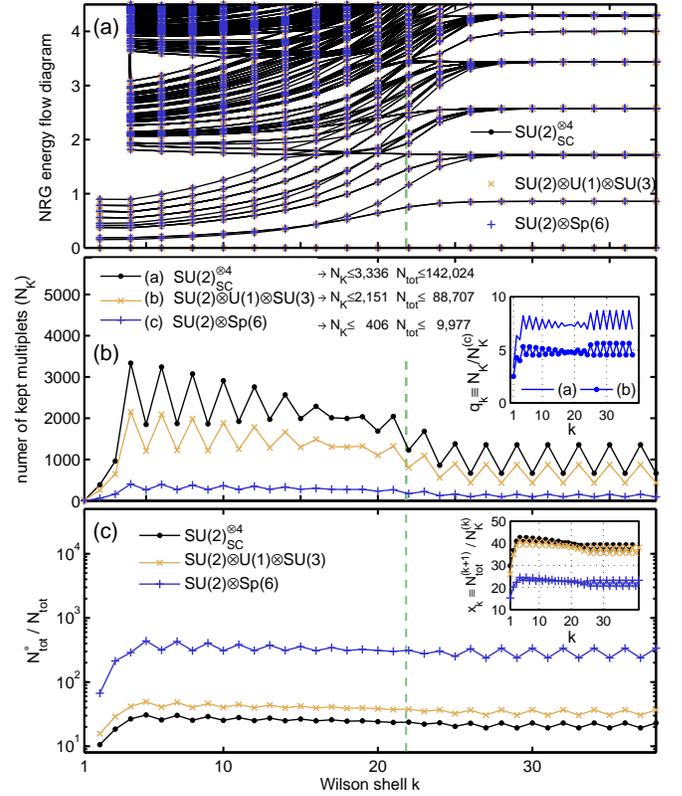

Figure 9. (Color online) Comparison of the efficiency of the symmetry settings as outlined in Eq. (30) for the calculation of the spectral data in Fig. 8 for the 3-channel model in Eq. (25). For a fair comparison, all calculations were performed using the same energy-based truncation with $E_K = 7$ for the same discretization $\Lambda = 4$ as in Fig. 8. The vertical dashed lines in all panels indicates the energy scale of T_K. Panel (a) compares the energy flow diagrams resulting for even iterations from the individual NRG runs, indicating perfect consistency for all symmetry settings. Panel (b) shows the number of kept multiplets for each iteration. For each symmetry setting, at the top of the panel the maximum dimension in the multiplet space over the entire NRG run is specified for kept (N_K) and total (N_{tot}), *i.e.* kept and discarded space, respectively. The inset shows the ratio $q_k \equiv N_K/N^{(c)}_K$ of the multiplets that needed to be kept for the symmetry settings in Eq. (30a) and Eq. (30b) relative to the case when the full Sp(6) is included [Eq. (30c)]. Panel (c) shows the ratio $N^*_{\mathrm{tot}}/N_{\mathrm{tot}}$ of the actual Hilbert-space dimension (N^*_{tot}) at a given iteration, which includes the internal multiplet dimensions, relative to the dimension of the multiplet space (N_{tot}). The inset shows the ratio x_k that describes the increase in the number of multiplets when adding a new site prior to truncation.

mensional reduction to multiplet spaces, that above NRG calculations had been feasible in practice, and this within a few hours of runtime. In contrast, the plain abelian setting simply would not have been able to deal with the underlying Hilbert state space dimension using state of the art workstations [*cf.* App. Tbl. VI].

Within the kept space, the multiplet dimension of the

first two settings in (30a) and (30b) relative the setting including the Sp(6) are shown in the inset to Fig. 9(b) [(a) and (b), respectively]. This clearly demonstrates a further reduction by a factor of about $5\ldots8$ when including the full Sp(6) symmetry. From the same inset, it is also clear that the symmetry setting in Eq. (30b) including SU(3)$_{\mathrm{channel}}$ allows, on average, a 40% further reduction of the number of multiplets in the simulation as compared to the SU(2)$_{\mathrm{SC}}^{\otimes4}$ setting.

Furthermore, the inset to Fig. 9(c) shows the ratio $x_k \equiv N_{\mathrm{tot}}^{(k+1)}/N_{\mathrm{K}}^{(k)}$ which indicates the increase in total number of multiplets when adding a new site *prior* to truncation. While this factor shows a clear reduction from the actual dimension of the local Hilbert space of a Wilson site of $4^3 = 64$ states, the ratio x_k is somewhat larger than what one may naively expect, considering that, depending on the symmetry, a Wilson site reduces to a total number of 4, 10 or 13 multiplets [see App. Tbl. V, Tbl. IV, and Tbl. III, respectively]. On the other hand, given non-abelian symmetries, the combination of two multiplets typically leads to clearly *more* than just one overall multiplet. In this sense, the major gain of using non-abelian symmetries is given by the state space reduction demonstrated in Fig. 9(b). For the first two symmetry settings in (30), the multiplet space increase by adding a new site in terms of a product space reduces the original abelian factor of 64 only modestly down to about 38. Only when using of the full Sp(6), this leads to a significant further reduction of the ratio x_k down to about 20, which thus becomes nearly comparable in numerical cost to a two-channel calculation with abelian symmetries, where a Wilson sites adds $4^2 = 16$ states to the system.

The SU(3) representations that are explicitly generated in the calculation of Fig. 8 using SU(2)$_{\mathrm{spin}} \otimes$ U(1)$_{\mathrm{charge}} \otimes$ SU(3)$_{\mathrm{channel}}$ are listed in Fig. 12 [App. C 3]. The largest Clebsch-Gordan space that is split off with respect to the SU(3) sector only is the $(4,4)$ representation with an internal multiplet dimension of 125. In other words, by explicitly accounting for SU(3) symmetries, in the present case, a 125-fold degeneracy in the Hamiltonian had been reduced to a *single* multiplet, with the SU(3) symmetry space taken care of separately with minor computational overhead. Nevertheless, the eigenstates in the SU(3) setting still show significant degeneracies. These can be entirely removed only by using the full Sp(6) symmetry, which allows to remove *original degeneracies in the Hamiltonian of several thousands.* Note that on top of above symmetries, the spin SU(2) multiplets present yet another independent multiplet space that enters as a tensor product, thus enlarging the overall symmetry space still further.

In terms of overall runtime on a state-of-the-art 8-core workstation, this translated to about 6 hours of runtime for the SU(2)$_{\mathrm{SC}}^{\otimes4}$ symmetries, as compared to about 4.5 hours of runtime when including SU(3)$_{\mathrm{channel}}$. Using the full symmetry as in (30c), on the other hand, took about 24 hrs. While significantly more efficient in terms of stor-

age requirements [*cf.* Tbl. VII] thus facilitating calculations on standard workstations, the huge CGC spaces in the last setting must be dealt with carefully. As can be seen from Fig. 9(b), the total number of kept multiplets hardly reaches 400, while the Sp(6) multiplets are fully comparable in terms of dimensionality, with some multiplets even much larger internally than the actual number of multiplets considered [*cf.* App. Tbl. VIII]. While the sparse algebra had been optimized by ourselves to also make use of the parallel shared memory capacity [*cf.* App. C 3 b], in contrast, the full multiplet spaces had access to the highly optimized shared BLAS libraries. The latter benefitted the first two symmetry settings (30a) and (30b) in terms of overall runtime. However, there is clearly room for further improvement in dealing with the sparse algebra for larger rank symmetries as in (30c).

VI. SUMMARY AND OUTLOOK

A generic and transparent framework has been presented for the implementation of non-abelian symmetries in tensor-networks in terms of QSpaces. For this, it was assumed that all participating state spaces are strictly orthonormal and can be assigned proper well-defined symmetry labels. Therefore the presented framework is straightforwardly applicable to the traditional DMRG as well as to the NRG. The latter was demonstrated in detail in this paper for an SU(3) symmetric 3-channel problem which, in the presence of particle-hole symmetry, can be enlarged into the symplectic symmetry Sp(6). By reducing the actual state space to the reduced multiplet space, while factorizing the Clebsch Gordan coefficient space, this allows an efficient description of all relevant tensors. While the explicit Clebsch Gordan algebra bears little overhead for combinations of lower rank symmetries, the average internal multiplet dimensions grow quickly with increasing rank r of a symmetry. In practice, one may roughly estimate that the typical internal multiplet dimension grows like $\mathcal{O}(10^r)$, for example, having $r = 0, 1, 2, 3$ for abelian, SU(2), SU(3), and Sp(6), respectively. Starting with $r = 3$, an efficient sparse scheme on all CGC spaces becomes crucial. For symmetries with rank larger than three, finally, it appears desirable to develop general strategies and sum rules for the contraction of extended complex networks of CGC spaces.

A detailed self-contained general introduction to non-abelian symmetries is given in the appendix, together with many explicit examples that arise in practice (App. A and C). Several further highlights explained in detail in the appendix (App. B) are: (i) a straightforward numerical recipe for the general calculation of Clebsch Gordan coefficients based on explicit product space decomposition in the presence of multiplicity, (ii) a generic recipe for the determination of irreducible operator sets, and last but not least, (iii) also a general algorithm to get the framework for several symmetries initialized from

plain Fock space. The latter does not require any initial detailed knowledge of specific symmetry labels other than the general action of the underlying generators. These are known in second-quantized form and thus also easily defined in Fock space.

P7.

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (Jan 2001)

[2] S. R. White, Phys. Rev. Lett. **69**, 2863 (Nov 1992)

[3] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (Oct 1975)

[4] S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997)

[5] A. Weichselbaum, F. Verstraete, U. Schollwock, J. I. Cirac, and J. von Delft, Phys. Rev. B **80**, 165117 (2009)

[6] A. W. Sandvik and G. Vidal, Phys. Rev. Lett. **99**, 220602 (2007)

[7] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. A **75**, 033605 (Mar 2007)

[8] J. I. Cirac and F. Verstraete, arXiv:0910.1130v1 [cond-mat.str-el](2009)

[9] S. Singh, R. N. C. Pfeifer, and G. Vidal, Phys. Rev. A **82**, 050301 (Nov 2010)

[10] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (Apr 1973)

[11] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys. Rev. Lett. **100**, 070502 (Feb 2008)

[12] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (Feb 2010)

[13] G. Vidal, Phys. Rev. Lett. **99**, 220405 (Nov 2007)

[14] G. Vidal, Phys. Rev. Lett. **101**, 110501 (Sep 2008)

[15] E. M. Stoudenmire and S. R. White, arXiv:1105.1374v2 [cond-mat.str-el](2011)

[16] I. P. McCulloch and M. Gulcsi, EPL (Europhysics Letters) **57**, 852 (2002)

[17] A. I. Tóth, C. P. Moca, Ö. Legeza, and G. Zaránd, Phys. Rev. B **78**, 245109 (2008)

[18] S. Singh, R. N. C. Pfeifer, and G. Vidal, arXiv:1008.4774v1 [cond-mat.str-el](2010)

[19] L. De Leo and M. Fabrizio, Phys. Rev. Lett. **94**, 236401 (Jun 2005)

[20] T. A. Costi, L. Bergqvist, A. Weichselbaum, J. von Delft, T. Micklitz, A. Rosch, P. Mavropoulos, P. H. Dederichs, F. Mallet, L. Saminadayar, and C. Bäuerle, Phys. Rev. Lett. **102**, 056802 (2009)

[21] W. J. de Haas, J. de Boer, and G. J. van dën Berg, Physica **1**, 1115 (1934)

[22] J. Kondo, Progress of Theoretical Physics **32**, 37 (1964)

[23] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007)

[24] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (Apr 2005)

[25] J. P. Elliott and P. G. Dawber, *Symmetry in physics, Vol. I+II* (Oxford University Press, New York, 1979)

[26] A. Alex, M. Kalus, A. Huckleberry, and J. von Delft, J. Math. Phys. **52**, 023507 (2011), ISSN 00222488

[27] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (Aug 2000)

[28] A. Weichselbaum, Phys. Rev. B **84**, 125130 (Sep 2011), http://xxx.lanl.gov/abs/1106.0916

[29] R. Bulla, T. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008)

[30] M. Yoshida, M. A. Whitaker, and L. N. Oliveira, Phys. Rev. B **41**, 9403 (1990)

[31] R. Žitko and T. Pruschke, Phys. Rev. B **79**, 085106 (2009)

[32] J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the solution of algebraic eigenvalue problems: a practical guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000) ISBN 0-89871-471-0

[33] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005)

[34] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006)

[35] R. Bulla, A. C. Hewson, and T. Pruschke, J. Phys.: Condens. Matter **10**, 8365 (1998)

[36] J. R. Schrieffer and P. A. Wolff, Phys. Rev. **149**, 491 (Sep 1966)

[37] A. Freyn and S. Florens, Phys. Rev. B **79**, 121102 (Mar 2009)

[38] C. Pope, *Geometry and Group Theory* (Online lecture notes, Texas A& M University, 2006)

[39] R. Gilmore, *Lie groups, Lie algebras, and some of their applications* (Dover Publications, 2006)

[40] J. J. Sakurai, *Modern Quantum Mechanics*, Revised ed. (Addison-Wesley Publishing Company, Reading, MA, 1994)

[41] M. A. A. van Leeuwen, A. M. Cohen, B. Lisser, and C. Arjeh, *LiE, A Package for Lie Group Computations* (Université de Poitiers, 1992) ISBN 90-74116-02-7, `http://www-math.univ-poitiers.fr/~maavl/LiE/`

Appendix A: Non-abelian symmetries 101

The general more pedagogical introduction of non-abelian groups in this appendix emerges from a practical numerical background of treating quantum many-body phenomena. It does not claim to cover non-abelian symmetries in every theoretical detail, yet requires certain elementary concepts which will be briefly reviewed. The main focus of this appendix then is on practical applications in quantum lattice models. Specifically, this targets the numerical renormalization group (NRG),[3,29] density matrix renormalization group (DMRG)[2,24] or more generally tensor networks,[9,18] yet also exact diagonalization, which itself may be formulated in a matrix product state language. The appendix offers a general treatment of continuous non-abelian symmetries, with modifications towards abelian, point-groups, or discrete non-abelian symmetries straightforward. Overall, this appendix should be self-contained, sufficient and hopefully helpful to deal with general abelian and non-abelian symmetries in numerical simulations.

The non-abelian symmetries of concern in this paper are continuous symmetries. An element $\hat{G}$ of the corresponding *Lie group* $\hat{\mathcal{G}}$ can be parameterized by a set of g continuous, independent, and real parameters a_σ,[25]

$$\hat{G}(a_1, \ldots, a_g) = \exp\Big(i \sum_{\sigma=1}^{g} a_\sigma \hat{S}_\sigma\Big), \qquad \text{(A1)}$$

with g the *dimension* of the symmetry group. Infinitesimal operations with $a_\sigma \ll 1$ then define the set of g generators $\{\hat{S}_\sigma\}$, the number of which thus also reflects the dimension of the group (note that the identity operator is a trivial operation which therefore is never part of the set of generators). For unitary symmetries, as considered throughout in this paper, the generators in Eq. (A1) are hermitian. Furthermore, when dealing with exponentially large yet *finite*-dimensional quantum-many-body Hilbert spaces, the non-abelian symmetries also must have finite-dimensional Lie algebras.

The commutator relations of the generators in Eq. (A1),

$$[\hat{S}_\sigma, \hat{S}_\mu] = \sum_\nu f_{\sigma\mu\nu} \hat{S}_\nu, \qquad \text{(A2)}$$

determine the tensor of the *structure constants* $f_{\sigma\mu\nu}$, which itself fully defines the underlying *Lie algebra*. The tensor $f_{\sigma\mu\nu}$ is antisymmetric in that by construction $f_{\sigma\mu\nu} = -f_{\mu\sigma\nu}$, yet not necessarily fully antisymmetric also *w.r.t.* to the last index ν [in principle, it can be made fully antisymmetric using the *Cartan-Killing metric*, while distinguishes between co- and contravariant indices in Eq. (A2);[38,39] for simplicity, however, this distinction is not made in this paper]. All generators are assumed to be connected to each other through above commutator relations. That is, if a subgroup of generators fully decouples in that it commutes with the rest

of the generators, then this subgroup forms a symmetry of its own. In this sense the group of generators for a specific *simple* symmetry is irreducible.

A set of matrices $\{R_\sigma\}$, that obeys exactly the same commutator relations as the generators (operators) $\{\hat{S}_\sigma\}$ in Eq. (A2), allows a one-to-one correspondence between the matrices $\{R_\sigma\}$ and the generators of the symmetry. It is called a *matrix representation* of the Lie algebra. If the *carrier space*, *i.e.* the vector space within which the matrix representation is defined, is fully explored through repeated application of the individual matrices of the representation, then this is called an *irreducible* matrix representation, to be denoted as $\{I_\sigma\}$ henceforth. It is unique up to an overall similarity transformation. Together with its carrier space it refers to an irreducible representation (IREP), specified by a unique label q. If, on the other hand, part of the carrier space of a matrix representation decouples, the representation is called *reducible*. This will be discussed in significantly more detail later in the context of state space decomposition in Secs. A 5 and A 6.

Consider an irreducible matrix representation $\{I_q^q\}$ for IREP q of dimension d_q. Its carrier space is spanned by the *multiplet* $|q\rangle \equiv \{|qq_z\rangle\}$, where q_z references the individual states within the multiplet q (consider, for example, spin multiplets, where $|qq_z\rangle \equiv |S, S_z\rangle$). The states $|qq_z\rangle$ forms an irreducible space *w.r.t.* the action of the generators, in that for an arbitrary symmetry operation $\hat{G}$ as in Eq. (A1),

$$\hat{G}|qq_z\rangle = \sum_{q_z'} G_{q_z,q_z'}^q |qq_z'\rangle, \qquad \text{(A3a)}$$

some linear superposition within the same multiplet space arises. The coefficients $G_{q_z,q_z'}^q$ form a $d_q \times d_q$ dimensional matrix, which represents the symmetry operation $\hat{G}$ within multiplet q, and is given by $G^q \equiv \exp(i \sum_\sigma a_\sigma I_\sigma^q)$ for some arbitrary but fixed values a_σ.

Similar to the multiplet space $|qq_z\rangle$ of dimension d_q, an *irreducible operator* (IROP) set $\hat{F}^q \equiv \{\hat{F}_{q_z}^q\}$ can be defined in a completely analogous manner. While it is not constrained to a specific carrier space, the IROP $\hat{F}^q$ consists of a set of d_q *operators* that are associated with multiplet q. As such, it can be written as a vector of operators, *i.e.* a generalized spinor. For a given symmetry operation $\hat{G}$ then, the IROP transforms analogously to Eq. (A3a), which for an operator implies

$$\hat{G}\hat{F}_{q_z}^q \hat{G}^{-1} = \sum_{q_z'} G_{q_z,q_z'}^q \hat{F}_{q_z'}^q. \qquad \text{(A3b)}$$

On the level of infinitesimal operations, $|a_\sigma| \ll 1$, in contrast to the plain action of generators on a ket-state as in Eq. (A3a), Eq. (A3b) shows that the transformation of an IROP directly translates to *commutator relations* [*l.h.s.* of Eq. (A3b)] with the generators of the symmetry, instead.

The practical relevance of above general statements will be discussed in much detail in what follows, together

with many examples relevant in actual numerical calculations.

1. Simple example: rotational symmetry

A simple and well-known example of a non-abelian symmetry is the rotational group O(3) in real space in three dimensions. An arbitrary rotation can be written as $G = e^{iS}$ with S an arbitrary hermitian, yet fully complex three-dimensional *matrix* (hence no hats). The latter is required for G to be real. Consequently, this leaves three real parameters (a_x, a_y, a_z), with $S = a_x S_x + a_y S_y + a_z S_z$. The generators[25]

$$S_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, S_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, S_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{(A4)}$$

represent infinitesimal rotations around the x, y, and z-axis, respectively. The O(3) symmetry therefore has dimension $g = 3$, and its Lie algebra is defined by,

$$[\hat{S}_\sigma, \hat{S}_\mu] = i \sum_\nu \varepsilon_{\sigma\mu\nu} \hat{S}_\nu, \quad \text{(A5)}$$

with $\sigma, \mu, \nu \in \{x, y, z\}$ and $\varepsilon_{\sigma\mu\nu}$ the Levi-Civita tensor, having switched to general operator notation [operator $\hat{S}$ (with hat) rather matrix S]. Being generators of O(3), the matrix representation in Eq. (A4) already represents a 3-dimensional IREP. As it is the simplest non-trivial IREP for O(3), it is also called its *defining representation*. By combining state spaces that share this symmetry then, many other IREPs can be generated, including, for example, the (trivial) scalar representation of dimension 1.

With respect to continuous functions $f(x, y, z)$ in three-dimensional space, the generators of infinitesimal rotations are given by the differential operator described by the angular momentum operator $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ with $\hat{\mathbf{p}} \sim \nabla_{\mathbf{r}}$. By construction, its three components $\hat{L}_i$ also obey exactly the same Lie algebra as the generators in Eq. (A5). The same also holds for the spin algebra SU(2) in complex space, which describes the symmetry for spinful particles such as electrons if rotational spin symmetry is not broken, *i.e.* in the absence of an external magnetic field. Hence the rotational group O(3) is isomorphic to the spin SU(2). In contrast to O(3), however, the defining representation of SU(2) is *two*-dimensional [*cf.* Eq. (A6)], and hence also allows half-integer spin multiplets, which are entirely absent in O(3). Having essentially twice as many multiplets in SU(2) as compared to O(3), SU(2) is thus called a double cover or 2:1 cover of O(3).

2. SU(2) **spin algebra**

In this paper, the setup and notation for non-abelian symmetries is generalized from SU(2). Therefore the symmetry SU(2) will be recapitulated in some more detail, introducing the semantics used for the general treatment of non-abelian symmetries. In this sense, the semantics used in this paper is somewhat more inclined towards the physics background, rather than strictly adhering to the mathematical language of Lie algebras. The latter, nevertheless, will be indicated in context.

Similar to the O(3) symmetry, an arbitrary unitary transformation in two-dimensional complex space is given by $G = e^{iS}$ with S an arbitrary two-dimensional hermitian matrix. This again has three independent real parameters (a_x, a_y, a_z), such that $S = a_x S_x + a_y S_y + a_z S_z$. Here $S_\sigma = \frac{1}{2}\tau_\sigma$, with $\sigma \in \{x, y, z\}$, is given by the standard Pauli spin matrices τ_σ,

$$\tau_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \tau_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \tau_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{(A6)}$$

For SU(2), this represents the smallest non-trivial matrix representation, therefore this also becomes its defining representation. The commutator relations of the matrices τ_σ are *exactly* the same as for O(3) in Eq. (A5), since SU(2) also refers to the same rotational symmetry. Therefore, the generators for SU(2) will again also be denoted by the operators $\{\hat{S}_\sigma\}$ with $\sigma \in \{x, y, z\}$ in what follows.

For a general irreducible representations of SU(2), *e.g.* a spin multiplet, the usual choice of basis is such that the z-component of the spin operator, $\hat{S}_z$, becomes diagonal in its matrix representation S_z, while the other two operators $\hat{S}_x$ and $\hat{S}_y$ remain non-diagonal (due to their non-commuting properties, only one spin component can be fully diagonalized, given the freedom of a similarity transformation for the whole representation). Using the notation $|qq_z\rangle \equiv |S, S_z\rangle$ for general spin multiplets, the multiplet label q (q-label) then can take the values $q = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ with the internal multiplet label (z-label) spanning the $2q + 1$ values $q_z \in \{-q, -q+1, \dots, +q\}$. The raising and lowering operators (RLOs) are defined as[40]

$$\hat{S}_\pm \equiv \hat{S}_x \pm i\hat{S}_y, \quad \text{(A7)}$$

such that $\hat{S}_- \equiv (\hat{S}_+)^\dagger$, with the commutator relations

$$[\hat{S}_z, \hat{S}_\pm] = \pm \hat{S}_\pm \quad \text{(A8a)}$$

$$[\hat{S}_+, \hat{S}_-] \equiv [\hat{S}_+, \hat{S}_+^\dagger] = 2\hat{S}_z. \quad \text{(A8b)}$$

For spin multiplets $|qq_z\rangle$ then, it holds[40]

$$\hat{S}_z |qq_z\rangle = q_z |qq_z\rangle$$
$$\hat{S}_\pm |qq_z\rangle = \sqrt{q(q+1) - q_z(q_z \pm 1)} \ |q, q_z \pm 1\rangle. \quad \text{(A9)}$$

While the operator set $\{\hat{S}_x, \hat{S}_y, \hat{S}_z\}$ generates the SU(2) symmetry group, this set itself *does not* yet represent an irreducible operator (IROP), in that it does not yet transform according to a specific symmetry multiplet. For this, a specific linear superposition of the original operators as in Eq. (A8b) is required. In particular,

the transformation of an IROP set under given symmetry is completely analogous to the transformation of the symmetry eigenstates in Eq. (A9). As indicated with Eq. (A3b), the major difference is that the action of a generator $\hat{S}_\sigma$ applied onto a state is simply replaced by the *commutator* of the generator with an operator. For example, for an IROP $\hat{F}^q$ given in terms of the set of operators $\{\hat{F}^q_{q_z}\}$ which transform like the (state) multiplet q, it follows for consistency with Eq. (A9),

$$[\hat{S}_z, \hat{F}^q_{q_z}] = q_z \hat{F}^q_{q_z} \tag{A10a}$$

$$[\hat{S}_\pm, \hat{F}^q_{q_z}] = \sqrt{q(q+1) - q_z(q_z \pm 1)} \cdot \hat{F}^q_{q_z \pm 1} \tag{A10b}$$

This allows, for example, to complete the operator $\hat{S}_z$ into an *irreducible* spin operator set as follows. Clearly, $[\hat{S}_z, \hat{S}_z] = 0 \cdot \hat{S}_z$, which implies that the operator $\hat{S}_z$ has z-label $q_z = 0$, *i.e.* $\hat{S}^q_0 \equiv \hat{S}_z$ with q still unknown. Applying the RLOs yields the operators corresponding to $q_z = \pm 1$,

$$\underbrace{[\hat{S}_\pm, \hat{S}^q_0]}_{=[\hat{S}_\pm, \hat{S}_z] = \mp \hat{S}_\pm} = \sqrt{q(q+1) + 0} \cdot \hat{S}^q_{\pm 1}.$$

With the further application of RLOs yielding zero, *i.e.* $[\hat{S}_+, \hat{S}_+] = [\hat{S}_-, \hat{S}_-] = 0$, the operator space is thus exhausted. The *irreducible* spin operator set therefore has three members $q_z \in \{-1, 0, +1\}$, and thus transforms like a spin multiplet $q = \max(q_z) = 1$,

$$\hat{S}^1 \equiv \{\hat{S}^1_{q_z}\} \equiv \begin{pmatrix} \hat{S}_{1,+1} \\ \hat{S}_{1,0} \\ \hat{S}_{1,-1} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \hat{S}_+ \\ \hat{S}_z \\ +\frac{1}{\sqrt{2}} \hat{S}_- \end{pmatrix}. \tag{A11}$$

Note that the signs and prefactors are crucial for consistency with the Wigner-Eckart theorem later.

In above derivation, the z-operator in Eq. (A10a) allowed to directly determine the z-label q_z. The RLOs in Eq. (A10b), on the other hand, served to explore the multiplet space, in that they generated the remaining operators $\hat{F}^q_{q_z'}$ with proper well-defined prefactors. In given case of spin SU(2), these factors are known [*cf. r.h.s.* of Eq. (A10b)]. In situations, where they may not be known right away, they can nevertheless be determined in a straightforward manner. For simplicity, in the absence of inner multiplicity for given multiplet, for canonical raising or lowering operator $S_\pm$ (see Sec. A 3 b) the combined application of $\hat{S}_\pm$ followed by $\hat{S}^\dagger_\pm$ onto an operator of given multiplet q results in the same operator, *i.e.*

$$[\hat{S}^\dagger_\pm, [\hat{S}_\pm, \hat{F}^q_{q_z}]] = a^2_\pm \hat{F}^q_{q_z},$$

from which the prefactor $a^2_\pm$ can be easily determined. The analogous situation for a state space multiplet $|qq_z\rangle$ is $\hat{S}^\dagger_\pm \hat{S}_\pm |qq_z\rangle = a^2_\pm |qq_z\rangle$, with $a^2_\pm \geq 0$ since $\hat{S}^\dagger_\pm \hat{S}_\pm$ is a positive operator; in case of spin SU(2), this exactly reflects the prefactor on the *r.h.s.* of Eq. (A10b), *i.e.*

$a^2_\pm = q(q+1) - q_z(q_z \pm 1) \geq 0$. Therefore if the application of $\hat{S}_\pm$ results in a new operator component in the multiplet, *i.e.* $a^2_\pm > 0$, then this operator is exactly given by

$$\hat{F}^q_{q_z'} = \frac{1}{\sqrt{a^2_\pm}} [\hat{S}_\pm, \hat{F}^q_{q_z}]. \tag{A12}$$

This already contains the correct normalization and sign, with the the latter strictly determined by the outcome of the commutator. The z-label q_z can be derived directly from the structure constants of the underlying Lie algebra, *i.e.* Eq. (A10a). For a more general discussion on IROPs and their general decomposition also in the presence of inner multiplicity for the IROP multiplet q, see Sec. A 7.

3. Generators and symmetry labels

Symmetries $\mathcal{S}$ within a quantum mechanical framework are described by a set of generators $\hat{S}_\sigma$ that leave the Hamiltonian $\hat{H}$ of the system invariant. Therefore it must hold for all generators of the symmetries considered that

$$[\hat{S}_\sigma, \hat{H}] = 0. \tag{A13}$$

Thus by definition, the Hamiltonian is a scalar operator. The generators of independent symmetries $\mathcal{S}$ and $\mathcal{S}'$ commute trivially, by definition, as they operate in independent symmetry sectors. Therefore, for simplicity, a single specific non-abelian symmetry $\mathcal{S}$ is considered in the following, also referred to as *simple* non-abelian symmetry, a prototypical example being SU(N).

Therefore let $\mathcal{S}$ be a simple non-abelian symmetry. By construction then, its set of generators $\{\hat{S}_\sigma\}$ is fully connected via the structure constants in Eq. (A2), *i.e.* is irreducible but not necessarily an IROP yet [*e.g.* see previous discussion for SU(2)]. With the symmetry reflected in the unitary transformation $\hat{G} = e^{i\varepsilon \hat{S}_\sigma}$ with hermitian $\hat{S}_\sigma$ [*cf.* Eq. (A1)], it follows that for infinitesimal $\varepsilon \ll 1$, the invariance of the Hamiltonian under this unitary transformation, *i.e.* $\hat{U}\hat{H}\hat{U}^\dagger = \hat{H}$, is trivially equivalent to Eq. (A13).

In order to ensure maximally independent generators, all operators in $\{\hat{S}_\sigma\}$ can be taken orthogonal with respect to each other and specifically also with respect to the identity matrix (which is always excluded from the set of generators $\{S_\sigma\}$). This requires a scalar or inner product for matrices, which is provided by

$$\langle A, B \rangle \equiv \text{tr}(A^\dagger B), \tag{A14}$$

together with the resulting Frobenius norm $\|A\|^2 = \langle A, A \rangle = \text{tr}(A^\dagger A)$. For the generators of the symmetry, thus one requires

$$\text{tr}(S^\dagger_\sigma S_{\sigma'}) = a_\sigma \delta_{\sigma\sigma'} \tag{A15a}$$

$$\text{tr}(S_\sigma) = \text{tr}(\mathbf{1}^{(\dagger)} S_\sigma) = 0, \tag{A15b}$$

The generators in Eq. (A15b) are understood as finite-dimensional matrix representations of the operators $\hat{S}_\sigma$ in some specific carrier space, here the defining representation. Moreover, the orthogonality *w.r.t.* to the identity in the last equation implies that all generators S_σ are traceless.

a. Z-operators (Cartan subalgebra)

For a given simple non-abelian symmetries, it is always possible to identify a maximal set of mutually commuting hermitian generators which form the so-called *Cartan subalgebra* of the symmetry's Lie algebra. These can be fully diagonalized simultaneously (together with the Hamiltonian), and hence can be considered diagonal. They shall be referred to as the *z-operators* [as they generalize the concept of the operator S_z for SU(2)],

$$[\hat{S}_z, \hat{S}_{z'}] = 0 \qquad \text{(z-operators)}. \qquad (A16)$$

For a given Hamiltonian $\hat{H}$ then, this implies that every eigenstate $\hat{H}|n\rangle = E_n|n\rangle$, in addition, can also be labeled with its respective set of symmetry labels $|n\rangle \to |qn; q_z\rangle$, leading to

$$\hat{H}|qn; q_z\rangle = E_{qn}|qn; q_z\rangle. \qquad (A17)$$

Here q identifies the multiplet, *i.e.* a set of states q_z that are connected in an irreducible manner through all of the generators of the symmetry. While the index n originally identified all states in given Hilbert space, it is now sufficient that it labels the multiplet within the space of multiplets that share the same q. The composite index (qn) then is referred to as *multiplet index*. Similarly, also the eigenenergies E_{qn} in Eq. (A17) acquire symmetry labels. These, however, are independent of q_z since, by construction, the states within a symmetry multiplet are degenerate in energy. More generally, with q_z entirely determined by symmetry for a given multiplet q, they can easily be generated and thus omitted where convenient.

Given a specific multiplet qn, the labels q_z are equal to the eigenvalues of the z-operators,

$$\hat{S}_z|qn; q_z\rangle = q_z|qn; q_z\rangle \qquad \text{(z-labels)}, \qquad (A18)$$

which will be referred to as *z-labels*. If more than one z-operator is associated with given symmetry $\mathcal{S}$, say a total of r z-operators, where r thus defines to the *rank* of the symmetry, then the z-label structure associated with a multiplet also consist of a collective set of r z-labels (note that r needs to be differentiated here from the rank r of a tensor or QSpace as used in the main text). For example, the symmetry group SU(N) has rank $r = N - 1$. Therefore the rank of SU(2) is 1, *e.g.* a single label q suffices to identify a state within an SU(2) spin multiplet. SU(3), on the other hand, already acquires a two-dimensional label structure for q_z, and thus also for q.

Note that the z-labels in Eq. (A18) for the states of a specific multiplet q may not necessarily be unique, in that the *same* q_z may occur multiple times. Let m_z describe how often a specific z-label occurs within given multiplet q. Then the presence of $m_z > 1$ for at least one z-label is called *inner multiplicity*. It is then necessary to introduce an extra label α that uniquely identifies the state within this degeneracy,

$$|qn; q_z\rangle \to |qn; q_z\alpha_z\rangle, \quad \text{(inner multiplicity)} \qquad (A19)$$

with $\alpha_z \in \{1, \ldots, m_z\}$. While inner multiplicity is absent for SU($N \leq 2$), it occurs on a regular basis for SU($N \geq 3$). The situation for *outer* multiplicity is analogous (see Sec. A 5).

The label for the entire multiplet q (to be referred to collectively as *q-labels*) is in principle arbitrary, yet must be unique to identify the multiplet. Since for a continuous symmetry infinitely many IREPs exist, it is natural that the q-labels inherit the r-dimensional label structure of the z-labels. In particular, it is possible to construct a set of r *scalar* operators, called Casimir operators (see Sec. **??**), that define a unique set of r constants for each multiplet. In practice, however, the q-labels are derived from $q \equiv \max\{q_z\}$, *i.e.* by the z-labels corresponding to the *maximum weight state* (see Sec. A 3 c) which in principle can be related to the constants derived from the Casimir operators.[25]

b. Raising and lowering operators (roots)

While for an arbitrary unitary element $\hat{G}$ of the symmetry *hermitian* $\{\hat{S}_\sigma\}$ are required, on the level of generators, in principle, arbitrary linearly-independent linear superpositions within the space of generators $\hat{S}_\sigma$ can be taken. Using such a reorganized set of generators, instead, this still preserves Eq. (A13), yet alters the structure constants $f_{\sigma\mu\nu}$ for given symmetry $\mathcal{S}$. This freedom is used in the following to define canonical raising and lowering operators, which are non-hermitian, in general.

Consider the action of a generator $\hat{S}_\sigma$ onto a symmetry eigenstate $|qn; q_z\rangle$. The z-operators are special, in that they are diagonal and hence return the same state, yet weighted by the eigenvalue q_z. The remaining generators, however, are non-diagonal, hence change the state and thus explore the multiplet space. In general, these generators can be reorganized such that all of them represent proper raising or lowering operators (RLOs), with the canonical commutator relations,

$$[\hat{S}_z, \hat{S}_\sigma] = f_{z\sigma\sigma}\hat{S}_\sigma \equiv f_{z\sigma}\hat{S}_\sigma, \qquad (A20)$$

with no summation over σ. The action of these canonical RLOs in z-label space, in the literature also referred to as *root space*, then defines the canonical form. By definition, the canonical RLOs $\{\hat{S}_\pm\}$ of a specific Lie algebra are expected to have the property that their application onto a

P7.

symmetry eigenstate in the multiplet with well-defined z-labels will generate another eigenstate of the z-operators, yet with raised or lowered, *i.e. well-defined different* z-labels. This is exactly what is expressed through the commutator relations in Eq. (A20). In particular, the structure constants take the simple form, where a non-zero contribution can only arise if the last two indices in $f_{z\sigma\sigma'}$ are identical, hence the shortcut notation $f_{z\sigma}$ in the last term in Eq. (A20). By construction, $f_{z\sigma}$ is fully antisymmetric. Note that Eq. (A20) also can be interpreted as an eigenvalue equation for the generators of the group. Since the z-operators $\hat{S}_z$ are symmetric, the resulting eigenvalue problem is always well-defined with real eigenvalues $f_{z\sigma}$.

As a specific example, Eq. (A20) was already encountered for SU(2) in Eq. (A8a). Here it states more generally that the commutator of an arbitrary generator $\hat{S}_\sigma$ with a z-operator yields the very same operator $\hat{S}_\sigma$ up to the scalar structure factor $f_{z\sigma}$. This factor can be zero, *e.g.* when $\hat{S}_\sigma$ refers to another z-operator as in Eq. (A16), therefore $f_{zz'} = 0$. For every z-operator, however, there must exist *at least* one RLO $\hat{S}_\sigma$ with $f_{z\sigma} \neq 0$, since otherwise the group of generators would be reducible.

With Eq. (A20), the application of a generator $\hat{S}_\sigma$ onto a symmetry eigenstate $|qn; q_z\rangle$ yields

$$\hat{S}_z \cdot \hat{S}_\sigma |qn; q_z\rangle = \underbrace{[\hat{S}_z, \hat{S}_\sigma]}_{\stackrel{(A20)}{=} f_{z\sigma}\hat{S}_\sigma}|qn; q_z\rangle + \hat{S}_\sigma \cdot \underbrace{\hat{S}_z|qn; q_z\rangle}_{=q_z|qn;q_z\rangle}$$
$$= (q_z + f_{z\sigma})\hat{S}_\sigma |qn; q_z\rangle. \quad (A21)$$

If $\hat{S}_\sigma$ is an RLO with $f_{z\sigma} \neq 0$, the state $\hat{S}_\sigma|qn; q_z\rangle$ is again a symmetry eigenstate, yet with a uniform shift in the z-labels,

$$q_z \rightarrow q_{z'} \equiv q_z + f_{z\sigma}. \quad (A22)$$

Therefore the action of an RLO $\hat{S}_\sigma$ in root space is generic, *i.e.* independent of the specific multiplet q or the state q_z under consideration. Nevertheless, the RLO may annihilate the state, *i.e.* $\hat{S}_\sigma|qn; q_z\rangle = 0$, which is essential to obtain a finite-dimensional multiplet space. Furthermore, Eq. (A22) allows to pair up raising and lowering operators. That is, if $\hat{S}_\sigma$ is a raising operator, then with

$$[\hat{S}_z, (\hat{S}_\sigma)^\dagger] = -[\hat{S}_z, \hat{S}_\sigma]^\dagger = -f_{z\sigma}(\hat{S}_\sigma)^\dagger, \quad (A23)$$

the operator $(\hat{S}_\sigma)^\dagger$ changes the z-labels exactly in the opposite direction as $\hat{S}_\sigma$ in Eq. (A22). In this sense, $(\hat{S}_\sigma)^\dagger \equiv \hat{S}_{-\sigma}$ represents the corresponding lowering operator. The actual definition of what is a raising or lowering operator is not entirely unique, as it depends on the specific underlying sorting scheme of the z-labels adopted in root space. This does not matter, however, as long as the sorting is done consistently throughout.[38,39]

In the presence of inner multiplicity a few complications arise. Most importantly, an RLO usually will generate a superposition in the $m_{z'}$-fold degenerate state space in the resulting $q_{z'}$,

$$\hat{S}_\sigma |qn; q_z \alpha_z\rangle = \sum_{\alpha_{z'}=1}^{m_{z'}} s^{[q\sigma]}_{q_z\alpha_z; q_{z'}\alpha_{z'}} |qn; q_{z'}\alpha_{z'}\rangle \quad (A24)$$

with some coefficients $s^{[q\sigma]}_{q_z\alpha_z; q_{z'}\alpha_{z'}}$. As a consequence, the application of a raising operator $\hat{S}_\sigma$ followed by its complimentary lowering operator $\hat{S}_\sigma^\dagger$ onto a symmetry eigenstate,

$$\hat{S}_\sigma^\dagger \hat{S}_\sigma |qn; q_z \alpha_z\rangle = \sum_{\alpha_{z'}=1}^{m_z} s^{[q\sigma]}_{q_z; \alpha_z \alpha_{z'}} |qn; q_z \alpha_{z'}\rangle \quad (A25)$$

with some other coefficients $s^{[q\sigma]}_{q_z; \alpha_z \alpha_{z'}}$, *does* return to the same symmetry labels q_z, yet *not necessarily to the same state*. If the resulting state in Eq. (A25) does not replicate the initial state $|qn; q_z\alpha_z\rangle$ up to an overall factor, then this allows to explore the other states in the degenerate subspace at q_z. This is relevant for the decomposition of state spaces, where the resulting state as in Eq. (A25) needs to be orthonormalized in a consistent fashion with respect to the already explored states of the multiplet including the state $|qn; q_z\alpha_z\rangle$ (see Sec. B 1, for more detail on the numerical implementation).

While all z-operators are required, *e.g.* for the definition of the z-labels, it is usually not required to explicitly construct all of the RLOs, as some of these operators can be generated through a product of a smaller set of RLOs. As will be seen below in the case of SU(N) or Sp(2n), the number of actually required RLOs can always be reduced to the rank of the symmetry, *i.e.* the number of z-operators. This minimal set of RLOs will be referred to as *simple* RLOs, consistent with their general notation in the literature as *simple roots* of the symmetry. In a sense, these simple RLOs are the ones that induce the smallest shifts in the z-labels.[38,39] Again, their definition is not entirely unique, depending on conventions such as normalization of generators or what specific sorting scheme is applied to the z-labels. The simple RLOs still fully generate and connect the state spaces of any IREP. The underlying intuitive reason is that an r-dimensional z-label structure only requires r linearly independent vectors to explore its space (for a rigorous proof, see for example Refs. 38 and 39). Therefore given r z-operators $\{Z_1, \ldots, Z_r\}$, it is sufficient to choose a specific subset of r raising operators $\{S_{1+}, \ldots, S_{r+}\}$, with the corresponding lowering operators $S_{i-} \equiv (S_{i+})^\dagger$. This reduction to simple RLOs is very useful in practice, yet does not restrict the non-abelian treatment in any way.

c. Maximum-weight state

Consider some multiplet q of internal dimension d_q for a given non-abelian symmetry group $\mathcal{S}$ of rank r. Then each of the d_q states carries a set of r z-labels. When depicted graphically as points in r-dimensional space, this

P7.

is called the *weight diagram* for the multiplet [for SU(3), for example, a collection of weight diagrams generated in an actual NRG run is shown in Fig. 12]. Since the z-operators are traceless, the values of the z-labels are naturally centered around the origin, *i.e.* $q_z = 0$. Inner multiplicity, if present, decreases as a function of distance $|q_z|$ to the origin, such that the outermost points in a weight diagram always refer to unique states without any remaining multiplicity. By choosing a *lexicographic* ordering in the r z-labels,[25] the *maximum weight* (MW) is defined by

$$q_{\mathrm{MW}} \equiv \max\{q_z\}. \tag{A26}$$

The state with $q_z = q_{\mathrm{MW}}$ is called the maximum-weight state. This state is guaranteed to be unique to the multiplet for non-abelian symmetries,[25,38,39] hence can be used as label for the entire multiplet, *i.e.* $q = q_{\mathrm{MW}}$. While the state space $|q_z| = \max(|q_z|)$ will not be unique, in general, since it refers to several states at the circumference of the weight diagram, $\max\{q_z\}$ does provide a unique set of z-labels. This underlines the importance of lexicographic ordering.

As an example, consider the well-known spin SU(2). The states within the multiplet q are labeled by $|q q_z\rangle$ where $q_z = -q, \ldots, +q$ identifies each state within the multiplet. This results in a one-dimensional weight diagram, with the multiplet itself labeled by the maximum weight states, $q = \max(q_z)$, indeed.

Clearly, the q-labels for a multiplet themselves are also not entirely unique and hence depend on convention. In particular, if the rank of a symmetry group $\mathcal{S}$ is $r > 1$, the order of the z-operators themselves is a priori arbitrary. Hence there is a certain freedom in the order of the z-labels, which in return affects the definition of the maximum weight state. Given a certain order in the z-operators then, the lexicographic sorting of sets of z-labels is typically done in *reverse* order, *i.e.* starting with the *last* of the r label for a given q_z. Moreover, having identified q_{MW}, this still leaves the freedom to use a linearly independent transform of q_{MW} as label for the entire multiplet for consistency with literature. For example, for SU(3) [Sp(6)] this is discussed with Eq. (A34) [Eq. (A91a)], respectively.

4. Example SU(N)

a. Defining representation

The symmetry SU(N) is defined as the unitary symmetry of an N-dimensional space. The defining representation, *i.e.* the IREP with smallest non-trivial dimension, is therefore given by $N \times N$ dimensional matrices. Since according to Eq. (A15b) all generators are traceless, only $N - 1$ diagonal z-operators exist, the diagonals of which form an N-dimensional orthogonal vector space that is also orthogonal to the diagonal of the identity matrix.

The raising (lowering) operators are chosen as $N \times N$ matrices with a single entry of 1. anywhere in the upper (lower) triangular space, respectively, away from the diagonal. For this, let

$$|e_i\rangle \equiv (0, \ldots, 0, 1_{(i)}, 0, \ldots, 0)^T, \tag{A27a}$$

with $i \in \{1, \ldots, N\}$ be the N-dimensional cartesian column basis vectors, and

$$E_{ij} \equiv |e_i\rangle\langle e_j|, \tag{A27b}$$

the matrices of the related operator basis, which also contain just a single entry of 1 in their $N \times N$ dimensional matrix space, *i.e.* $(E_{ij})_{i'j'} = \delta_{ii'}\delta_{jj'}$. Then the generators can be written as follows,

$$S_{i \neq j}^{\mathrm{SU}(N)} = E_{ij} = \begin{cases} \text{raising operator for } i < j \\ \text{lowering operator for } i > j \end{cases} \tag{A28a}$$

$$S_{z, k < N}^{\mathrm{SU}(N)} = \left(\sum_{i=1}^{k} E_{ii}\right) - k E_{k+1, k+1}. \tag{A28b}$$

These matrices are orthogonal as in Eq. (A15), while the (arbitrary) normalization was chosen such that, for convenience, all entries are integers. The choice of generators for SU(N) in Eq. (A28) guarantees canonical RLOs, and thus simplifies the groups commutator relations *w.r.t.* z-operators exactly the way as indicated in Eq. (A20). This can be easily seen by observing that for a diagonal operator of the type $(\hat{Z})_{ij} = z_i \delta_{ij}$, the matrix elements of the commutator with an arbitrary operator $(\hat{S})_{ij} = s_{ij}$ is given by

$$[\hat{Z}, \hat{S}]_{ij} = s_{ij}(z_i - z_j),$$

that is, existing non-zero matrix elements in $\hat{S}$ are weighted by differences in diagonal elements of $\hat{Z}$, while there cannot arise any new matrix elements unequal zero in $[\hat{Z}, \hat{S}]$ as compared to $\hat{S}$. Clearly, if $\hat{S}_\pm$ only has a single non-zero entry as for the operators in Eq. (A28a), it follows $[\hat{Z}, \hat{S}_\pm] = \mathrm{const} \cdot \hat{S}_\pm$, in agreement with Eq. (A20).

From Eq. (A28a) above, a total of $\frac{1}{2}N(N-1)$ different raising operators arise. However, not all of these are required to fully explore the multiplet space. Consider, for example, the subset of $r = N - 1$ raising operators

$$\{S_+^{\mathrm{SU}(N)}\}_r \equiv \{\hat{S}_{12}^{\mathrm{SU}(N)}, \hat{S}_{23}^{\mathrm{SU}(N)}, \ldots, \hat{S}_{N-1,N}^{\mathrm{SU}(N)}\}, \tag{A29}$$

which thus matches the rank r of the symmetry group SU(N) and thus also the number of z-operators. From repeated application of these operators, it is easily seen that the remaining raising operators not contained in Eq. (A29) can be generated. For example, $\hat{S}_{13}^{\mathrm{SU}(N)}$ is generated by $\hat{S}_{12}^{\mathrm{SU}(N)} \cdot \hat{S}_{23}^{\mathrm{SU}(N)}$. Therefore, above minimal set of r raising operators with their hermitian conjugate set of lowering operators is sufficient, indeed, to explore all multiplet spaces.

b. The symmetry SU(3)

The defining representation for SU(3) is chosen as in Eq. (A28), with the z-operators given by,

$$Z_1 \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; Z_2 \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad \text{(A30)}$$

Their diagonals can be collected as rows into a matrix,

$$z = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix}, \quad \text{(A31)}$$

the columns of which give the z-labels (z_1, z_2) for the three states in the defining representation (see large black dots in Fig. 10). This represents the weight diagram for the defining representation.

The canonical commutator relations as in Eq. (A20) yield the structure constants $(f)_{z,\sigma} \equiv f_{z\sigma}$ for $z \in \{1, 2\}$ and $\sigma \in \{12, 23, 31\}$,

$$f = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 3 & -3 \end{pmatrix}. \quad \text{(A32)}$$

The columns in Eq. (A32) thus define the roots, *i.e.* the shift in z-labels when applying either S_{12}, S_{23}, or S_{31}, respectively. These vectors (roots) are depicted in Fig. 10 by large thick arrows. Clearly, the three points in the weight diagram of the defining representation can be connected by these roots, equivalent to (repeatedly) applying raising or lowering operators.

With the convention, that z-labels are lexicographically sorted starting with the last z-label, *i.e.* sorting w.r.t. z_2 first and then z_1, the three states in the defining representation are already properly sorted from largest to smallest [left to right in Eq. (A30)]. Furthermore, Eq. (A32) shows that S_{12} and S_{23} correspond to positive roots, since $(2, 0) > (0, 0)$ and also $(-1, 3) > (0, 0)$. As their application makes z-labels *larger*, they represent raising operators, indeed, while S_{31} is a lowering operator, all in agreement with Eq. (A28a). The third raising operator thus would be S_{13} with root $(1, 3)$ which, however, is not a simple root and hence can be dropped.

Finally, SU(3) still contains well-known SU(2) subalgebras. That is, for example, by using S_{12} as a raising operator for the (x, y) subspace together with its corresponding z-operator $[S_{12}, S_{12}^\dagger] =: 2S_z^{(12)} = Z_1$ while keeping the y component abelian, this shows that every *line of points* in the (z_1, z_2) plane in Fig. 10 parallel to S_{12} must correspond to a proper SU(2) multiplet. The same also holds for the two remaining permutations of (x, y, z) using S_{23} or S_{31} for the SU(2) subspace, instead. These SU(2) subalgebras clearly obey the standard commutator relations for SU(2).
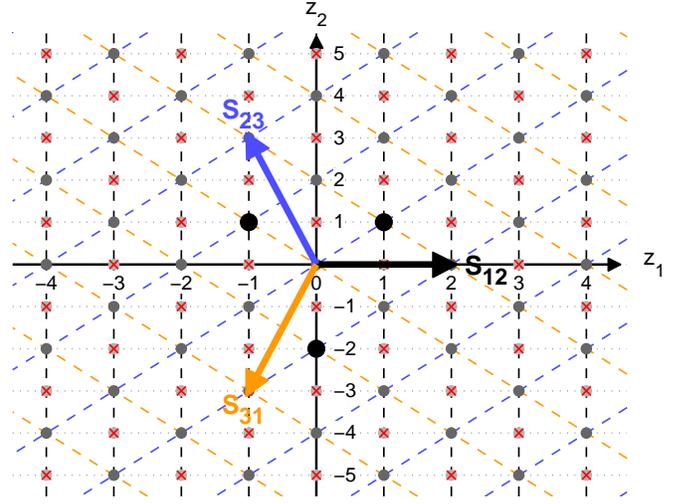
Figure 10. (Color online) Root space (z_1, z_2) for SU(3). The three large black dots depict the weight diagram of the three-dimensional defining representation of SU(3). Large arrows indicate the shifts in z-labels due to the action of the raising operators S_{12}, S_{23}, and the lowering operator S_{31}, while dashed lines close to orthogonal to these arrows indicate lines of constant S_{12}, S_{23}, and S_{31} (color match with corresponding arrows). Dark symbols indicate accessible z-labels, while light crossed-out symbols are not accessible within SU(3) (see text).

c. Symmetry labels for SU(3)

The q-labels for a given IREP within SU(3) are derived from its maximum-weight labels $q_{\mathrm{MW}} \equiv \max\{(z_1, z_2)\}$. With the z-labels additive through tensor products (see latter in the appendix), the z-labels of *arbitrary* multiplets must be integer multiples of the z-labels of the defining representation. This immediately excludes the z-labels (points) in Fig. 10 that are crossed out. In particular, with the columns of Eq. (A31) being linearly dependent, one may therefore use the columns of

$$\tilde{z} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad \text{(A33)}$$

as basis for the maximum weight labels, for consistency with literature.[25,41] Given q_{MW}, the actual label of the multiplet then is determined by

$$q \equiv (q_1, q_2) \equiv \tilde{z}^{-1} \cdot q_{\mathrm{MW}} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} q_{\mathrm{MW}}. \quad \text{(A34)}$$

This prescription makes the q-labels independent of the specific normalization conventions chosen for the z-operators. Furthermore, with $\tilde{z} = z \cdot (1, 0, -1)^T$ and the vectors in the columns in Eq. (A33) being positive by the adopted sorting scheme, this guarantees plain positive integers for the multiplet labels q. These labels also

lie dense, in the sense that any (q_1, q_2) with $q_i \geq 0$ results in a valid multiplet.

The defining representation with $q_{\mathrm{MW}} = (1, 1)$ has the q-label $(1, 0)$. Its weight diagram together with many further examples for multiplets, as generated, in practice, from an actual NRG run using SU(3), are presented in Fig. 12. Note, however, that weight diagrams are mainly a matter of presentation of multiplets, while in practice a listing of z-labels suffices to describe the multiplet space.

5. Decomposition into irreducible representations

The generators of a specific symmetry group $\mathcal{S}$ represent an irreducible finite set of operators $\{\hat{S}_\sigma\}$, assumed to act in the full Hilbert space of a given physical system. Within (small) subspaces of the system, finite dimensional matrix representations can be constructed that obey exactly the same commutator relations as the generators in terms of their structure constants in Eq. (A2). As such, a given matrix representation $\{R_\sigma\}$ inherits all the properties of the generators. In particular, the matrix representation has the same number of operators as $\{\hat{S}_\sigma\}$ with a one-to-one correspondence in the symmetry label σ. Therefore the z-operators as well as the RLOs share exactly the same interpretation within the D-dimensional carrier space of $\{R_\sigma\}$.

Consider some arbitrary matrix representation $\{R_\sigma\}$ that may have emerged, for example, from a tensor product space. As it operates in a D-dimensional carrier space, all of its matrices share the same dimension $D \times D$. Assume a well-defined symmetry eigenstate within this space is available, to be called *seed state*, with a typically easy example being a maximum weight state. Then repeated application of RLOs from the set $\{R_\sigma\}$ generates a (sub)space which eventually describes a full symmetry multiplet, *i.e.* an IREP. By construction this subspace already diagonalizes the z-operators. Thus the z-labels are known, which also provides the q-labels for the multiplet, *e.g.* by simply taking the maximum weight labels, $q_1 = q_{\mathrm{MW}}$.

If this multiplet q_1 with d_{q_1} symmetry eigenstates spans the entire D-dimensional carrier space, then the matrix representation $\{R_\sigma\}$ is already irreducible. If only a subspace of the D-dimensional carrier space was generated, *i.e.* $d_{q_1} < D$, the matrix representation $\{R_\sigma\}$ is reducible. Multiplet q_1 then defines a fully separated space, given the symmetry operations in $\{R_\sigma\}$. Combining the orthonormal state space of multiplet q_1 as columns into a matrix V_1, the matrix representation $\{R_\sigma\}$ can be cast into the space of multiplet q_1, $R_\sigma \rightarrow I^{[q_1]} \equiv V_1^\dagger R_\sigma V_1$, which thus constructs the irreducible matrix representation $I^{[q_1]}$ for IREP q_1.

In case the D-dimensional vector space is not exhausted yet, above procedure can be repeated with another seed state within the remainder of the vector space, generating further irreducible multiplets $q_2, q_3, \ldots$, until the D-dimensional vector space is fully exhausted. By combining the state spaces of the multiplets thus generated, the resulting unitary matrix $V \equiv [V_1, V_2, \ldots]$ allows to block-decompose the original matrix representation $\{R_\sigma\}$ in terms of its irreducible representations,

$$V^\dagger R_\sigma V = \bigoplus_q M_q I_\sigma^{[q]}. \tag{A35}$$

where q runs through all IREPs $I^{[q]}$. Note that a given IREP may be generated multiple times in the decomposition, which is indicated by the *outer multiplicity* $M_q \in [0, 1, 2, \ldots]$. The presence of outer multiplicity therefore refers to the situation that $M_q > 1$ for at least one q in the decomposition. In this case, also inner multiplicity may occur, which refers to non-uniqueness of z-labels within an irreducible multiplet [*cf.* Eq. (A19)], both of which are specifically relevant, for example, for $\mathrm{SU}(N > 2)$ or $\mathrm{Sp}(2n > 2)$.

As seen from above construction, the matrix representation $I^{[q]}$ of IREP q is tightly connected to the symmetry multiplet q. In general, $I^{[q]}$ is unique only up to a global similarity transformation, as this does not affect commutator relations. By using its related multiplet state space, however, this space (i) can be chosen such that it diagonalizes all z-operators, and (ii) can put into a well-defined order as provided, for example, by the lexicographic ordering in the z-labels used to define the maximum weight state. Based on this basis, the matrix representation $I^{[q]}$ can be determined uniquely. This procedure on obtaining unique irreducible matrix representations will be adopted throughout.

The decomposition in Eq. (A35), finally, can be done fully numerically along the same lines as already sketched above. Particular attention, however, must be paid to issues related to inner and outer multiplicity for overall consistency. This will be discussed in more detail in Sec. B 1.

6. Tensor product spaces

Consider two irreducible matrix representations $I^{[q_1]}$ and $I^{[q_2]}$ of some non-abelian symmetry group $\mathcal{S}$, with their matrix elements written in the basis of the symmetry eigenstates $|q_1 q_{1z}\rangle$ and $|q_2 q_{2z}\rangle$ of the two IREPs q_1 and q_2, respectively. The two multiplets are assumed to live in different spaces, so they can be joined through a tensor product, *i.e.* $|q_1 q_{1z}\rangle|q_2 q_{2z}\rangle \equiv |q_1 q_{1z}; q_2 q_{2z}\rangle$. Then the generators of the symmetry in the combined space are defined in an *additive* fashion, which derives from the origin of the generators in infinitesimal symmetry operations, *cf.* Eq. (A1),

$$R_\sigma^{\mathrm{tot}} \equiv I_\sigma^{[q_1]} \otimes \mathbf{1}^{[q_2]} + \mathbf{1}^{[q_1]} \otimes I_\sigma^{[q_2]}. \tag{A36}$$

Note that the additivity of the symmetry generators *directly* also implies the additivity of z-labels for non-abelian symmetries in general. And even if the non-abelian part of the SU(N) symmetry is broken, *e.g.* re-

duced to an abelian symmetry with quantum labels q_z, these are, of course, still additive.

By construction, the tensor product representation $\{R_\sigma^{\text{tot}}\}$ in Eq. (A36) is also a representation of the symmetry, as it obeys the same commutator relations within the combined system as the IREPs $I_\sigma^{[q_i]}$ within their individual space,

$$[R_\sigma^{\text{tot}}, R_{\sigma'}^{\text{tot}}] = \underbrace{[R_\sigma^{[q_1]}, R_{\sigma'}^{[q_1]}]}_{=\sum_\mu f_{\sigma\sigma'\mu} R_\mu^{[q_1]}} \otimes \mathbf{1}^{[q_2]} + \mathbf{1}^{[q_1]} \otimes \underbrace{[R_\sigma^{[q_2]}, R_{\sigma'}^{[q_2]}]}_{=\sum_\mu f_{\sigma\sigma'\mu} R_\mu^{[q_2]}}$$
$$= \sum_\mu f_{\sigma\sigma'\mu} R_\mu^{\text{tot}}. \tag{A37}$$

The product representation $\{R_\sigma^{\text{tot}}\}$, however, is typically reducible. The resulting decomposition into IREPs is done exactly the same way as in Eq. (A35).

The unitary transformation that rotates the product space $|q_1 q_{1z}; q_2 q_{2z}\rangle$ into its combined symmetry multiplets $|q q_z; (\alpha)\rangle$ is given by the Clebsch-Gordan coefficients (CGC),

$$|q q_z; (\alpha)\rangle = \sum_{q_{1z}, q_{2z}} |q_1 q_{1z}; q_2 q_{2z}\rangle \underbrace{\langle q_1 q_{1z}; q_2 q_{2z} | q q_z; (\alpha)\rangle}_{\equiv C_{q_{1z} q_z; (\alpha)}^{[q_{2z}]}},$$
$$\tag{A38}$$

with the shorthand notation $C_{q_{1z} q_z; (\alpha)}^{[q_{2z}]}$ for CGCs, consistent with the MPS tensors in the main body of the paper [cf. Eq. (10)]. Note that the CGCs implicitly also carry the multiplet labels q_1, q_2, and q. The index α has been added to account for possible outer-multiplicity [cf. Eq. A35], in that for input multiplets q_1 and q_2 the same output multiplet q can appear $M_q^{[q_1, q_2]}$ times, therefore $\alpha = 1, \ldots, M_q^{[q_1, q_2]}$ for a given q. If outer multiplicity is absent, the index α can be omitted, hence the round brackets around α in Eq. (A38) or Eq. (A39).

As outer multiplicity also refers to *different* multiplets and hence state spaces, the Clebsch-Gordan coefficients, reflecting a unitary transformation, obey the general orthogonality condition,

$$\sum_{q_{1z} q_{2z}} C_{q_{1z} q_z; (\alpha)}^{[q_{2z}]} C_{q_{1z} q_z'; (\alpha')}^{[q_{2z}]} = \delta_{q_z, q_z'} \cdot (\delta_{\alpha, \alpha'}). \tag{A39}$$

This holds within the same output multiplet q, whereas the overlap between different output multiplets is strictly zero. While outer multiplicity is intrinsically connected to the underlying symmetry and hence to CGCs, in addition, this also affects the output multiplet space which must accommodate the additional multiplets [e.g. see QSpace discussion in the main text; note also that Eq. (A39) is completely analogous in structure to the orthogonality relation of A-tensors as in Eq. (11)].

7. Irreducible operator sets and Wigner-Eckart theorem

Consider a set of generators $\{\hat{S}_\sigma\}$ of some symmetry group $\mathcal{S}$ that a Hamiltonian $\hat{H}$ commutes with. Then all energy eigenstates of the Hamiltonian can be categorized with well-defined quantum-labels, as indicated in Eq. (A17). In order to maintain an effective book keeping of the quantum labels when calculating matrix elements of operators, it must be possible to similarly categorize the operators themselves. Typically, the operators of interest are closely related to the Hamiltonian, *i.e.* consist of operators that also appear in the Hamiltonian or are composites thereof, such as creation, annihilation, occupation, spin operators, etc. Since the Hamiltonian can be properly constructed within the given symmetry, so can be its constituents.

An *irreducible operator set* is constructed in a completely analogous fashion as an irreducible state space, with an explicit example already derived for the spin operator in Eq. (A11) using Eq. (A10). Consider the generic setup of a set of generators $\{\hat{S}_\sigma\}$ including RLOs. Then irreducible state multiplets can be generated through iterative application of these operators,

$$\hat{S}_\sigma |q q_z\rangle = s_{q_z q_z'}^{[q\sigma]} |q q_z'\rangle, \tag{A40}$$

as in Eq. (A24), while ignoring inner multiplicity for the sake of the argument and having dropped the energy multiplet index n for simplicity. Given an operator $\hat{F}$, on the other hand, its transformation according to a symmetry is fully reflected in its *commutator relations* with the generators of the symmetry. This is easily motivated through infinitesimal symmetry operations as in Eq. (A3b). The commutator relations, on the other hand, also emerge naturally when analyzing the effect of a generator of the symmetry acting onto a symmetry state $|q q_z\rangle$ that already also has the operator $\hat{F}$ applied to it,

$$\hat{S}_\sigma \cdot \hat{F} |q q_z'\rangle = [\hat{S}_\sigma, \hat{F}] |q q_z'\rangle + \hat{F} \cdot \hat{S}_\sigma |q q_z'\rangle. \tag{A41}$$

The second term on *r.h.s.* clearly describes the symmetry properties of the state $|q q_z\rangle$, while the first term yields the transformation properties of the operator $\hat{F}$ which are independent of the carrier space. This is similar to what has already been seen in Eq. (A21) for the combined action of two generators.

Now *iff* an operator set $\hat{F}^q \equiv \{\hat{F}_{q_z}^q\}$ transforms exactly the same way as the state space of IREP q in Eq. (A40), that is

$$[\hat{S}_\sigma, \hat{F}_{q_z}^q] = s_{q_z q_z'}^{[q\sigma]} \hat{F}_{q_z'}^q, \tag{A42}$$

then the operator set $\hat{F}^q$ is called an *irreducible operator* (IROP) set that transforms like the multiplet q. It carries the symmetry labels (q, q_z) the same way as an irreducible state multiplet does.

a. IROP decomposition

In the case that a specific member of the IROP set is already known, then Eq. (A42) allows to generate the full IROP set exactly the same way as a state multiplet can be generated. Again, the maximum weight label determined the multiplet q that the IROP represents. This was exactly the procedure adopted, for example, to obtain the spin operator in Eq. (A11). Furthermore, Eq. (A11) also serves as a simple demonstration that the space of generators itself clearly also can be cast into a single IROP. The corresponding multiplet is called the *regular representation* then.

More generally, it is instructive to realize that irreducible operator sets (IROPs) and symmetry multiplets (IREPs) can be treated on a nearly equal footing. In particular, the notion of proper orthonormalization of state spaces can be directly applied also to IROP sets, up to a global normalization factor. This is motivated by the observation, that given a scalar multiplet $|0\rangle$ for which $|F_{q_z'}^q\rangle \equiv F_{q_z'}^q |0\rangle \neq 0$, *i.e.* does not vanish, then $|F_{q_z'}^q\rangle$ represents the multiplet *vector space* for IREP q. With proper overall normalization of the IROP F^q, it follows

$$\delta_{q_z,q_z'} = \langle F_{q_z}^q | F_{q_z'}^q \rangle = \langle 0 | F_{q_z}^{q\dagger} F_{q_z'}^q | 0\rangle.$$

The last equation also holds, if the scalar multiplet $|0\rangle$ is replaced by an arbitrary other symmetry eigenstate $|qq_z\rangle$. For matrix representations of IROPs and operators more generally, this motivates the scalar or inner product for two matrices as in Eq. (A14). Thus equipped with scalar product and norm for matrices, an IROP decomposition can be done *exactly* the same way as the multiplet decompositions for symmetry multiplets starting from a specific symmetry eigenstate (IROP component). This is important, in particular, in the presence of inner multiplicity in the multiplet of an IROP for consistency with the Wigner-Eckart theorem.

b. Wigner Eckart theorem

It follows from Eqs. (A41) and (A42), that the states resulting from the IROP $\{\hat{F}_{q_{1z}}^{q_1}\}$ applied to a multiplet $|q_2 q_{2z}\rangle$,

$$\hat{S}_\sigma \cdot \hat{F}_{q_{1z}}^{q_1} |q_2 q_{2z}\rangle$$
$$= \underbrace{[\hat{S}_\sigma, \hat{F}_{q_{1z}}^{q_1}]}_{=s_{q_{1z}q_{1z}'}^{[q_1\sigma]}\hat{F}_{q_{1z}'}^{q_1}} |q_2 q_{2z}\rangle + \hat{F}_{q_{1z}}^{q_1} \cdot \underbrace{\hat{S}_\sigma |q_2 q_{2z}\rangle}_{=s_{q_{2z}q_{2z}'}^{[q_2\sigma]}|q_2 q_{2z}'\rangle} ,$$

transforms exactly the same way under given symmetry as a tensor product of two state multiplets,

$$\hat{S}_\sigma \cdot |qq_z\rangle_1 |qq_z\rangle_2$$
$$= \underbrace{\hat{S}_{1\sigma}|qq_z\rangle_1}_{=s_{q_{1z}q_{1z}'}^{[q_1\sigma]}|qq_z'\rangle_1} \otimes |qq_z\rangle_2 + |qq_z\rangle_1 \otimes \underbrace{\hat{S}_{2\sigma}|qq_z\rangle_2}_{=s_{q_{2z}q_{2z}'}^{[q_2\sigma]}|qq_z'\rangle_2} ,$$

using Eqs. (A36) and (A40). Therefore the action of an IROP $\hat{F}^{q_1}$ onto the state space of an IREP q_2 shares exactly the same algebraic structure in terms of symmetries like the product space of the two multiplets q_1 and q_2.

This motivates the Wigner-Eckart theorem. With the definition of the Clebsch-Gordan coefficients in Eq. (A38), it is thus clear that up to scalar factors depending on the normalization of the operator set, the *same* CGCs also apply for the state space decomposition arising out of $\hat{F}_{q_{1z}}^{q_1}|q_2 q_{2z}\rangle$. In particular, it follows for the matrix elements of the operator *w.r.t.* a given state space,

$$\langle qq_z; (\alpha) | \hat{F}_{q_{1z}}^{q_1} | q_2 q_{2z}\rangle \equiv \langle qq_z; (\alpha)| \cdot \left(\hat{F}_{q_{1z}}^{q_1} | q_2 q_{2z}\rangle \right)$$
$$= \langle q; (\alpha) \| \hat{F}^{q_1} \| q_2 \rangle \cdot C_{q_{1z}q_z;(\alpha)}^{[q_{2z}]}, \tag{A43}$$

where, again, α accounts for possible outer multiplicity. $\langle q; (\alpha) \| \hat{A}_{q_1} \| q_2 \rangle$ is called the *reduced matrix element*. It is entirely independent of the z-labels, *i.e.* the internal structure of the IREPs q_1, q_2, and q.

The first line in Eq. (A43) specifies the adopted convention for matrix elements given the Wigner-Eckart theorem: the operator is acting to the *right* ket-state, the symmetry labels of which are combined. The resulting object is contracted with the bra-states. This is important for consistency, since the IROP $\hat{F}^q$ is subtly different from the IROP $(\hat{F}^\dagger)^q$. Therefore one *must* be careful with expressing a matrix element through $\langle qq_z | \hat{F} | q_2 q_{2z}\rangle = \langle q_2 q_{2z} | \hat{F}^\dagger | qq_z\rangle^*$. Even though usually $(\hat{F}_{q_z}^q)^\dagger \sim (\hat{F}^\dagger)_{-q_z}^q$, further signs may bee needed to ensure for consistency within the Clebsch-Gordan coefficients [*e.g.* see discussion around Eq. A49 later].

8. Several independent symmetries

A physical system often exhibits several symmetries. Each of the $\lambda = 1, \ldots, n_S$ symmetries $\mathcal{S}^\lambda$ is completely described by its own set of generators $\{\hat{S}_\sigma^\lambda\}$. As these symmetries act independently of each other, this implies that their generators must commute,

$$[\hat{S}_\sigma^\lambda, \hat{S}_{\sigma'}^{\lambda'}] = 0 \qquad \text{for } \lambda \neq \lambda'. \tag{A44}$$

This allows to assign independent quantum labels $(q^\lambda q_z^\lambda)$ with respect to each individual symmetry [*cf.* discussion following Eq. (1) in the main paper]. On the multiplet level, the symmetries are given compactly by the combined q-labels, $q \equiv (q^1, q^2, \ldots, q^{n_S})$, while similarly their *z-labels* are given by $q_z \equiv (q_z^1, q_z^2, \ldots, q_z^{n_S})$. Here the elementary multiplet labels q^λ and q_z^λ can already consist of a set of labels themselves, the number of which is determined by the rank r of the respective symmetry $\mathcal{S}^\lambda$ [*cf.* Eq. (A18)].

When a non-abelian symmetry is broken, it it is reduced to simpler subalgebras. In particular, it may be reduced to its abelian core of z-operators (Cartan subalgebra). For example, consider the rotational spin SU(2)

P7.

symmetry. This symmetry can be broken by applying a magnetic field. The system still maintains, however, a continuous rotational symmetry around the axis of the magnetic field, leaving the q_z symmetry intact, while the multiplet label q becomes irrelevant. Similarly, if particle-hole symmetry (see later) is broken, only the abelian quantum number of total charge [*i.e.* the z-label] remains.

Abelian symmetries therefore fit seamlessly into the general non-abelian framework outlined in this paper. With the multiplet label q irrelevant, the q_z are promoted to the status of a q-label, instead, with no z-labels remaining [with all multiplets being one-dimensional, the z-labels are no longer required, *i.e.* can be set to zero, for simplicity]. This then allows to write the abelian symmetry in terms of trivial scalar Clebsch-Gordan coefficients. The latter, nevertheless, are important as they account for the proper addition rules *w.r.t.* the abelian z-labels,

$$\langle q_{1(z)} q_{2(z)} | q_{(z)} \rangle = \delta_{q_{(z)}, q_{1(z)} + q_{2(z)}}. \tag{A45}$$

9. Symmetries in physical systems

In the following, several examples of symmetries in simple physical systems will be given, with the associated spinors and irreducible operator sets explained in detail. In particular, this concerns fermionic systems with spin or particle-hole symmetry.

For the model Hamiltonians in strongly correlated electron systems, correlation through interaction plays an important role, while the terms describing interaction typically preserve certain underlying global symmetries. Since the arguments of demonstrating symmetries of a specific Hamiltonian, however, are rather similar, in general, it suffices to consider a simple non-interacting Hamiltonian. Simple issues related to interactions are discussed with Eq. (A58) below.

For simplicity, therefore much of the following discussion will be exemplified in terms of the Hamiltonian of a plain spinful fermionic tight-binding chain,

$$\hat{H} = \sum_k t_k \underbrace{\sum_\sigma (\hat{c}^\dagger_{k\sigma} \hat{c}_{k+1,\sigma} + \text{H.c.})}_{\equiv \hat{h}_{k,k+1}}, \tag{A46}$$

where $\hat{c}^\dagger_{k\sigma}$ creates a particle at site k with spin $\sigma \in \{\uparrow, \downarrow\}$. The Hamiltonian in Eq. (A46) has spin-independent hopping amplitudes t_k, hence possesses spin-SU(2) symmetry, SU(2)$_{\text{spin}}$ in short. Furthermore, it is particle-hole symmetric, implying particle-hole SU(2) symmetry, also called charge-SU(2) symmetry, or SU(2)$_{\text{charge}}$ in short.

a. SU(2) *spin symmetry*

Using the two-dimensional spinor

$$\hat{\psi}_{S,k} \equiv \begin{pmatrix} \hat{c}_{k\uparrow} \\ \hat{c}_{k\downarrow} \end{pmatrix} \tag{A47}$$

for each site k, the Hamiltonian in Eq. (A46) can be rewritten as

$$\hat{H} = \sum_k t_k \big(\hat{\psi}^\dagger_{S,k} \cdot \hat{\psi}_{S,k+1} + \text{H.c.} \big), \tag{A48}$$

where the sum over σ was incorporated in the scalar product of the vector of operators in $\hat{\psi}_{S,k}$. Clearly, the two-dimensional scalar product is invariant under an arbitrary unitary two-dimensional transformation U, *i.e.* $\psi^\dagger_k \psi_{k+1} = (U\psi_k)^\dagger (U\psi_{k+1})$, thus exhibiting spin-SU(2) symmetry. The spinor in Eq. (A47) is defined in a site specific manner. When concentrating on a single site, therefore the index k can be dropped for convenience.

The generators of spin-SU(2) symmetry are constructed in terms of the two-dimensional defining representation of $\{S_\sigma\} \equiv \{S_+, S_z, S_-\}$ [*cf.* Eq. (A7)]. These can be written as operators (distinguished by the hat) through second quantization in the full Hilbert space,

$$\hat{S}_\sigma = \hat{\psi}^\dagger_S S_\sigma \hat{\psi}_S,$$

which up to prefactors leads to the spin IROP $\hat{S}^1 \equiv \{-\frac{1}{\sqrt{2}} \hat{S}_+; \hat{S}_z; +\frac{1}{\sqrt{2}} \hat{S}_-\}$, already derived in Eq. (A11). The raising operator, for example, is given by

$$\hat{S}_+ = \hat{\psi}^\dagger_S \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{\psi}_S = \hat{c}^\dagger_\uparrow \hat{c}_\downarrow,$$

which flips a down-spin to an up-spin for given site. Similarly, the z-operator is given by

$$\hat{S}_z = \hat{\psi}^\dagger (\tfrac{1}{2} \tau_z) \hat{\psi} = \tfrac{1}{2} (\hat{c}^\dagger_\uparrow \hat{c}_\uparrow - \hat{c}^\dagger_\downarrow \hat{c}_\downarrow) \equiv \tfrac{1}{2} (\hat{n}_\uparrow - \hat{n}_\downarrow).$$

Furthermore, $[\hat{S}_+, \hat{c}^\dagger_\downarrow] = [\hat{c}^\dagger_\uparrow \hat{c}_\downarrow, \hat{c}^\dagger_\downarrow] = \hat{c}^\dagger_\uparrow$ shows that the spinor $\hat{\psi}^\dagger_S$ already represents an IROP for the $q = \frac{1}{2}$ multiplet of SU(2)$_{\text{spin}}$,

$$(\hat{\psi}^\dagger_S)^{[1/2]} = \begin{pmatrix} \hat{c}^\dagger_\uparrow \\ \hat{c}^\dagger_\downarrow \end{pmatrix}. \tag{A49a}$$

This is already properly sorted *w.r.t.* z-labels, in that the second component correspond to the lower $q_z = -\frac{1}{2}$ element of the multiplet, since $[\hat{S}_z, \hat{c}^\dagger_\downarrow] = (-\frac{1}{2}) \cdot \hat{c}^\dagger_\downarrow$.

In contrast, the IROP for the spinor $\hat{\psi}_S$, *i.e. without* the dagger, is similar, yet has subtle differences. In particular, with $[\hat{S}_+, \hat{c}_\uparrow] = [\hat{c}^\dagger_\uparrow c_\downarrow, \hat{c}_\uparrow] = -\hat{c}_\downarrow$, the role of spin within the multiplet is reversed, *i.e.* $q_z \to -q_z$, while also an additional sign is acquired,

$$(\hat{\psi}_S)^{[1/2]} = \begin{pmatrix} -\hat{c}_\downarrow \\ \hat{c}_\uparrow \end{pmatrix}. \tag{A49b}$$

This extra sign is important in context of the Wigner-Eckart theorem in Eq. (A43), where the particular order of first applying, *i.e.* combining an operator with the ket-state is directly related to the order in the Clebsch-Gordan coefficients. This is convention, of course, but consistency is crux.

In terms of the proper IROPs in Eqs. (A49), finally, the Hamiltonian in Eq. (A46) can be written in either IROP while, however, one must not mix them,

$$\hat{H} = \sum_k t_k \left(\left[(\hat{\psi}_S)^{[1/2]} \right]^\dagger \cdot (\hat{\psi}_S)^{[1/2]} + \text{H.c.} \right) \quad (\text{A50a})$$

$$= \sum_k t_k \left((\hat{\psi}_S^\dagger)^{[1/2]} \cdot \left[(\hat{\psi}_S^\dagger)^{[1/2]} \right]^\dagger + \text{H.c.} \right). \quad (\text{A50b})$$

The second line is essentially the same as the spinor expression in Eq. (A48), yet with the difference, that here the underlying IROP structure has been pointed out explicitly.

b. SU(2) *particle-hole symmetry for spinful system*

The particle-hole symmetry $\text{SU}(2)_{\text{charge}}$ of the Hamiltonian in Eq. (A46) can be made apparent in a similar way as for the spin symmetry above. Consider the spinor in the charge sector,

$$\hat{\psi}_{C,k\sigma} \equiv \begin{pmatrix} \hat{c}_{k\sigma} \\ s_k \hat{c}_{k,-\sigma}^\dagger \end{pmatrix}$$

with alternating phases $s_k = (-1)^k$ along the chain in Eq. (A46). Again, the Hamiltonian can be written as sum over *scalar products* in the spinors,

$$\sum_\sigma \hat{\psi}_{C,k\sigma}^\dagger \cdot \hat{\psi}_{C,k+1,\sigma} = \sum_\sigma \left(\hat{c}_{k\sigma}^\dagger \hat{c}_{k+1,\sigma} - \hat{c}_{k,-\sigma} \hat{c}_{k+1,-\sigma}^\dagger \right)$$

$$= \hat{h}_{k,k+1},$$

suggesting another underlying SU(2) symmetry. Note that the alternating sign s_k is crucial to recover the correct hopping structure in Eq. (A46). Given the spinor in the charge sector, the raising operator becomes

$$\hat{\psi}_{C,k\sigma}^\dagger \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{\psi}_{C,k\sigma} = s_k \hat{c}_{k\sigma}^\dagger \hat{c}_{k,-\sigma}^\dagger$$

which, up to a sign, is the *same* for both spins. It is therefore sufficient in the charge sector to consider a spinor for one specific σ in $\hat{\psi}_{C,k\sigma}$ only. Therefore, again concentrating on a single site and hence dropping the site index k, now with fixed $\sigma = \uparrow$, the spinor in the charge sector is given by,

$$\hat{\psi}_C \equiv \begin{pmatrix} \hat{c}_\uparrow \\ s\hat{c}_\downarrow^\dagger \end{pmatrix}. \quad (\text{A51})$$

The associated raising operator becomes

$$\hat{C}_+ = s\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger, \quad (\text{A52})$$

which now creates a pair of particles with opposite spin, while the z-operator is

$$\hat{C}_z = \hat{\psi}_C^\dagger (\tfrac{1}{2}\tau_z) \hat{\psi}_C = \tfrac{1}{2}(\hat{c}_\uparrow^\dagger \hat{c}_\uparrow - \hat{c}_\downarrow \hat{c}_\downarrow^\dagger)$$

$$\equiv \tfrac{1}{2}(\hat{n}_\uparrow + \hat{n}_\downarrow - 1). \quad (\text{A53})$$

With $\hat{n} \equiv \hat{n}_\uparrow + \hat{n}_\downarrow$, the z-operator $\hat{C}_z$ counts the total charge on given fermionic site relative to half-filling. With

$$[\hat{C}_+, \hat{c}_\uparrow] = [s\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger, \hat{c}_\uparrow] = -s\hat{c}_\downarrow^\dagger \quad (\text{A54a})$$

$$[\hat{C}_+, \hat{c}_\downarrow] = [s\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger, \hat{c}_\downarrow] = s\hat{c}_\uparrow^\dagger, \quad (\text{A54b})$$

this allows to construct the $q = \frac{1}{2}$ IROPs for $\text{SU}(2)_{\text{charge}}$,

$$(\hat{\psi}_C)^{[1/2]} = \begin{pmatrix} s\hat{c}_\downarrow^\dagger \\ -\hat{c}_\uparrow \end{pmatrix} \quad (\text{A55a})$$

$$(\hat{\psi}_C^\dagger)^{[1/2]} = \begin{pmatrix} s\hat{c}_\uparrow^\dagger \\ \hat{c}_\downarrow \end{pmatrix}, \quad (\text{A55b})$$

again associating the lower component with the $q_z = -\frac{1}{2}$ element of the $q = (1/2)$ multiplet [*cf.* Eqs. (A49)]. An irrelevant overall minus sign has been applied to the spinor in Eq. (A55a) for later convenience. With this, the hopping term in the Hamiltonian in Eq. (A46) can be rewritten in terms of the scalar products

$$\hat{h}_{k,k+1} = \left[(\hat{\psi}_{Ck})^{[1/2]} \right]^\dagger \cdot (\hat{\psi}_{C,k+1})^{[1/2]}$$

$$+ \left[(\hat{\psi}_{Ck}^\dagger)^{[1/2]} \right]^\dagger \cdot (\hat{\psi}_{C,k+1}^\dagger)^{[1/2]} \quad (\text{A56})$$

The spinors in the charge sector do mix spin components, which essentially also requires full spin symmetry [see later discussion of symplectic group $\text{Sp}(2m)$ in Sec. A 10]. More importantly, the construction of the $\text{SU}(2)_{\text{charge}}$ symmetry allows it to fully commute with the spin-SU(2) symmetry introduced earlier,

$$[\hat{S}_z, \hat{C}_z] = \tfrac{1}{4}[\hat{n}_\uparrow - \hat{n}_\downarrow, \hat{n} - 1] = 0$$

$$[\hat{S}_z, \hat{C}_+] = \tfrac{s}{2}\underbrace{[\hat{c}_\downarrow^\dagger \hat{c}_\uparrow, \hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger]}_{=\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger} - \tfrac{s}{2}\underbrace{[\hat{c}_\downarrow^\dagger \hat{c}_\downarrow, \hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger]}_{=\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger} = 0$$

$$[\hat{S}_+, \hat{C}_z] = \tfrac{1}{2}\underbrace{[\hat{c}_\uparrow^\dagger \hat{c}_\downarrow, \hat{c}_\uparrow^\dagger \hat{c}_\uparrow]}_{=-\hat{c}_\uparrow^\dagger \hat{c}_\downarrow} + \tfrac{1}{2}\underbrace{[\hat{c}_\uparrow^\dagger \hat{c}_\downarrow, \hat{c}_\downarrow^\dagger \hat{c}_\downarrow]}_{=\hat{c}_\uparrow^\dagger \hat{c}_\downarrow} = 0$$

$$[\hat{S}_+, \hat{C}_+] = s[\hat{c}_\uparrow^\dagger \hat{c}_\downarrow, \hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger] = s\hat{c}_\uparrow^\dagger \hat{c}_\uparrow^\dagger = 0. \quad (\text{A57})$$

That is, the two symmetries act completely independent of each other and thus can coexist simultaneously, written as the overall symmetry $\text{SU}(2)_{\text{spin}} \otimes \text{SU}(2)_{\text{charge}}$.

If interactions are present in the system, such as local Coulomb interaction $U\hat{n}_\uparrow \hat{n}_\downarrow$, then the particle-hole symmetric regime requires a specific altered onsite energy

P7.

relative to the chemical potential. With $\hat{n}_\sigma^2 = \hat{n}_\sigma$, and $\hat{n} \equiv \hat{n}_\uparrow + \hat{n}_\downarrow$, it follows that $\hat{n}_\uparrow \hat{n}_\downarrow = \frac{1}{2}(\hat{n}-1)^2 + \frac{1}{2}(\hat{n}-1)$, and therefore

$$\varepsilon_d \hat{n} + U\hat{n}_\uparrow \hat{n}_\downarrow = (\varepsilon_d + \tfrac{U}{2})\underbrace{(\hat{n}-1)}_{=\hat{C}_z} + \tfrac{U}{2}(\hat{n}-1)^2 + \text{const.}$$

$$(A58)$$

The first term on the *r.h.s.* is proportional to the $\hat{C}_z$ operator, which thus acts like a magnetic field for $\mathrm{SU(2)_{spin}}$. Therefore for full particle-hole symmetry to hold, this term must be zero, which requires $\varepsilon_d = -\frac{U}{2}$. In particular, in the absence of interaction, this implies $\varepsilon_d = 0$. The actual interaction term, *i.e.* the second term on the *r.h.s.* in Eq. (A58), also resembles $\hat{C}_z$. Yet it is quadratic, and for this it also holds, $\hat{C}_z^2 = \hat{C}_x^2 = \hat{C}_y^2$. Therefore, this term can actually be written as $\hat{C}^2$ which itself, like spin $\hat{S}^2$ for $\mathrm{SU(2)_{spin}}$, represents the Casimir operator for $\mathrm{SU(2)}$, and thus is compatible with $\mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)_{charge}}$.

The actual IROP for particle creation and annihilation given $\mathrm{SU(2)_{SC}^{\otimes 2}} \equiv \mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)_{charge}}$ symmetry can be generated using above symmetry operations. This generates a four-dimensional spinor. As it turns out, the resulting IROP is the combination of the two IROPs generated in the spin symmetric case in Eqs. (A49) as well as in the particle-hole symmetric case in Eqs. (A55),

$$\hat{\psi}_{\mathrm{CS}}^{\left[\frac{1}{2},\frac{1}{2}\right]} \equiv \begin{pmatrix} s\hat{c}_\uparrow^\dagger \\ \hat{c}_\downarrow \\ s\hat{c}_\downarrow^\dagger \\ -\hat{c}_\uparrow \end{pmatrix}. \qquad (A59)$$

The multiplet labels $\left[\frac{1}{2}, \frac{1}{2}\right]$ will be derived with Eq. (A61) below. The signs for the individual components in above IROP have been properly adjusted, considering that the raising operator in the charge sector itself, Eq. (A52), carries the alternating sign $s_{(k)} = (-1)^k$. For example, commuting $\hat{C}_+$ onto the fourth component, yields the third component of the spinor $\hat{\psi}_{\mathrm{CS}}$ [*cf.* Eq. (A54a)], while commuting $\hat{S}_+$ onto the third component yields the first component, and so on. Again, keeping track of the alternating sign $s_k = (-1)^k$ is crucial to recover the hopping structure in Eq. (A46),

$$\begin{aligned}
&\left(\hat{\psi}_{\mathrm{CS},k}\right)^\dagger \cdot \hat{\psi}_{\mathrm{CS},k+1} \\
&= -\underbrace{\hat{c}_{k\uparrow}\hat{c}_{k+1,\uparrow}^\dagger}_{=-\hat{c}_{k+1,\uparrow}^\dagger \hat{c}_{k\uparrow}} + \hat{c}_{k\downarrow}^\dagger \hat{c}_{k+1,\downarrow} - \underbrace{\hat{c}_{k\downarrow}\hat{c}_{k+1,\downarrow}^\dagger}_{=-\hat{c}_{k+1,\downarrow}^\dagger \hat{c}_{k\downarrow}} + \hat{c}_{k\uparrow}^\dagger \hat{c}_{k+1,\uparrow} \\
&= \hat{h}_{k,k+1},
\end{aligned} \qquad (A60)$$

The full tight-binding Hamiltonian simply becomes $\hat{H} = \sum_k t_k (\hat{\psi}_{\mathrm{CS},k})^\dagger \hat{\psi}_{\mathrm{CS},k+1}$ where the hermitian conjugate term has been incorporated already in the spinor structure. This also reflects the irrelevance of taking the hermitian conjugate version of the IROP in Eq. (A59) as this results in essentially the same object after properly reordering of its components and taking care of signs.

With Eq. (A60) being a scalar product in a four-dimensional spinor space, one may be tempted to think that a plain tight binding chain actually has a non-abelian symmetry with a defining representation of dimension 4. This cannot be the symmetry $\mathrm{SU(4)}$, however, since $\mathrm{SU(4)}$ has rank-3 and thus requires *three* commuting abelian z-operators. The symmetries discussed here, however, only have *two* abelian z-operators, namely total spin and total charge. The symmetry that appears compatible with this scenario, at second glance, is the symplectic symmetry $\mathrm{Sp(4)}$ [see Sec. A 10 below]. Nevertheless, even the latter can be excluded, since raising and lowering operators are severely constrained by the fact that the creation and annihilation operators appear in *pairs* for the *same* fermionic particle in the IROP of Eq. (A59). Consequently, quadratic operators of the type $(\hat{c}_\sigma)^\dagger \hat{c}_\sigma^\dagger = (\hat{c}_\sigma^\dagger)^\dagger \hat{c}_\sigma = 0$ are immediately excluded. With this, the symmetry of the spinor in Eq. (A59) has to remain the product of two symmetries, *i.e.* $\mathrm{SU(2)_{SC}^{\otimes 2}} \equiv \mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)_{charge}}$.

Having determined the IROP $\hat{\psi}_{\mathrm{CS}}$ in Eq. (A59) by repeated application of RLOs $\hat{S}_\pm$ and $\hat{C}_\pm$, the z-labels for each of the four components, on the other hand, can be determined through the z-operators $\hat{C}_z \equiv \frac{1}{2}(\hat{c}_\uparrow^\dagger \hat{c}_\uparrow - \hat{c}_\downarrow \hat{c}_\downarrow^\dagger)$ and $\hat{S}_z \equiv \frac{1}{2}(\hat{c}_\uparrow^\dagger \hat{c}_\uparrow - \hat{c}_\downarrow^\dagger \hat{c}_\downarrow)$, resulting in the z-labels $q_z \equiv (C_z, S_z)$, respectively. The results are summarized in the following table.

[z-operator, IROP component]	(C_z, S_z)
$[\hat{C}_z, s\hat{c}_\uparrow^\dagger] = +\frac{1}{2}(s\hat{c}_\uparrow^\dagger)$ $[\hat{S}_z, s\hat{c}_\uparrow^\dagger] = +\frac{1}{2}(s\hat{c}_\uparrow^\dagger)$	$(+\frac{1}{2}, +\frac{1}{2})$
$[\hat{C}_z, \hat{c}_\downarrow] = -\frac{1}{2}(\hat{c}_\downarrow)$ $[\hat{S}_z, \hat{c}_\downarrow] = +\frac{1}{2}(\hat{c}_\downarrow)$	$(-\frac{1}{2}, +\frac{1}{2})$
$[\hat{C}_z, s\hat{c}_\downarrow^\dagger] = +\frac{1}{2}(s\hat{c}_\downarrow^\dagger)$ $[\hat{S}_z, s\hat{c}_\downarrow^\dagger] = -\frac{1}{2}(s\hat{c}_\downarrow^\dagger)$	$(+\frac{1}{2}, -\frac{1}{2})$
$[\hat{C}_z, -\hat{c}_\uparrow] = -\frac{1}{2}(-\hat{c}_\uparrow)$ $[\hat{S}_z, -\hat{c}_\uparrow] = -\frac{1}{2}(-\hat{c}_\uparrow)$	$(-\frac{1}{2}, -\frac{1}{2})$

$$(A61)$$

These z-labels demonstrate that both the charge and spin multiplet contained in $\hat{\psi}_{\mathrm{CS}}$ corresponds to a $q = \frac{1}{2}$ multiplet. The maximum weight state has the z-labels $(\frac{1}{2}, \frac{1}{2})$, which thus labels the spinor, as was already indicated in Eq. (A59).

Similarly, the local state space of a fermionic site must be organized consistent with the $\mathrm{SU(2)_{spin}} \otimes \mathrm{SU(2)_{charge}}$ symmetry above. The local state space consists of the empty state $|0\rangle$, the singly occupied states $|\uparrow\rangle$ and $|\downarrow\rangle$, and the doubly occupied state $s |\uparrow\downarrow\rangle$. Note that the sign in the last state is crucial, as it is generated by the raising

P7.

operator $\hat{C}_+$ acting on the empty state $|0\rangle$. In summary,

$$
\begin{array}{l|cc|cc}
\text{state} & C & C_z & S & S_z \\
\hline
|0\rangle & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\
|\uparrow\rangle \equiv \hat{c}_\uparrow^\dagger|0\rangle & 0 & 0 & \frac{1}{2} & +\frac{1}{2} \\
|\downarrow\rangle \equiv \hat{c}_\downarrow^\dagger|0\rangle & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\
s\,|\uparrow\downarrow\rangle \equiv s\hat{c}_\uparrow^\dagger\hat{c}_\downarrow^\dagger|0\rangle & \frac{1}{2} & +\frac{1}{2} & 0 & 0.
\end{array}
\tag{A62}
$$

Therefore the local four-dimensional state space of fermionic site is spanned by the two multiplets, $q \equiv (C,S) \in \{(0,\frac{1}{2}),(\frac{1}{2},0)\}$. If particle-hole symmetry is broken yet particle number still preserved, then $2C_z$ from the middle column describes the total number of particles relative to half-filling [*cf.* Eq. (A53)].

With the ordering convention of state labels being $|C,C_z;S,S_z\rangle$ and $\sigma \in \{\uparrow,\downarrow\} \equiv \{+1,-1\}$, the non-zero matrix elements of the 4-component spinor in Eq. (A59) can be calculated. For example,

$$
\begin{aligned}
+s\sigma &= \langle -\sigma|\hat{c}_\sigma \cdot s|\uparrow\downarrow\rangle = s\langle\uparrow\downarrow\,|s\cdot s\hat{c}_\sigma^\dagger\cdot|-\sigma\rangle^{(*)} \\
&\equiv s\langle\tfrac{1}{2},\tfrac{+1}{2};0,0|\hat{\psi}_{(\frac{1}{2}\frac{+1}{2};\frac{1}{2}\frac{\sigma}{2})}|0,0;\tfrac{1}{2},\tfrac{-\sigma}{2}\rangle \\
&= s\underbrace{\langle\tfrac{1}{2},\tfrac{+1}{2}|\tfrac{1}{2},\tfrac{+1}{2};0,0\rangle}_{=1 \text{ (charge)}}\underbrace{\langle 0,0|\tfrac{1}{2},\tfrac{\sigma}{2};\tfrac{1}{2},\tfrac{-\sigma}{2}\rangle}_{=\frac{+\sigma}{\sqrt{2}} \text{ (spin)}}\langle\tfrac{1}{2},0\|\psi\|0,\tfrac{1}{2}\rangle \\
&\Rightarrow \langle\tfrac{1}{2},0\|\psi\|0,\tfrac{1}{2}\rangle = \sqrt{2}.
\end{aligned}
$$

The order inversion of the matrix element in the first line was used since the spinor $\hat{\psi}_{\mathrm{CS}}$ in Eq. (A59) contains the creation operator $s\hat{c}_\sigma^\dagger$ and not its hermitian conjugate. The overall complex conjugation $\langle\cdot\rangle^{(*)}$, however, is irrelevant since all matrix elements are real, hence the notation of putting the asterisk in brackets.

The second non-zero reduced matrix element can be calculated in a similar fashion,

$$
\begin{aligned}
1 &= \langle 0|\hat{c}_\sigma|\sigma\rangle = s\langle\sigma|\cdot s\hat{c}_\sigma^\dagger\cdot|0\rangle^{(*)} \\
&\equiv s\langle 0,0;\tfrac{1}{2},\tfrac{\sigma}{2}|\hat{\psi}_{(\frac{1}{2}\frac{+1}{2};\frac{1}{2}\frac{\sigma}{2})}|\tfrac{1}{2},\tfrac{-1}{2};0,0\rangle \\
&= s\underbrace{\langle 0,0|\tfrac{1}{2},\tfrac{+1}{2};\tfrac{1}{2},\tfrac{-1}{2}\rangle}_{=+\frac{1}{\sqrt{2}} \text{ (charge)}}\underbrace{\langle\tfrac{1}{2},\tfrac{\sigma}{2}|\tfrac{1}{2},\tfrac{\sigma}{2};0,0\rangle}_{=1 \text{ (spin)}}\langle 0,\tfrac{1}{2}\|\psi\|\tfrac{1}{2},0\rangle \\
&\Rightarrow \langle 0,\tfrac{1}{2}\|\psi\|\tfrac{1}{2},0\rangle = s\sqrt{2}
\end{aligned}
$$

Overall, this leads to the reduced matrix elements in the charge-spin sectors $(C,S) \in \{(0,\frac{1}{2}),(\frac{1}{2},0)\}$

$$
\psi_{\mathrm{CS}}^{[1/2,1/2]} = \begin{pmatrix} 0 & s\sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}.
\tag{A63}
$$

Note that although the spinor in Eq. (A59) has four components, *i.e.* is of rank-3, on the reduced multiplet level in Eq. (A63) the spinor becomes a two-dimensional object as expected from an IROP. The further internal structure is entirely taken care of by rank-3 Clebsch Gordan coefficients, which have been omitted in Eq. (A63) [for

a full description of $\psi_{\mathrm{CS}}^{[1/2,1/2]}$ including Clebsch-Gordan coefficients in terms of a QSpace see Tbl. C5].

The operator in Eq. (A63) is non-hermitian. In the context of two-site hopping, however, this nevertheless leads to a hermitian term in the Hamiltonian, as required. Indicating the local symmetry eigenspace for site k by $|\sigma\rangle_k$, the matrix elements of the hopping term in the tensor-product basis $|\sigma_{k+1}\rangle\,|\sigma_k\rangle$ (in this order, assuming site $k+1$ is added after site k) are given by

$$
\langle\sigma_k|\langle\sigma_{k+1}|\hat{\psi}_k^\dagger\hat{\psi}_{k+1}|\sigma'_{k+1}\rangle|\sigma'_k\rangle = \psi_k^\dagger \otimes [z\psi]_{k+1}, \tag{A64}
$$

where the ψ's to the right without the hat denote the matrix elements in the local $|\sigma\rangle$ basis. Note, that fermionic signs apply, when $\hat{\psi}_k^\dagger$ is moved, for example, to the left of $\langle\sigma_{k+1}|$, such that the tensor-product on the *r.h.s.* of Eq. (A64) contains $[z\psi]_{k+1}$ rather than ψ_{k+1}, where $\hat{z}_k \equiv (-1)^{\hat{n}_k}$ is diagonal in $|\sigma_k\rangle$ and adds signs corresponding to the number of particles in $|\sigma_k\rangle$. Note that with the particle number being related to $\hat{C}_z = \frac{1}{2}(\hat{n}-1)$, the operator $\hat{z}$ is well-defined in terms of the symmetry labels. It is a scalar operator, since $(-1)^{\hat{n}-1} = (-1)^{(\hat{n}-1)^2} = (-1)^{4\hat{C}_z^2}$, hence does not alter the Clebsch-Gordan content of the operator $\hat{\psi}$ but rather acts on the multiplet level only. For the hopping $\hat{\psi}_k^\dagger\hat{\psi}_{k+1}$ between two nearest-neighbor sites, Eq. (A63) finally leads to

$$
H_{k,k+1} = \begin{pmatrix} 0 & \sqrt{2} \\ s_k\sqrt{2} & 0 \end{pmatrix} \otimes \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & s_{k+1}\sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}}_{=\begin{pmatrix} 0 & -s_{k+1}\sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}},
$$

written as a plain tensor product on the level of the multiplet spaces of two fermionic sites. For the sake of the argument, the product space here is not yet described in terms of proper combined symmetry multiplets of sites k and $k+1$.

With $s_k = (-1)^k$, $H_{k,k+1}$ in the last equation clearly yields a hermitian object for all iterations. For example, for even k, the hopping term is given by

$$
H_{k,k+1}^{[k \text{ even}]} = \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}
$$

similar in structure to a hermitian object of the type $\tau_x \otimes \tau_x$ in terms of Pauli matrices, while for odd k,

$$
H_{k,k+1}^{[k \text{ odd}]} = \begin{pmatrix} 0 & \sqrt{2} \\ -\sqrt{2} & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -\sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}
$$

similar in structure to the hermitian $(i\tau_y)\otimes(i\tau_y) = -\tau_y\otimes\tau_y$. Hence for every even (odd) iteration, one has a $\tau_x\otimes\tau_x$ ($\tau_y \otimes \tau_y$) structure, respectively, a prescription that is periodic with every pair of iterations. This intrinsic even-odd behavior is not specifically surprising, considering, for example, that two particles are needed to return to

P7.

the same charge quantum numbers related to particle hole symmetry.

In summary, using Eq. (A60), the hopping in the Hamiltonian is given by

$$\hat{H} = \sum_k t_k \hat{\psi}_k^\dagger \hat{\psi}_{k+1}. \tag{A65}$$

In a sense, the net effect of incorporating spin SU(2) was to eliminate the spin index, while incorporation of particle-hole SU(2) eliminates the hermitian conjugate term in the Hamiltonian. Together they reduce the four terms initially required for a single hopping in Eq. (A46) to the single scalar term $\hat{\psi}_k^\dagger \hat{\psi}_{k+1}$.

c. Particle-hole SU(2) symmetry for several channels

The alternating sign in the raising operator $\hat{C}_{k,+} = s\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger$ in Eq. (A52) defines the doubly occupied states as $|\frac{1}{2}; \frac{+1}{2}\rangle = \hat{C}_{k,+}|\frac{1}{2}; \frac{-1}{2}\rangle = s\hat{c}_\uparrow^\dagger c_\downarrow^\dagger|0\rangle$; for even sites, $s = +1$, therefore $|\frac{1}{2}; \frac{+1}{2}\rangle = \hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger|0\rangle$. For odd sites, on the other hand, $|\frac{1}{2}; \frac{+1}{2}\rangle = -\hat{c}_\uparrow^\dagger \hat{c}_\downarrow^\dagger|0\rangle = \hat{c}_\downarrow^\dagger \hat{c}_\uparrow^\dagger|0\rangle$. In practice, for consistency, usually a certain well-defined fermionic order is adopted. Above raising operator $\hat{C}_{k,+}$ thus suggests that it may be useful to *reverse the fermionic order of every other site* for the local state space included there.

Fully reversing the fermionic order of a given site k with several fermionic channels $i = 1, \ldots, m$ implies for the *matrix elements* for particle creation or annihilation operators,

$$c_{k,i\sigma} \to \tilde{c}_{k,i\sigma} \equiv z_k c_{k,i\sigma}$$
$$c_{k,i\sigma}^\dagger \to \tilde{c}_{k,i\sigma}^\dagger \equiv -z_k c_{k,i\sigma}^\dagger.$$

This transformation is equivalent to a unitary transformation local to site k. Similar to Eq. (A64), $\hat{z}_k \equiv (-1)^{\hat{n}_k}$ with $\hat{n}_k \equiv \sum_{i\sigma} \hat{n}_{k,i\sigma}$ and $\hat{n}_{k,i\sigma} \equiv \hat{c}_{k,i\sigma}^\dagger \hat{c}_{k,i\sigma}$ again takes care of fermionic signs for the full multi-level site k. Being a scalar operator, $\hat{z}_k$ is independent of the fermionic order.

Now consider the effect of flipping the fermionic order for the odd sites in the tight-binding chain that carry the sign $s = -1$, assuming particle-hole symmetry in every channel. For a specific channel, this (i) takes away the sign in the raising operator $\hat{C}_{k,+}$, and (ii) implies, for example, for the 4-component spinor in Eq. (A59) for a single channel,

$$\hat{\psi}_{\text{CS},k \text{ odd}} \equiv \begin{pmatrix} -\hat{c}_\uparrow^\dagger \\ \hat{c}_\downarrow \\ -\hat{c}_\downarrow^\dagger \\ -\hat{c}_\uparrow \end{pmatrix} \to \begin{pmatrix} +\hat{z}\hat{c}_\uparrow^\dagger \\ \hat{z}\hat{c}_\downarrow \\ +\hat{z}\hat{c}_\downarrow^\dagger \\ -\hat{z}\hat{c}_\uparrow \end{pmatrix} \equiv \hat{z}_k \hat{\psi}_{\text{CS},k \text{ even}}, \tag{A66}$$

having intermittently dropped the index k for readability. Therefore, up to the local operator $\hat{z}_k$ which assigns

fermionic signs to the full Hilbert space of a local site, the matrix elements of the spinor for the odd sites are *exactly* the same as the matrix elements of the spinor for even sites. Therefore with $\hat{\psi}$ taken as the spinor for even sites in the chain, the required spinor for odd sites becomes $\hat{z}\hat{\psi}$. Together with the additional fermionic signs in the nearest-neighbor hopping term as already encountered in Eq. (A64), the hopping structure $\hat{h}_{k,k+1}$ of the tight-binding Hamiltonian in Eq. (A65) becomes,

$$\psi_k^\dagger \otimes [z \cdot (z\psi)]_{k+1} = \psi_k^\dagger \otimes \psi_{k+1} \text{ for } k \text{ even}$$
$$(z\psi)_k^\dagger \otimes [z \cdot \psi]_{k+1} = (z\psi)_k^\dagger \otimes (z\psi)_{k+1} \text{ for } k \text{ odd}.$$

This result generalizes to any number of channels with particle-hole symmetry. As such it much simplifies the structure and thus the treatment of the two different kinds of spinors for even and odd sites, respectively, that had been required initially.

d. Symmetric three-channel system

P7.

Consider the generalization of the spinful one-channel setup in Eq. (A46) to a spinful three-channel system,

$$\hat{H} = \sum_k t_k \cdot \underbrace{\sum_{i=1}^{m=3} \sum_\sigma (\hat{c}_{k,i\sigma}^\dagger \hat{c}_{k+1,i\sigma} + \text{H.c.})}_{\equiv \hat{h}_{k,k+1}}, \tag{A67}$$

where $\hat{c}_{k,i\sigma}^\dagger$ creates a particle at site k in channel i with spin σ. This model is relevant for the system analyzed in the main body of the paper where the specific number of three channels, for example, originates from the underlying orbital band structure in terms of a partially filled d-shell. The Hamiltonian in Eq. (A67) can also be complemented with interaction terms that are compatible with the symmetries discussed in the following. This can include onsite interaction U at half-filling [*cf.* Eq. (A58)], or uniform local Hund's coupling J_H [*e.g.* see Eq. (25b)]. Here, however, the focus of the discussion is on symmetries, for which the Hamiltonian in Eq. (A67) suffices.

The Hamiltonian in Eq. (A67) possesses SU(2) spin symmetry, SU(2) particle-hole symmetry in each channel, and also SU(3) channel symmetry, while not all of these symmetries necessarily are independent of, *i.e.* commute with each other. All of these symmetries can be defined within the Hilbert space of a local site, hence again focusing the discussion on a single site k in the following, while dropping the site index k, for simplicity. For each of the three channels, the associated spinful fermionic level is represented by the four states as in Eq. (A62), leading to a total of $4^3 = 64$ state for a given site.

The total spin-SU(2) symmetry of a site is described

by the generators

$$\hat{S}_+ = \sum_i \hat{S}_{i+} = \sum_i \hat{c}^\dagger_{i\uparrow}\hat{c}_{i\downarrow}$$
$$\hat{S}_z = \sum_i \hat{S}_{iz} = \tfrac{1}{2}\sum_i(\hat{n}_{i\uparrow} - \hat{n}_{i\downarrow}), \tag{A68}$$

where $\hat{S}_{i\sigma}$ represents the spin operators for the fermionic level i, with $\hat{S}_z = \tfrac{1}{2}[\hat{S}_+, \hat{S}^\dagger_+]$, as expected for SU(2).

The particle-hole symmetry exists for every channel i, and is described by the SU(2) symmetry,

$$\hat{C}_{i+} = s\hat{c}^\dagger_{i\uparrow}\hat{c}^\dagger_{i\downarrow}$$
$$\hat{C}_{iz} = \tfrac{1}{2}[\hat{C}_{i+}, \hat{C}^\dagger_{i+}] = \tfrac{1}{2}(\hat{n}_i - 1), \tag{A69a}$$

which includes the same sign-factor $s_k = (-1)^k$ as in Eq. (A52) to correctly represent the hopping structure in the Hamiltonian Eq. (A67). The total charge relative to half-filling is given by (up to a factor of 2)

$$\hat{C}_z \equiv \sum_i \hat{C}_{iz}. \tag{A69b}$$

P7.

Finally, the channel symmetry is given by the minimal set of two raising operators $\{\hat{T}_+, \hat{U}_+\} \equiv \{\hat{S}_{12}, \hat{S}_{23}\}$ together with the z-operator $\{\hat{T}_z, \hat{Y}\} \equiv \{\hat{Z}_1, \hat{Z}_2\}$ as introduced through Eqs. (A28) in Eq. (A30),

$$\hat{T}_+ = \sum_\sigma \hat{c}^\dagger_{1\sigma}\hat{c}_{2\sigma}, \quad \hat{T}_z = \sum_\sigma(\hat{n}_{1\sigma} - \hat{n}_{2\sigma}),$$
$$\hat{U}_+ = \sum_\sigma \hat{c}^\dagger_{2\sigma}\hat{c}_{3\sigma}, \quad \hat{Y} = \sum_\sigma(\hat{n}_{1\sigma} + \hat{n}_{2\sigma} - 2\hat{n}_{3\sigma}). \tag{A70}$$

Here the notation for the generators of SU(3) has been changed to another notation frequently also found in literature, so these generators can be better distinguished from the generators for spin and particle-hole symmetry. In particular, the operators $\hat{T}_+$ and $\hat{T}_z$ generate an SU(2) subalgebra, that is linked to the full SU(3) symmetry through the generators $\hat{U}_+$ and $\hat{Y}$. The normalization of the z-operators, however, has been chosen consistent with Eqs. (A28), such that plain integer matrix elements arise.

The spin symmetry clearly commutes with the particle-hole symmetry in each channel, which follows from the previous one-channel discussion in Eqs. (A57). Therefore it remains to analyze the compatibility of the SU(3) channel symmetry. All z-operators clearly commute. For the SU(3) raising operators, it follows with respect to the spin symmetry,

$$[\hat{T}_+, \hat{S}_+] = \sum_{\sigma,i}[c^\dagger_{1\sigma}\hat{c}_{2\sigma}, \hat{c}^\dagger_{i\uparrow}\hat{c}_{i\downarrow}]$$
$$= \sum_\sigma(\delta_{\sigma\uparrow} - \delta_{\sigma\downarrow}) \cdot \hat{c}^\dagger_{1\uparrow}\hat{c}_{2\downarrow} = 0 \tag{A71}$$

with a similar expression for $\hat{U}_+$ instead of $\hat{T}_+$ with a shift in the channel indices. Note that in order for the *r.h.s.* to vanish, the sum over the spin σ is essential

which shows the importance of the summation over σ in Eqs. (A70). As a consequence, the SU(3) channel symmetry in Eq. (A70) commutes with the SU(2) spin symmetry, indeed.

The compatibility of the SU(3) channel symmetry with the SU(2) particle-hole symmetry, however, cannot be established, since

$$[\hat{T}_+, \hat{C}_{i+}] = \sum_\sigma[c^\dagger_{1\sigma}\hat{c}_{2\sigma}, s\hat{c}^\dagger_{i\uparrow}\hat{c}^\dagger_{i\downarrow}]$$
$$= s\delta_{i2}(\hat{c}^\dagger_{1\uparrow}\hat{c}^\dagger_{2\downarrow} - \hat{c}^\dagger_{1\downarrow}\hat{c}^\dagger_{2\uparrow}) \neq 0 \tag{A72}$$

cannot be made to vanish for all channels i at the same time. Therefore the non-abelian channel and particle-hole symmetries cannot coexist independently of each other. Nevertheless, the generators of each individual symmetry do commute with the Hamiltonian, *which thus suggests a larger symmetry*, with Eq. (A72) already indicating one of the additional generators. As it turns out, this symmetry is $\mathrm{Sp}(2m)$ with m the number of channels. This symmetry will be introduced and discussed in the next section.

By reducing the non-abelian particle-hole symmetry to its abelian conservation of total charge, however, this abelian symmetry does commute with the SU(3) channel symmetry,

$$\sum_i[\hat{T}_+, \hat{C}_{iz}] = \sum_{\sigma,i\sigma'}\tfrac{1}{2}[c^\dagger_{1\sigma}\hat{c}_{2\sigma}, c^\dagger_{i\sigma'}\hat{c}_{i\sigma'}]$$
$$= \sum_i \tfrac{1}{2}(\delta_{i2} - \delta_{i1})\sum_\sigma \hat{c}^\dagger_{1\sigma}\hat{c}_{2\sigma} = 0.$$

In order to get a commuting abelian charge symmetry, the z-operators for the channel-specific particle-hole symmetry must be summed over all channels i. With all commuting symmetries combined, this leads to the overall symmetry $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$, consisting of the SU(2) total spin symmetry in Eq. (A68), the abelian total charge of the system in Eq. (A69b), and the channel SU(3) symmetry in Eq. (A70).

A more conventional symmetry setup can be obtained by giving up the channel SU(3) symmetry. Bearing in mind that the channel-specific SU(2) particle-hole symmetries commute with total spin, this also allows the symmetry setup $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)^{\otimes 3}_{\mathrm{charge}}$.

The symmetry combinations above can be motivated also by a simple counting argument with respect to conserved abelian quantum numbers. Note that the preserved abelian quantum numbers in the Hamiltonian Eq. (A67) are the particle number in each of the three channels together with the total spin S_z. This results in a total of four z-operators, and thus four z-labels. Now, by including non-abelian flavors, the number of z-operators clearly cannot increase, but will remain the same. Total spin has one z-operator, the channel SU(3) symmetry has two z-operators, and the channel-specific particle-hole symmetries have three z-operators, which combined results in $1+2+3 = 6$ z-operators. This set of z-operators

therefore *cannot be independent* of each other, as already seen in the earlier discussion. Yet, in fact, both of the alternative symmetry setups above do have a total of four z-operators. For $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$, these are $1+1+2$ from spin, charge, and channel symmetry, respectively, while for $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)^{\otimes 3}_{\mathrm{charge}}$ these are $1+3$ from spin and each channel.

For the symmetry setting $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ then, the hopping term in the Hamiltonian in Eq. (A67) is described by a 6-component IROP $\hat{\psi}_k^{[6]}$ (annihilation operators for spin-up and spin-down combined), that can be obtained, for example, numerically as described in Sec. B 2. This leads to

$$\hat{h}_{k,k+1} = \hat{\psi}_n^{[6]\dagger} \cdot \hat{\psi}_k^{[6]} + \text{H.c.} \qquad (A73a)$$

In contrast, for the second symmetry setting $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)^{\otimes 3}_{\mathrm{charge}}$, the IROPs $\hat{\psi}_{k,i}^{[4]}$ required for the hopping term are already exactly the 4-component spinors in Eq. (A59), *i.e.* one for each individual channel, $i = 1, \ldots, 3$. The hopping in the Hamiltonian is thus described by

$$\hat{h}_{k,k+1} = \sum_{i=1}^{m=3} \hat{\psi}_{k,i}^{[4]\dagger} \hat{\psi}_{k+1,i}^{[4]}, \qquad (A73b)$$

i.e. without the hermitian conjugate part as this is already included through particle-hole symmetry. Furthermore, note that particle-hole symmetry also acquires even-odd alternations for the spinors along a chain [see App. A 9 b].

10. The symplectic group $\mathrm{Sp}(2m)$

All Hamiltonians considered in this paper are time-independent, hence obey time-reversal symmetry. Time-reversal symmetry then is described by an anti-unitary operator $\hat{T} = \hat{\Sigma}_y \hat{K}$,[40] that includes a standard unitary operation $\hat{\Sigma}_y$ together with the operator $\hat{K}$, which stands for complex conjugation [the notation of $\hat{\Sigma}_y$ has been chosen for latter convenience; see Eq. (A77) below]. The time-reversal operator obeys $\hat{T}^2 = \pm 1$, where for spin-half particles, such as electrons as considered throughout in this paper, it holds $\hat{T}^2 = -1$. The latter is important for the symmetry $\mathrm{Sp}(2m)$, since it implies that the unitary $\hat{\Sigma}_y$ must be antisymmetric. This follows simply by looking at the matrix elements of the time-reversal operator for arbitrary states $|a\rangle$ and $|b\rangle$ in some real basis i,

$$\langle a|\hat{T}b\rangle = \sum_{i,j} a_i^* (\Sigma_y)_{ij} b_j^*,$$

yet it also holds,

$$\langle a|\hat{T}b\rangle = \langle \underbrace{\hat{T}^2}_{-1} b|\hat{T}a\rangle = -\sum_{i,j} b_j^* (\Sigma_y)_{ji} a_i^*.$$

As this applies for arbitrary states $|a\rangle$ and $|b\rangle$, this shows that, given $\hat{T}^2 = -1$, the unitary $\hat{\Sigma}_y$ must be antisymmetric, indeed.

Since a time-independent Hamiltonian obviously commutes with the time-reversal operator, it follows that all eigenstates of the Hamiltonian can also be written as eigenstates of the the time-reversal operator $\hat{T}$. As a consequence, all unitary symmetry operations $\hat{G} = \exp(i \sum_\sigma a_\sigma \hat{S}_\sigma)$ can be *constrained* to unitaries which also leave the time-reversal operator invariant. That is,

$$\hat{T} \overset{!}{=} \hat{G}\hat{T}\hat{G}^{-1} = \hat{G}\hat{\Sigma}_y \underbrace{\hat{K}\hat{G}^\dagger}_{=\hat{G}^T \hat{K}}$$

$$\Rightarrow \quad \hat{\Sigma}_y = \hat{G}\hat{\Sigma}_y \hat{G}^T. \qquad (A74)$$

For the generators $\hat{S}_\sigma$ of a symmetry group this implies (*e.g.* by expansion of the exponential in $\hat{G}$ to first order in a_σ), that

$$\hat{S}_\sigma \hat{\Sigma}_y + \hat{\Sigma}_y \hat{S}_\sigma^T = 0, \qquad (A75)$$

This exactly corresponds to the definition of the Lie algebra $\mathrm{Sp}(2m)$. Having a unitary, *i.e.* non-singular, yet also antisymmetric $\hat{\Sigma}_y$, this requires a global Hilbert space of *even* dimension N, since $\det(\Sigma_y) = \det(\Sigma_y^T) = (-1)^N \det(\Sigma_y) \neq 0$. While this argument holds on the entire Hilbert space, for a specific symmetry subspace (carrier space) of an irreducible representation of $\mathrm{Sp}(2m)$ this is not necessarily the case. Specifically, there are IREPs with *odd* dimensions, a simple example being the scalar representation with dimension 1. Within such an irreducible representation, a non-singular antisymmetric Σ_y does not exist. This is not a problem, however, since the existence of Σ_y is required only globally, and also in the *defining* representation, which thus has to be of even dimension.

Consider such a matrix representation of $\mathrm{Sp}(2m)$ of even dimension, which allows to explicitly construct the non-singular antisymmetric Σ_y. In this case, an arbitrary matrix $S_{(\sigma)}$ within the space of the generators of the symmetry can be written as a tensor-product with a two-dimensional space, which itself can be expanded in terms of the Pauli matrices τ_σ [*cf.* Eq. (A6)],

$$S_{(\sigma)} \equiv \sum_{x=0}^3 \tau_x \otimes \mathrm{S}_x^{(\sigma)}, \qquad (A76)$$

where $x \in \{0,1,2,3\} \equiv \{0,x,y,z\}$ and $\tau_0 \equiv 1^{(2)}$ the two-dimensional identity matrix. Here the same letter S is used left and right in Eq. (A76), as their interpretation is related. Nevertheless, they refer to different objects. So in order to distinguish them, the generators on the *l.h.s.* are written with Greek-letter subscripts (σ), while their decomposition $\mathrm{S}_x^{(\sigma)}$ is denoted in roman font with roman or numeric subscripts. Moreover, for readability, the index σ referring to a specific generator will be skipped in the following where not explicitly required (hence the σ has been put in brackets).

P7.

Now, with representations of a symmetry unique up to similarity transformation, one is free to choose the form of the matrix representation of the operator $\hat{\Sigma}_y$ in Eq. (A75). In the two-dimensional (block) space described by the Pauli matrices then, Σ_y is chosen as follows,[38,39]

$$\Sigma_y = \tau_y \otimes 1^{(m)} \equiv \begin{pmatrix} 0^{(m)} & -i1^{(m)} \\ i1^{(m)} & 0^{(m)} \end{pmatrix}, \qquad (A77)$$

where the last term explicitly denotes the tensor block-decomposition of $m \times m$ matrices, with $0^{(m)}$ $[1^{(m)}]$ an $m \times m$ dimensional zero [identity] matrix, respectively. This Σ_y fulfills the minimal requirement that it is (i) unitary and (ii) antisymmetric. Using the Σ_y in Eq. (A77) in the defining equation for Sp($2m$), Eq. (A75), and the fact that the generators $S_{(\sigma)}$ in Eq. (A76) shall refer to hermitian operators to start with, this implies for the decomposition S_x, that $S_0 \equiv iA$ is a purely imaginary and antisymmetric matrix, while the remaining S_x for $x = (1,2,3)$ must be real symmetric matrices. In summary, this allows to rewrite the matrix block-decomposition in Eq. (A76) in the form,[38,39]

$$S = \begin{pmatrix} iA + S_3 & S_1 - iS_2 \\ S_1 + iS_2 & iA - S_3 \end{pmatrix} \equiv \begin{pmatrix} C & D^\dagger \\ D & -C^T \end{pmatrix}, \qquad (A78)$$

where $C \equiv iA + S_3$ ($D \equiv S_1 + iS_2$) is an arbitrary hermitian (symmetric) $m \times m$ matrix, respectively. The resulting number of free parameters is $m^2 + m$ for the matrix D (where the $+m$ comes from the fact that the diagonal can be fully complex), and m^2 for the hermitian matrix C. The total number of free parameters of the ($N \equiv 2m$)-dimensional matrices therefore is,

$$g = m(2m+1) \equiv \frac{N}{2}(N+1). \qquad (A79)$$

In case of the defining representation, by construction, this also corresponds to the dimension of the symmetry group Sp($2m$). For comparison, for example, the orthogonal group $O(N)$ has dimension $\frac{N}{2}(N-1)$.

Setting the off-diagonal block-matrix D in Eq. (A78) to zero, and using arbitrary hermitian *yet also traceless* matrices C, this directly demonstrates that $SU(m)$ is contained as a subalgebra within Sp($2m$). This subalgebra $SU(m)$ has rank $m-1$, *i.e.* has $m-1$ z-operators. Now, the full Sp($2m$) symmetry also includes the *tracefull* hermitian matrix C. This introduces the remaining m-th z-operator, $Z_m = \tau_z \otimes 1^{(m)}$. With a total of m z-operators, Sp($2m$) therefore has rank m, with the z-operators given by

$$Z_k \equiv \tau_z \otimes Z_k^{(m)}, \qquad (A80a)$$

where

$$Z_k^{(m)} = \begin{cases} (Z_k^{(m)})^{\mathrm{SU}(m)} & k = 1, \ldots, m-1 \\ 1^{(m)} & k = m \end{cases}, \qquad (A80b)$$

with $\left(Z_k^{(m)}\right)^{\mathrm{SU}(m)}$ the standard $m \times m$ dimensional z-operators for SU(m). By construction, all of these z-operators can be considered diagonal, as they form a mutually commuting set of matrices.

Leaving the space of strictly hermitian generators, the canonical RLOs from the SU(m) subalgebra are given by

$$S_{ij} \equiv \begin{pmatrix} S_{ij} & 0 \\ 0 & -S_{ij}^T \end{pmatrix}, \qquad (i \neq j) \qquad (A81)$$

with $S_{ij} \equiv E_{ij}$ given by the non-symmetric matrices in Eq. (A27b). This encodes both, raising and lowering operators, depending on $i < j$ or $i > j$, respectively. Having $(m^2 - 1) + 1 = m^2$ generators from the SU(m) subalgebra together with Z_m, the remaining $m(m+1)$ operators are split equally into complimentary raising and lowering operators. The corresponding canonical RLOs can be chosen as follows,[38,39]

$$\tilde{S}_{ij}^\pm \equiv \frac{1}{2}(\tau_x \pm i\tau_y) \otimes \tilde{S}_{ij}, \qquad (\text{all } i, j) \qquad (A82)$$

with the *symmetric* matrices $\tilde{S}_{ij} \equiv \frac{1}{2}(E_{ij} + E_{ji})$. Here the tilde serves to differentiate the RLOs from the SU(m) subalgebra in Eq. (A81). Having symmetric $\tilde{S}_{ij}$, *i.e.* $\tilde{S}_{ij} = \tilde{S}_{ji}$, this describes a total of $\frac{1}{2}m(m+1)$ raising operators. Complemented by $\frac{1}{2}m(m+1)$ lowering operators, indeed, this completes the group of generators for the Lie algebra Sp($2m$).

Using the canonical representation for SU(m) together with above extension to Sp($2m$), this provides the canonical representation for Sp($2m$) as in Eq. (A20). For example, with

$$(\vec{z}_i)_k \equiv z_{k,i}^{(m)} \equiv (Z_k^{(m)})_{ii} \qquad (A83)$$

referring to the i-th diagonal matrix element of the diagonal matrices $Z_k^{(m)}$, it follows

$$Z_k^{(m)}\tilde{S}_{ij} = \tilde{S}_{ij}Z_k^{(m)} = \underbrace{(z_{k,i}^{(m)} + z_{k,j}^{(m)})}_{\equiv (\vec{z}_i + \vec{z}_j)_k} \cdot \tilde{S}_{ij},$$

and thus

$$\begin{aligned} [Z_k, \tilde{S}_{ij}^\pm] &= [\tau_z \otimes Z_k^{(m)}, \tfrac{1}{2}(\tau_x \pm i\tau_y) \otimes \tilde{S}_{ij}] \\ &= \underbrace{\tfrac{1}{2}[\tau_z, \tau_x \pm i\tau_y]}_{= \pm(\tau_x \pm i\tau_y)} \otimes \left((\vec{z}_i + \vec{z}_j)_k \tilde{S}_{ij}\right) \\ &= \pm(\vec{z}_i + \vec{z}_j)_k \cdot \tilde{S}_{ij}^\pm, \qquad (A84a) \end{aligned}$$

(no summation over i or j). Similarly, for the RLOs S_{ij} from the SU(m) subalgebra, with

$$[Z_k^{(m)}, S_{ij}] = (\vec{z}_i - \vec{z}_j)_k \cdot S_{ij}$$

it follows,

$$\begin{aligned}
\left[Z_k, S_{ij}\right] &= \left[\tau_z \otimes Z_k^{(m)}, \begin{pmatrix} S_{ij} & 0 \\ 0 & -S_{ij}^T \end{pmatrix}\right] \\
&= \begin{pmatrix} \left[Z_k^{(m)}, S_{ij}\right] & 0 \\ 0 & +\left[Z_k^{(m)}, S_{ij}^T\right] \end{pmatrix} \\
&= (\vec{z}_i - \vec{z}_j)_k \cdot S_{ij},
\end{aligned} \tag{A84b}$$

since $\left[Z, S^T\right] = -\left[Z, S\right]^T$. This confirms that the z-operators together with the raising and lowering operators are in the expected canonical form, indeed.

a. Internal multiplet ordering

The block-decomposition of Eq. (A76) is not yet ordered w.r.t. to the RLOs, i.e. the z-labels [here, by definition, it is assumed that a raising (lowering) operator leads to a larger (smaller) z-label in root space which directly links to the underlying sorting implemented in root space]. The starting point, however, is correct: (i) The $(D = 2m)$ dimensional first state $|e_1\rangle$ [cf. Eq. (A27a)] does represent the maximum weight state, indeed, and (ii) by applying the $m - 1$ lowering operators from the $SU(m)$ subalgebra, this iteratively demotes the MW-state through the states $|e_2\rangle$, ..., $|e_m\rangle$. So far the state order is correct.

However, the next lower state is obtained by the m-th lowering operator, i.e. the one that links to the full $Sp(2m)$ symmetry. This will generate the state $|e_D\rangle$, which thus is not in order. Through another sequence of lowering operators from the $SU(m)$ subalgebra, finally this proceeds through the states $|e_{D-1}\rangle$, ..., $|e_{D-m+1}\rangle$ with additional alternating signs. The full sequence of normalized states thus obtained starting from the MW-state, can be collected as columns into a unitary matrix U,

$$U \equiv \begin{pmatrix} 1^{(m)} & 0 \\ 0 & \Sigma^{(m)} \end{pmatrix}, \tag{A85a}$$

with the $m \times m$ dimensional matrix $\Sigma^{(m)}$

$$\Sigma^{(m)} \equiv \begin{pmatrix} \cdot & \cdot & \cdot & \cdot^{\cdot^{\cdot}} \\ \cdot & \cdot & +1 & \cdot \\ \cdot & -1 & \cdot & \cdot \\ +1 & \cdot & \cdot & \cdot \end{pmatrix}, \tag{A85b}$$

to be distinguished from Σ_y in Eq. (A77) associated with time-reversal symmetry. The unitary U in Eq. (A85a) maps the basis into the correct order w.r.t. to sorted z-labels, as is assumed throughout this paper. Therefore this basis convention will be used henceforth, which requires U to be applied to all generators.

The transformation of an arbitrary symmetry operation S in Eq. (A78) then leads to $S \to U^\dagger S U$, that is

$$\begin{pmatrix} C & D^\dagger \\ D & -C^T \end{pmatrix} \to \begin{pmatrix} C & (\Sigma^T D)^\dagger \\ \Sigma^T D & -C^t \end{pmatrix}. \tag{A86}$$

In $\Sigma^T D$, Σ^T flips the order of the rows in D with alternating signs, starting with $+1$ on the new first row. The transformation $C^t \equiv \Sigma^T C^T \Sigma$ in the lower right block, finally, corresponds to inversion of C w.r.t. its center with alternating checker-board like minus signs applied, starting with plus signs along the regular matrix diagonal. With C hermitian, when taken real, C^t is equivalent to transposition w.r.t. the *minor* diagonal,[39] thus indicated by superscript lowercase t [this is in contrast to the standard transposition $(\cdot)^T$ around the regular diagonal].

All generators inherited from the $SU(m)$ subalgebra thus become

$$S \to \begin{pmatrix} S_i & 0 \\ 0 & -S_i^t \end{pmatrix}. \tag{A87}$$

In particular, all z-operators have the diagonal in the lower-right diagonal flipped to reverse order. The *simple* RLOs from the $SU(m)$ subalgebra now have two strictly positive entries $+1$ at the first upper subdiagonal at symmetric positions w.r.t. the *center* of the matrix. The remaining simple raising operator completing the $Sp(2m)$ algebra (see below) is given by the matrix $\tilde{S}_{mm} = E_{mm} \to E_{mm}\Sigma = +E_{m1}$ in the upper right block, thus naturally completing the set of simple raising operators of the type

$$S^+_{(\alpha=1)} = \begin{pmatrix} 0 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & \ddots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \ddots & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}, \tag{A88}$$

with $\alpha = 1, \ldots, m$ indicating the position of the entries of 1 moving towards the center of the first upper off-diagonal.

b. Multiplet labels for $Sp(2m)$ for $m = 3$

With the RLOs defined to have at most two matrix elements exactly equal to 1, the canonical commutator relations in Eqs. (A84) directly depict the diagonal elements of the z-matrices. As already indicated in Eq. (A83), these diagonals can be combined as *rows* into an $r \times D$ matrix $z_{k,i}$, to be referred to as z-matrix, with $r = m$ being the rank of the symmetry and $D = 2m$ the dimension of the defining matrix representation. The vectors $\vec{z}_i$ in Eq. (A83) thus refer to the *columns* in the z-matrix,

P7.

and therefore directly reflect the q_z-labels, *i.e.* the root space.

For Sp(6), this 3×6 dimensional z-matrix reads

$$z = \underbrace{\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix}}_{\equiv \tilde{z}} \left.\begin{array}{ccc} 0 & 1 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & -1 \end{array}\right). \qquad \text{(A89)}$$

By construction, all matrix elements are integers, for simplicity. The z-labels of the defining representation are *directly* specified by the columns $\vec{z}_i$ of the z-matrix. Moreover, since the z-labels are additive for tensor-product spaces, this implies that the z-labels for *arbitrary* IREPs also contain integers only.

Consequently, the root space is fully spanned by simple linear integer combinations of the vectors $\vec{z}_i$. Furthermore, also the action of the RLOs themselves can be expressed as simple shifts in root space [*cf.* Eqs. (A84)]. While in the defining representation, the z-labels in the carrier space are clearly unique, they are not linearly independent. In particular, it is sufficient to focus the discussion on the linearly independent subset of the vectors $\vec{z}_i$ in terms of the leading 3×3 block $\tilde{z}$ of the z-matrix in Eq. (A89).

In terms of the three column vectors $\vec{z}_i$ in $\tilde{z}$, the simple roots are given (i) by the simple roots of SU(m), which (ii) is complemented by one further root involving $\vec{z}_3$,

$$\begin{array}{llll} \vec{\alpha}_1 & = \vec{z}_1 - \vec{z}_2 & = (\ \ 2, \ \ 0, \ \ \ 0\ )^T & \hat{=}\ S_{12} \\ \vec{\alpha}_2 & = \vec{z}_2 - \vec{z}_3 & = (\ -1, \ \ 3, \ \ \ 0\ )^T & \hat{=}\ S_{23}, \qquad \text{(A90)} \\ \vec{\alpha}_3 & = 2\vec{z}_3 & = (\ \ \ 0, \ -4, \ \ \ 2\ )^T & \hat{=}\ \tilde{S}_{33}^+ \end{array}$$

where the correspondence with the raising operators indicated in the last column follows from Eq. (A84). Having $\vec{\alpha}_i \cdot \vec{\alpha}_j \leq 0$ for $i \neq j$ together with taking smallest integer combinations derived from the action of RLOs in Eq. (A84), this suggests *simple* roots.[38,39]

Similar to SU(N), the convention on the sorting of the z-labels is chosen lexicographic, yet as always, starting from the last z-label. In this sense, the vectors $\vec{\alpha}_i$ in Eq. (A90) are greater than $(0, 0, 0)^T$, hence positive. The corresponding operators thus *increase* the z-labels, *i.e.* correspond to raising operators, indeed. Moreover, having reduced the symmetry to its simple roots, equivalently, this also defines the set of *simple* RLOs that are sufficient to fully explore multiplet spaces. Note that above convention on the sorting of the z-labels is already also consistent with the state order in the defining representation in Eq. (A89): the z-labels strictly decrease, starting from the MW-state (the very left column) all the way to the last state represented by the very right column.

In principle, the z-labels of the MW-state already could be used as labels for the entire multiplet. However, using the vectors $\hat{z}_i$ as (non-)orthogonal basis that spans the root space, also $\vec{q} \equiv \tilde{z}^{-1} \max\{\vec{z}\}$ could be used as multiplet label, instead. The latter has the advantage

that it guarantees that the multiplet labels are strictly positive integers or zero. For consistency with literature, however, the multiplet labels for Sp($2m$) are still modified somewhat further, and thus finally derived from the MW-state as follows,

$$\vec{q} \equiv \underbrace{M\tilde{z}^{-1}}_{\equiv Q} \cdot \max\{\vec{z}\} \qquad \text{(A91a)}$$

where the matrix M,

$$M \equiv \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \text{(A91b)}$$

has been added as a further minor modification for consistency with standard literature[41] which further ensures that the multiplet labels lie dense, *i.e.* with $q = (q_1, q_2, q_3)$ any $q_i \geq 0$ will result in a valid multiplet. Overall,

$$Q \equiv \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}. \qquad \text{(A91c)}$$

For example, when applied from the right to the z-matrix in Eq. (A89), all resulting matrix elements (z-labels) are either ± 1 or 0. In particular, the MW-state of the defining representation of Sp($2 \cdot 3$) has the q-labels $(1, 0, 0)$.

c. Construction of Sp($2m$) for m-channel setup

Given the three-channel setup in the previous section with $m = 3$, the resulting defining representation for Sp($2m$) is $(2m = 6)$-dimensional. As seen from the earlier introduction of this model in Sec. A 9 d, this contains an SU(3) subalgebra, together with a third z-operator, namely total particle conservation. This subalgebra of a total of 9 generators can now be completed by 6 raising operators together with their hermitian conjugates, *i.e.* their corresponding lowering operators. This leads to a total of 21 generators, consistent with the dimension of the group Sp($2 \cdot 3$).

Using a sorted z-label space, this requires that the unitary U in Eq. (A85a) is applied to all generators of the defining representation, as well as to the initial spinor $\hat{\psi}^{[2m]} \equiv (\hat{c}_{1\uparrow}, \ldots, \hat{c}_{m\uparrow}, \hat{c}_{1\downarrow}^\dagger, \ldots, \hat{c}_{m\downarrow}^\dagger)^T$ derived from Eq. (A76). In case of $m = 3$, the properly sorted 6-dimensional spinor (IROP) spinor becomes,

$$\hat{\psi}^{[6]}_{(\uparrow)} \equiv \begin{pmatrix} \hat{c}_{1\uparrow} \\ \vdots \\ \hat{c}_{m\uparrow} \\ +\hat{c}_{m\downarrow}^\dagger \\ -\hat{c}_{m-1,\downarrow}^\dagger \\ \vdots \\ (-1)^{m-1}\hat{c}_{1\downarrow}^\dagger \end{pmatrix} \qquad \text{(A92)}$$

This naturally generalizes particle-hole symmetry in the presence of channel symmetry. The symmetry preserving hopping term in Eq. (A67), for example, can now be written as scalar contraction $\hat{h}_{k,k+1} = \sum_\sigma \big(\hat{\psi}_{k\sigma}^{[6]}\big)^\dagger \cdot \hat{\psi}_{k+1,\sigma}^{[6]}$. Note that if, in addition, also $SU(2)_{\text{spin}}$ is present, this would further double the dimension of the IROP in Eq. (A92) to a set of 12 operators, such that the hopping term in Eq. (A67) can be written as *single* scalar contraction $\hat{h}_{k,k+1} = \big(\hat{\psi}_k^{[12]}\big)^\dagger \cdot \hat{\psi}_{k+1}^{[12]}$.

All generators are given in second quantization by the quadratic form $\hat{S}_\sigma \equiv \hat{\psi}^\dagger S_\sigma \hat{\psi}$, with S_σ a $2m$-dimensional generator from the defining representation. Specifically, the remaining $\frac{1}{2}m(m+1)$ raising operators for the m-channel setup in Eq. (A82) that complete $Sp(2m)$ are given by

$$\tilde{S}_{ij}^+ = \tfrac{1}{2}\left(\tau_x + i\tau_y\right) \otimes \left(\tilde{S}_{ij}\Sigma\right) \equiv \begin{pmatrix} 0 & \tilde{S}_{ij}\Sigma \\ 0 & 0 \end{pmatrix}, \qquad \text{(A93)}$$

which leads to

$$\hat{S}_{ij}^+ \equiv \hat{\psi}^\dagger \tilde{S}_{ij}^+ \hat{\psi} = \tfrac{1}{2}\big(\hat{c}_{i\uparrow}^\dagger \hat{c}_{j\downarrow}^\dagger + \hat{c}_{j\uparrow}^\dagger \hat{c}_{i\downarrow}^\dagger\big). \quad \text{(all } i,j) \quad \text{(A94)}$$

This generates a pair of particles, the nature of which originates from the underlying general particle-hole symmetry. With $\{\hat{\psi}_i, \hat{\psi}_j^\dagger\} = \delta_{ij}$ for $\nu = 1, \ldots, 2m$, and therefore

$$\begin{aligned} \big[\hat{S}_\sigma, \hat{S}_{\sigma'}\big] &\equiv \big[\hat{\psi}_i^\dagger \left(S_\sigma\right)_{ij} \hat{\psi}_j, \hat{\psi}_{i'}^\dagger \left(S_{\sigma'}\right)_{i'j'} \hat{\psi}_{j'}\big] \\ &= \hat{\psi}^\dagger \left[S_\sigma, S_{\sigma'}\right] \hat{\psi}, \end{aligned} \qquad \text{(A95)}$$

the commutator relations within the matrix representations of the defining representation earlier directly carry over to the quadratic second-quantized operators as in Eq. (A94).

Appendix B: Numerical implementation

Tensor-product spaces are an essential ingredient to numerical renormalization group techniques such as NRG or DMRG. State spaces are enlarged iteratively by adding a small local state space at a time, *i.e.* a physical site with a few degrees of freedom. With respect to the description of strongly-correlated entangled quantum many-body states, this leads to a description which is well-known as matrix product states (MPS). Both, the existing state space (iteratively constructed itself) as well as the newly added state-space, have finite dimension and well-defined symmetry labels. New representations can therefore only emerge through the tensor product of the two spaces. In particular, all iteratively constructed quantum many body states strictly derive from the IREPs of the elementary sites. With operators usually acting locally, these are also expressed in the symmetries of the local basis. Furthermore, the local state space of a site is usually small. For example, a fermionic site has the four states described in Eq. (A62). Therefore

the IREPs present within the local state space are usually just the smallest non-trivial IREPs, often just the defining representation itself. For identical sites, the local symmetry space can be setup once and for all at the beginning of the calculation.

Having identified and labeled all symmetries on the local site level, this sets the stage for generic iterative algorithms such as NRG or DMRG. The remainder is a large exercise on tensor-product spaces. By construction, the iteratively combined spaces are finite, yet as they grow rapidly, they are eventually truncated on the multiplet level while leaving the symmetry content of the individual multiplets, *i.e.* the CGC spaces, fully intact.

1. Tensor product decomposition of symmetry spaces

The decomposition of the tensor-product space of two IREPs into irreducible multiplets has already been discussed more generally in Secs. A 5 and A 6. In the actual numerical implementation, however, in particular the presence of inner and out multiplicity must be taken care of meticulously for overall consistency. This will be discussed in the following.

Similar to Sec. A 6, consider a specific arbitrary non-abelian symmetry group $\mathcal{S}$ whose Clebsch-Gordan coefficients may not necessarily be easily accessible analytically for arbitrary multiplets. Assume two of its IREPs, q_1 and q_2, with dimensions d_{q_1} and d_{q_2}, respectively, are known together with their irreducible representations of the generators $I_\sigma^{[q_1]}$ and $I_\sigma^{[q_2]}$, specifically the z-operators (Cartan subalgebra) and the simple RLOs (simple roots). In practice, these representations either refer to small IREPs such as the defining representation, or have been generated through prior iterative calculations. As in Eq. (A36), consider their tensor-product,

$$R_\sigma^{\text{tot}} \equiv I_\sigma^{[q_1]} \otimes \mathbf{1}^{[q_2]} + \mathbf{1}^{[q_1]} \otimes I_\sigma^{[q_2]}, \qquad \text{(B1)}$$

resulting in matrices of dimension $D = d_{q_1} d_{q_2}$. Clearly the commutator relations are preserved, and the z-labels are *additive* under this operation [*cf.* Sec. A 6].

In order to determine the decomposition into IREPs, a tempting route may be through the construction of the group's Casimir operators in the combined state space and their simultaneous diagonalization together with the z-operators. However, in the presence of outer or inner multiplicity, subspaces exist that are fully degenerate in Casimir operators as well as in the z-operators. In this case, for overall consistency a *unique deterministic* algorithm must be constructed that (i) separates multiplets in the presence of outer multiplicity, and (ii) fixes a choice of basis for degenerate spaces within a multiplet in the presence of inner multiplicity. Moreover, the explicit construction of the Casimir operators bears some efforts of its own. In practice, therefore a more straightforward approach has been adopted, instead, as will be explained in the following.
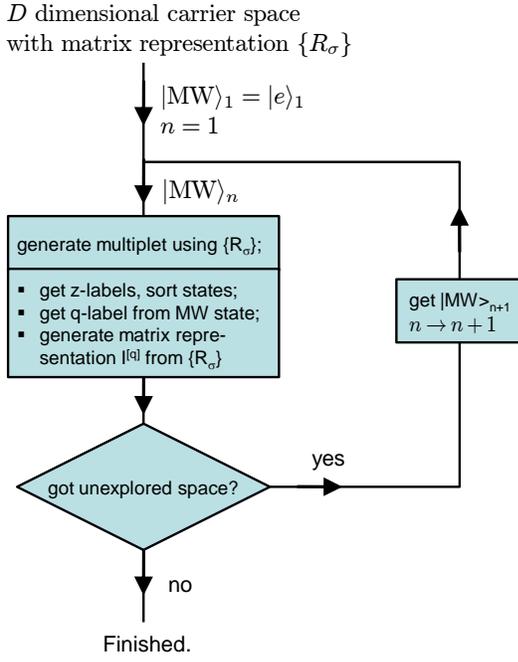
P7.

Figure 11. Schematic procedure of state space decomposition of given D dimensional vector space with known set of generators $\{R_\sigma^{(tot)}\}$.

The main hurdles in the decomposition of the tensor-product in Eq. (B1) into IREPs is the possible occurrence of outer and inner multiplicity. The strategy employed here to deal with this situation is based on the uniqueness and accessibility of the MW (maximum-weight) states as introduced in Sec. A 3 c. For this, throughout the procedure below, the *same* lexicographic sorting scheme of the z-labels, used to obtain the MW-state in Eq. (A26), is employed to order all states within an IREP. The sorting is descending, such that the MW-state appears *first* within a multiplet.

Since the z-labels are additive, it also follows for a tensor product of two such representations that the first state automatically also represents a MW-state of some multiplet,

$$|MW\rangle_1 \equiv |e_1\rangle, \tag{B2}$$

where the vectors $|e_k\rangle$ [*cf.* Eq. (A27a)] form the cartesian basis for the D-dimensional space of the representation R_σ^{tot} in Eq. (B1). Given that the MW-state of a representation is guaranteed to be unique,[25,38,39] the state $|MW\rangle_1$ is already a proper symmetry eigenstate, *i.e.* an eigenstate of all z-operators. This was always double checked, in practice, as a safety measure. The further procedure then is schematically depicted in the work flow diagram in Fig. 11: starting with $|MW\rangle_{n=1}$,

1. the symmetry eigenstate $|MW\rangle_n$ is used as the *seed state* to sequence its complete IREP (the *current multiplet*). This is done by repeatedly applying

an *arbitrary but fixed order of simple lowering operators only* to the current set of vectors in the multiplet. Therefore starting with the MW-state $|MW\rangle_n$ and adding the newly acquired symmetry states one at a time, this introduces a well-defined state order, independent of whether their z-labels are degenerate or not. In the presence of inner multiplicity, it is important to notice, however, that it is not guaranteed that a newly acquired state is automatically orthogonal to the already existing states within the current multiplet. Therefore, a newly acquired state, if it represents a new vector space component, must be orthonormalized with respect to the existing states. This is repeated, until the current multiplet space is exhausted.

2. The states in the multiplet thus generated, by construction, already have well-defined z-labels (this again was double-checked, in practice); the states are sorted with respect to these labels in descending lexicographic order while keeping subspaces that are degenerate in the z-labels in their *original order* in order to remain deterministic. Within this order, the first state defines the label for the generated multiplet, *i.e.* $q = q_{MW}$. In addition, the matrix representation in Eq. (B1), when cast into the current IREP space results in the newly generated irreducible matrix representation $\hat{I}_\sigma^{[q]}$.

3. If the D-dimensional vector space is not fully exhausted yet, a new seed state is determined by finding the *smallest* k for which $|e_k\rangle$ exhibits a new vector component *w.r.t.* the symmetry states already collected. Having started with $k = 1$ above, it follows $k > 1$. After proper orthonormalization with respect to the previously explored space, this state becomes the next seed state. If it already does represent a MW-state, which is typically the case in that it is destroyed by all raising operators, then $|MW\rangle_{n+1}$ has been found. Otherwise repeatedly apply simple raising operators on the current seed state until the unique new maximum weight state $|MW\rangle_{n+1}$ is reached. Continue with (1), setting $n \to n+1$.

4. If on the other hand, the D-dimensional vector space in Eq. (B1) is already fully exhausted, the decomposition of the tensor-product space into n irreducible representations is completed, and the procedure terminates.

Note that no explicit reference to z-labels has been made, except for step (2). That step, however, is actually not required right away for the decomposition, with its results only relevant for subsequent calculations. By construction, therefore this procedure is deterministic and does not dependent on dealing with degeneracies in the z-labels or inner and outer multiplicities. The MW-states are accessible by keeping IREPs sorted in their z-labels

throughout. They represent the entry point in sequencing its IREP, which *guarantees* that inner and outer multiplicities are dealt with in a consistent fashion. Finally, note that the choice of the seed states $|\mathrm{MW}\rangle_n$, *i.e.* starting with $+|e\rangle_k$, also provides the sign convention.

The resulting unitary transformation into the irreducible symmetry subspaces directly determines (i) the Clebsch Gordan coefficients, and (ii) the matrix-representations of the newly generated IREPs. With only a few Clebsch-Gordan coefficients usually unequal zero and of order 1, small numbers below a numerical noise threshold for double-precision (10^{-12}) are neglected, *i.e.* set to zero. Moreover, a non-zero Clebsch-Gordan coefficient can typically be expressed as a rational number, or the square root of a rational number, an efficient approximation of which can be found through continued fraction techniques. Therefore if an excellent fractional approximation was found within the same accuracy of 10^{-12}, this rational approximation also was used, instead.

2. State space initialization and operator compactification

In the presence of several symmetries, a given state space is represented by a certain set of multiplet combinations. For a single fermionic site in the presence of spin-symmetry and particle-hole symmetry, this still can be easily characterized by hand [*cf.* Eq. (A62)]. The situation, however, can quickly become more involved. For example, for a spinful three-channel calculation with SU(3) channel symmetry as in Eq. (25a), a site is represented by $4^3 = 64$ states (4 fermionic states for each of the 3 channels). If for example, particle number, spin symmetry and channel symmetry is preserved, then this system exhibits $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ symmetry, as discussed in Sec. A 9 d. Given these symmetries, the 64-dimensional Hilbert space of a site cannot be decomposed into a tensor product of convenient smaller units with already well-defined $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ symmetry labels themselves. For the channel symmetry it is essential, of course, that all three channels are present, while it is also essential for the spin symmetry that both spin species are present. Therefore in the example above, the 64-dimensional space of site already appears as the smallest building block. It can be reduced to a set of irreducible multiplet combinations, of course, but explicit determination can quickly become tedious if done by hand, while the problem can be tackled completely generally and straightforwardly on a numerical level.

In order to get started numerically, a simple and natural starting point is the Fock space representation. While this usually does not represent the symmetry eigenbasis, of course, nevertheless all generators of the symmetries present, in particular its raising, lowering, and z-operators, are known in second-quantized form and can be equally constructed in Fock space.

The z-operators typically have a simple form. In particular, for 3-channel setup mentioned above, the z-operators are already all diagonal in the Fock space, *cf.* Eqs. (A68), (A69b), or (A70). This thus already provides the z-labels. Next, note that the order of the states *w.r.t.* to their z-labels is important for consistency with the Clebsch Gordan coefficients later, which suggests using the same lexicographic order as for the determination of the MW-states. Sorting the states in this order and applying the same prescription for state space decomposition as explained in Sec. B 1, this suffices to fully identify all symmetry multiplets within the given D-dimensional Hilbert space.

a. Compactifying operators using Wigner-Eckart theorem

Irreducible operator sets can be equally constructed starting from the Fock space representation of a *seed operator* that is part of some irreducible operator set. This seed operator is typically known, yet can be completed to an IROP set, by using the RLOs in Fock space representation and numerically evaluating the commutators in Eq. (A42) (see also subsequent discussion in Sec. A 7 a). Using the same unitary transformation that brings the Fock space into the correct symmetry eigenbasis as described above, the IROP set is rotated into the space of symmetry eigenstates. With this, however, this IROP set is still represented in the *fully expanded* multiplet space, *i.e.* this space still references both multiplet labels and their corresponding z-labels on the same flat level. However, through the Wigner-Eckart theorem, Eq. (A43),

$$\langle qq_z|\hat{F}^{q_1}_{q_{1z}}|q_2q_{2z}\rangle = \langle q\|\hat{F}^{q_1}\|q_2\rangle \cdot C^{[q_{2z}]}_{q_{1z}q_z(\sigma)}, \qquad \text{(B3)}$$

many of the matrix elements can be related to each other through Clebsch Gordan coefficients. The IROP set can therefore be *compactified* as a tensor-product of reduced matrix-elements $\langle q\|\hat{A}_{q_1}\|q_2\rangle$ in the multiplet space times the CGC space $C^{[q_{2z}]}_{q_{1z}q_z(\sigma)}$.

The CGC spaces are known from a separate numerical calculation, *e.g.* they can be generated by several iterations of tensor-product decompositions starting from the defining representation. Therefore, the final compactification in Eq. (B3) of the fully expanded matrix elements of the IROP also serves as a major consistency check. The first non-zero matrix-element $\langle qq_z|\hat{F}^{q_1}_{q_{1z}}|q_2q_{2z}\rangle$ for the already known multiplet spaces $(q; q_1, q_2)$ can be used to determine the reduced matrix element $\langle q\|\hat{F}^{q_1}\|q_2\rangle$, with its corresponding Clebsch Gordan coefficient known. This, however, immediately predicts the existence of a set of other non-zero matrix elements within the same multiplet spaces $(q; q_1, q_2)$. These matrix elements must exist and agree within numerical noise. The matched matrix elements are marked and considered taken care of. If the same value of a matrix element occurs several times within the multiplets $(q; q_1, q_2)$ for the same z-labels, the

first one that matches is taken. This check is thus not entirely unique, but nevertheless a strong one, and sufficient to obtain the space of reduced matrix elements. Finally, having identified all non-zero matrix elements, the multiplet matrix element space $\langle q \| \hat{F}^{q_1} \| q_2 \rangle$ are stored together with their referenced CGC space in terms of a QSpace as discussed in the main text.

Appendix C: Example QSpaces

QSpaces represent an efficient numerical description of tensors of arbitrary rank in the presence of arbitrary quantum symmetries [*cf.* Eq. (5)]. This includes both abelian and non-abelian symmetries, with the extension to further symmetries such as point symmetries being straightforward. The QSpaces are decomposed into a set of reduced multiplet spaces together with their respective CGC (Clebsch Gordan coefficient) spaces. In the following several elementary examples of QSpaces are given as they appeared in practice. Elementary QSpaces typically have rank-2 (such as scalar operators with identity CGC spaces) or rank-3 (IREPs and IROPs with reference to standard rank-3 CGC spaces), while combinations of these through subsequent algebraic operations can easily result in higher-rank intermediate objects.

The notation regarding the elementary data arrays will be as follows. Plain matrices of dimension $m \times n$ will be written as $a = [a_{11}, \ldots, a_{1n}; \ldots; a_{m1}, \ldots, a_{mn}]$, *i.e.* m rows of equal length n separated by semicolons. The commas within a row are considered optional. In order to deal with $m \times n \times k$ dimensional rank-3 objects, the notation $\{a_1, a_2, \ldots, a_k\}$ is used, which shall indicate that the matrices $a_1, \ldots, a_k$, all of the same dimension $m \times n$, are concatenated along the third dimension. Trailing singleton dimensions will be considered implicit if required, *e.g.* a scalar such as 1. can stand for an arbitrary rank-r object in that a number also represents a $1 \times 1 \times \ldots \times 1$ object. Identity matrices of dimension n will be denoted by $\mathbf{1}^{(n)}$.

1. Fermionic site with $U(1)_{\text{charge}} \otimes SU(2)_{\text{spin}}$ symmetry

Consider the state space of a single fermionic site with the four states: empty $|0\rangle$, singly occupied $|\uparrow\rangle$ and $|\downarrow\rangle$, and double occupied $|\uparrow\downarrow\rangle$. The symmetries considered are particle conservation $U(1)_{\text{charge}}$, and full spin symmetry $SU(2)_{\text{spin}}$. The z-operators are $\hat{C}_z \equiv \frac{1}{2}(\hat{n}_\uparrow + \hat{n}_\downarrow - 1)$ and $\hat{S}_z \equiv \frac{1}{2}(\hat{n}_\uparrow - \hat{n}_\downarrow)$, with the corresponding quantum labels C_z for charge and S for total spin. For consistency with later, here the charge is treated as the reduction of the non-abelian particle-hole symmetry to its abelian part, which also reduces the set of symmetry operations to the z-operator $\hat{C}_z$ only [hence the factor $\frac{1}{2}$]. Consequently, the z-label of the underlying non-abelian symmetry is promoted into a q-label, while the CGC space

becomes trivial (1.) with internal *multiplet* dimension of 1. In order to stress the difference between the original z-label which can become negative, and the SU(2) q-labels of multiplets which are positive, by definition, the q-labels are therefore written as $(+C_z, S)$, emphasizing the origin of the q-label C_z being derived from a z-operator.

a. Symmetry space and operators of one site

The states $|0\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, and $|\uparrow\downarrow\rangle$ already represent the correct symmetry eigenstates [*cf.* Eq. (A62)],

$$
\begin{array}{c|c}
\text{multiplet space} & \text{dimension} \\
|C_z; S\rangle & d_{C_z} \times d_S = d_{\text{tot}} \\
\hline
\left|-\frac{1}{2}; 0\right\rangle \equiv |0\rangle & 1 \times 1 = 1 \\
\left|+\frac{1}{2}; 0\right\rangle \equiv |\uparrow\downarrow\rangle & 1 \times 1 = 1 \\
\left|0; \frac{1}{2}\right\rangle \equiv \{|\uparrow\rangle, |\downarrow\rangle\} & 1 \times 2 = 2
\end{array} \tag{C1}
$$

The matrix elements of a generic Hamiltonian in this basis can be written as QSpace [see definition in Eq. (5)],

$$
H \equiv \left\{
\begin{array}{cc|c|cc}
(C_z; S) & (C_z'; S') & \|H\| & \text{CGC spaces} & \\
\hline
-\frac{1}{2}; 0 & -\frac{1}{2}; 0 & h_{-\frac{1}{2}, 0} & 1. & 1. \\
+\frac{1}{2}; 0 & +\frac{1}{2}; 0 & h_{+\frac{1}{2}, 0} & 1. & 1. \\
0; \frac{1}{2} & 0; \frac{1}{2} & h_{0, \frac{1}{2}} & 1. & \mathbf{1}^{(2)}
\end{array}
\right\} \tag{C2}
$$

The Hamiltonian is a scalar operator, hence its rank as an IROP can be reduced from three to two, as it is the only operator in its irreducible set. Consequently, all CGC spaces reduce to the identity, as reflected in the last two columns of the QSpace (C2). Each of the remaining two indices explicitly refers to symmetry states, hence the QSpace requires the two sets of q-labels $q \equiv (C_z; S)$ and $q' \equiv (C_z'; S')$ referring to the first (second) index shown in the first (second) column, respectively. With the Hamiltonian preserving the symmetries, it must be block-diagonal, *i.e.* $q = q'$ for all records in (C2). Both of the symmetry spaces $(\pm\frac{1}{2}; 0)$ have a single state only, therefore the corresponding entries in the multiplet space $h_{\pm 1/2, 0}$ are 1×1 dimensional blocks, *i.e.* numbers. The last symmetry multiplet $(0; \frac{1}{2})$ has two states owing to the SU(2) symmetry, see (C1). By means of the Wigner Eckart theorem, the space of reduced matrix elements, $h_{0,1/2}$, is therefore again a number while the CGC space becomes a 2-dimensional identity matrix. Therefore the most general representation of a scalar operator for a single fermionic level in the presence of $U(1)_{\text{charge}} \otimes SU(2)_{\text{spin}}$ symmetry is given by the three numbers $\{h_{-1/2,0}, h_{+1/2,0}, h_{0,1/2}\}$ in the multiplet space. The remaining matrix elements are constrained due to symmetry.

As an example for a non-scalar IROP, consider the spinor of particle creation operators $\hat{\psi}_S^\dagger = \{\hat{c}_\uparrow^\dagger, \hat{c}_\downarrow^\dagger\}$ that

Table I. Example QSpaces in the presence of $U(1)_{\text{charge}} \otimes SU(2)_{\text{spin}}$ symmetry for a single fermionic site ($\hat{\psi}_S^\dagger$ and $\hat{\psi}_S$), and for the combination of the state space of two sites (A-tensor). Having the two symmetries of abelian $U(1)_{\text{charge}}$ and non-abelian $SU(2)_{\text{spin}}$, the respective CGC spaces C (trivial) and S appear in the right columns. The record index ν in the first column, as well as the explicit specification of the dimensions of the reduced multiplet space and the combined CGC spaces are just added for better clarity. For comparison, Tbl. II shows how the QSpaces (C3a) and (C4) are modified for the case that the abelian charge symmetry also becomes a non-abelian $SU(2)_{\text{charge}}$ particle-hole symmetry.

$$\psi_S^\dagger \equiv \left\{ \begin{array}{c|ccc|cc|ccc} \text{record} & & & & \multicolumn{2}{c|}{\text{reduced matrix elements}} & \multicolumn{3}{c}{\text{CGC spaces}} \\ \text{index } \nu & (C_z; S) & (C_z'; S') & (C_z''; S'') & \|\psi_S^\dagger\|, & \text{dimension} & \text{C} & \text{S} & \text{dimension} \\ \hline 1. & 0; \frac{1}{2} & \frac{-1}{2}; 0 & \frac{+1}{2}; \frac{1}{2} & 1. & 1 \times 1 \times 1 & 1., & \{[1; 0], [0; 1]\} & 2 \times 1 \times 2 \\ 2. & \frac{+1}{2}; 0 & 0; \frac{1}{2} & \frac{+1}{2}; \frac{1}{2} & \sqrt{2} & 1 \times 1 \times 1 & 1., & \frac{1}{\sqrt{2}}\{[0\ 1], [-1\ 0]\} & 1 \times 2 \times 2 \end{array} \right\} \quad \text{(C3a)}$$

$$\psi_S \equiv \left\{ \begin{array}{c|ccc|cc|ccc} \text{record} & & & & \multicolumn{2}{c|}{\text{reduced matrix elements}} & \multicolumn{3}{c}{\text{CGC spaces}} \\ \text{index } \nu & (C_z; S) & (C_z'; S') & (C_z''; S'') & \|\psi_S\|, & \text{dimension} & \text{C} & \text{S} & \text{dimension} \\ \hline 1. & 0; \frac{1}{2} & \frac{+1}{2}; 0 & \frac{-1}{2}; \frac{1}{2} & 1. & 1 \times 1 \times 1 & 1., & \{[1; 0], [0; 1]\} & 2 \times 1 \times 2 \\ 2. & \frac{-1}{2}; 0 & 0; \frac{1}{2} & \frac{-1}{2}; \frac{1}{2} & -\sqrt{2} & 1 \times 1 \times 1 & 1., & \frac{1}{\sqrt{2}}\{[0\ 1], [-1\ 0]\} & 1 \times 2 \times 2 \end{array} \right\} \quad \text{(C3b)}$$

QSpaces $\hat{\psi}_S^\dagger = \{\hat{c}_\uparrow^\dagger, \hat{c}_\downarrow^\dagger\}$ and $\hat{\psi}_S = \{-\hat{c}_\downarrow; \hat{c}_\uparrow\}$ representing IROPs for a single spinful fermionic level [cf. Eqs. (A49)]. Note that the IROP $\hat{\psi}_S^\dagger$ is interpreted differently compared to the IROP $(\hat{\psi}_S)^\dagger$, hence $\hat{\psi}_S^\dagger \neq (\hat{\psi}_S)^\dagger$ [e.g. note the sign in multiplet space $\|\hat{\psi}_S\|$ in the second record of $\hat{\psi}_S$ or the reverted signs in the q-labels for $(C_z''; S'')$ associated with the IROP in the third column; see text].

P7.

$$A \equiv \left\{ \begin{array}{c|ccc|cc|ccc} \text{record} & \text{site 1} & \text{site 2} & \text{combined} & \multicolumn{2}{c|}{\text{multiplet space}} & \multicolumn{3}{c}{\text{CGC Spaces}} \\ \text{index } \nu & (C_z; S) & (C_z'; S') & (C_z''; S'') & \|A\| & \text{dimension} & \text{C} & \text{S} & \text{dimension} \\ \hline 1. & \frac{-1}{2}; 0 & \frac{-1}{2}; 0 & -1; 0 & 1. & 1 \times 1 \times 1 & 1. & 1. & 1 \times 1 \times 1 \\ \hline 2. & \frac{-1}{2}; 0 & 0; \frac{1}{2} & \frac{-1}{2}; \frac{1}{2} & \{[1], [0]\} & 1 \times 1 \times 2 & 1. & \{[1\ 0], [0\ 1]\} & 1 \times 2 \times 2 \\ 3. & 0; \frac{1}{2} & \frac{-1}{2}; 0 & \frac{-1}{2}; \frac{1}{2} & \{[0], [1]\} & 1 \times 1 \times 2 & 1. & \{[1; 0], [0; 1]\} & 2 \times 1 \times 2 \\ \hline 4. & \frac{-1}{2}; 0 & \frac{1}{2}; 0 & 0; 0 & \{[1], [0], [0]\} & 1 \times 1 \times 3 & 1. & 1. & 1 \times 1 \times 1 \\ 5. & 0; \frac{1}{2} & 0; \frac{1}{2} & 0; 0 & \{[0], [1], [0]\} & 1 \times 1 \times 3 & 1. & [0\ \frac{-1}{\sqrt{2}};\ \frac{1}{\sqrt{2}}\ 0] & 2 \times 2 \times 1 \\ 6. & \frac{1}{2}; 0 & \frac{-1}{2}; 0 & 0; 0 & \{[0], [0], [1]\} & 1 \times 1 \times 3 & 1. & 1. & 1 \times 1 \times 1 \\ \hline 7. & 0; \frac{1}{2} & 0; \frac{1}{2} & 0; 1 & 1. & 1 \times 1 \times 1 & 1. & \begin{array}{c}\{[1\ 0; 0\ 0], \\ [0\ \frac{1}{\sqrt{2}};\ \frac{1}{\sqrt{2}}\ 0], \\ [0\ 0; 0\ 1]\}\end{array} & 2 \times 2 \times 3 \\ \hline 8. & 0; \frac{1}{2} & \frac{1}{2}; 0 & \frac{1}{2}; \frac{1}{2} & \{[1], [0]\} & 1 \times 1 \times 2 & 1. & \{[1; 0], [0; 1]\} & 2 \times 1 \times 2 \\ 9. & \frac{1}{2}; 0 & 0; \frac{1}{2} & \frac{1}{2}; \frac{1}{2} & \{[0], [1]\} & 1 \times 1 \times 2 & 1. & \{[1\ 0], [0\ 1]\} & 1 \times 2 \times 2 \\ \hline 10. & \frac{1}{2}; 0 & \frac{1}{2}; 0 & +1; 0 & 1. & 1 \times 1 \times 1 & 1. & 1. & 1 \times 1 \times 1 \end{array} \right\} \quad \text{(C4)}$$

QSpace of identity A-tensor combining two fermionic sites. Site 1 with symmetries $(C_z; S)$ and site 2 with symmetries $(C_z'; S')$ are combined into the global symmetry $(C_z''; S'')$. The specific order of the records is irrelevant and hence arbitrary. Here, the records have been sorted with respect to the combined quantum labels $q'' \equiv (C_z''; S'')$, where groups with the same q'' are indicated by horizontal lines for clarity. The dimensions in the last (third) index are therefore the same within a group that shares the same $(C_z''; S'')$.

encodes $SU(2)$ spin symmetry [cf. Eq. (A49)], with its QSpace representation shown in (C3a). The z-labels of the IROP set $\hat{\psi}_S^\dagger$ are determined through the z-operators $\hat{C}_z$ and $\hat{S}_z$ acting on the components of $\hat{\psi}_S^\dagger$,

$$[\hat{C}_z, \hat{c}_\sigma^\dagger] = \frac{+1}{2} \cdot \hat{c}_\sigma^\dagger$$
$$[\hat{S}_z, \hat{c}_\sigma^\dagger] = \frac{\sigma}{2} \cdot \hat{c}_\sigma^\dagger,$$

with $\sigma \equiv \{\uparrow, \downarrow\} \equiv \{+1, -1\}$. The IROP $\hat{\psi}_S^\dagger$ is therefore

identified with the multiplet $q'' \equiv (C_z''; S'') = \left(\frac{+1}{2}; \frac{1}{2}\right)$, as indicated in the third column of (C3a). The QSpace representation of $\hat{\psi}_S^\dagger$ derives from the matrix-elements

$$\psi_S^\dagger \to \langle C_z S| \cdot \left((\hat{\psi}_S^\dagger)^{\left(\frac{+1}{2}; \frac{1}{2}\right)} |C_z' S'\rangle\right)$$

using the Wigner Eckart theorem as in Eq. (A43).

The operator index in the QSpace (C3a) is listed third, by convention. The two non-zero matrix elements of each $\hat{c}_\sigma^\dagger$ within the four-dimensional space of single fermionic site implies a total of four non-zero matrix elements in $\psi_S^\dagger$, all having norm 1, with one matrix-element being negative. These matrix elements can be directly identified in QSpace (C3a). Since the reduced matrix elements $\|\psi_S^\dagger\|$ and the CGC spaces are to be interpreted as tensor product, the $\sqrt{2}$ factors in the last line cancel. With $\hat{\psi}_S^\dagger$ representing non-hermitian operators, the first column $q \equiv (C_z; S)$ is in general different form the second column $q' \equiv (C_z'; S')$. Moreover, since $\hat{\psi}_S^\dagger$ creates one particle, the first column, for example, cannot contain the empty state $(\frac{-1}{2}; 0)$, while the second column cannot contain the double occupied state $(\frac{+1}{2}; 0)$.

In contrast, the QSpace representation of the IROP ψ_S, *i.e.* without the dagger, is shown in (C3b). Note that for $\hat{\psi}_S \equiv \{-\hat{c}_\downarrow; \hat{c}_\uparrow\}$ to be an irreducible operator as compared to $\hat{\psi}_S^\dagger = \{\hat{c}_\uparrow^\dagger, \hat{c}_\downarrow^\dagger\}$, the reverse order in spin and the minus sign in the first component is essential [see discussion along with Eqs. (A49)]. In terms of the QSpace (C3b), this leads to the extra minus signs in the *multiplet space* of the second row. Moreover, the z-labels of the operator $\hat{\psi}_S$ itself flipped sign *w.r.t.* $\hat{\psi}_S^\dagger$ as expected as it removes a particle rather than adding one [see the multiplet labels $q'' \equiv (C_z''; S'')$ in the third column of (C3b)]. This is to emphasize that the application of the Wigner Eckart theorem must be performed consistently, *i.e.* switching sides in the application of an operator as in $\langle C_z S| \cdot (\psi^\dagger |C_z' S'\rangle) = (\psi |C_z S\rangle)^\dagger \cdot |C_z' S'\rangle$ must be dealt with carefully.

b. Identity A-tensor for two fermionic sites

Consider the combination of two fermionic sites. Alluding to Fig. 1, let site 1 (2) be described by $|i\rangle$ ($|\sigma\rangle$), respectively, both representing a 4-dimensional state space $\{|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$ of their own. The decomposition of the combined space in terms of the overall symmetry $\text{U}(1)_{\text{charge}} \otimes \text{SU}(2)_{\text{spin}}$ is fully described by the rank-3 QSpace (C4).

Given the $\text{U}(1)_{\text{charge}} \otimes \text{SU}(2)_{\text{spin}}$ symmetries, the abelian charge quantum number C_z simply adds up, while for the $\text{SU}(2)$ spin symmetry, the usual $\text{SU}(2)$ addition algebra applies. The overall number of multiplets in the combined space q'' is given by the last number (index 3) in the dimensions specified with the multiplet space. The specific input combinations entering a certain

combined space q'' are easily verified. The $q'' = (-\frac{1}{2}; \frac{1}{2})$ sector, for example, derives from the two configurations $\{q, q'\} = \{(-\frac{1}{2}; 0), (0; \frac{1}{2})\}$ and $\{(0; \frac{1}{2}), (-\frac{1}{2}; 0)\}$. Therefore the dimension of the reduced multiplet space for this q'' is 2. Each of these multiplets has an internal z-space which is itself of dimension 2 [last column]. The combined total dimension of the $q'' = (\frac{-1}{2}; \frac{1}{2})$ sector is therefor given by the product $2 \cdot 2 = 4$. Consistently, the dimension of the two 4-dimensional sites combined add up correctly to 16 states total. That is, multiplying the last dimension in the reduced multiplet space with the last dimension in the combined CGC spaces for each block separated by horizontal lines, bearing in mind that the multiplet space and the CGC spaces are to be combined in a tensor-product, yields the overall dimension of the combined space, $1 \cdot 1 + 2 \cdot 2 + 3 \cdot 1 + 1 \cdot 3 + 2 \cdot 2 + 1 \cdot 1 = 16$.

The A-tensor in (C4) is an identity A-tensor, in that up to permutations, plain identity matrices are split-up on the reduced multiplet level. By considering, for example, the $q'' = \left(\pm\frac{1}{2}; \frac{1}{2}\right)$ symmetry sector in records $2 - 3$ or $8 - 9$ of the QSpace (C4), the multiplet space when viewed together, *i.e.* ignoring all brackets, resemble the structure of a 2-dimensional identity matrix. Similar so for the $q'' = (0; 0)$ space in records $4 - 6$, having essentially a 3-dimensional identity matrix in the multiplet space. Allowing for arbitrary unitaries in the multiplet space in QSpace (C4), this then becomes the most general unitary transformation of the product space of two fermionic sites that also respects the symmetries considered.

2. Fermionic sites in the presence of particle-hole symmetry

The tensors introduced in the previous section for $\text{U}(1)_{\text{charge}} \otimes \text{SU}(2)_{\text{spin}}$ symmetry will now be written more compactly still by assuming the stronger particle-hole $\text{SU}(2)_{\text{charge}}$ symmetry instead of the plain abelian $\text{U}(1)_{\text{charge}}$. The symmetry considered in the following is therefore $\text{SU}(2)_{\text{SC}}^{\otimes 2} \equiv \text{SU}(2)_{\text{spin}} \otimes \text{SU}(2)_{\text{charge}}$. The z-operator for charge, $\hat{C}_z$, is now complemented by the raising operator $\hat{C}_+$ for charge $\text{SU}(2)$ [*cf.* Eq. (A52)]. The combined symmetries are given by the multiplet label for both charge and spin $\text{SU}(2)$, *i.e.* the non-negative labels $q = (C, S)$ with the z-labels of the charge symmetry now also taken care of by the CGC spaces.

The basis for a single fermionic level given given $\text{SU}(2)_{\text{SC}}^{\otimes 2}$ symmetry has been introduced in Eq. (A62). Therefore the full space of the four states $\{0, \uparrow, \downarrow, \uparrow\downarrow\}$ can be reduced to the two symmetry multiplets

$$
\begin{array}{l|l}
\text{multiplet space} & \text{dimension} \\
|C, S\rangle & d_C \times d_S = d_{tot} \\
\hline
\left|\frac{1}{2}; 0\right\rangle \equiv \{|0\rangle, s|\uparrow\downarrow\rangle\} & 2 \times 1 = 2 \\
\left|0; \frac{1}{2}\right\rangle \equiv \{|\uparrow\rangle, |\downarrow\rangle\} & 1 \times 2 = 2
\end{array}
\tag{C7}
$$

The most general scalar operator such as the Hamiltonian

P7.

Table II. Example QSpaces in the presence of $\mathrm{SU(2)}_{\mathrm{SC}}^{\otimes 2} \equiv \mathrm{SU(2)}_{\mathrm{spin}} \otimes \mathrm{SU(2)}_{\mathrm{charge}}$ symmetry for a single fermionic site $(\hat{\psi}_{CS})$, and for the combination of the state space of two sites (A-tensor). The CGC spaces for $\mathrm{SU(2)}_{\mathrm{charge}}$ and $\mathrm{SU(2)}_{\mathrm{spin}}$ are indicated by the C and S, respectively. The record index ν as well as the explicit specification of the dimensions are just added for clarity. For comparison, Tbl. I shows the same QSpaces for the case where the particle-hole symmetry is reduced to abelian charge conservation.

$$
\psi_{\mathrm{CS}} \equiv \left\{
\begin{array}{c|ccc|cc|cccc}
\text{record} & & & & \text{red. matrix elements} & & \text{CGC spaces} & & & \text{combined} \\
\text{index } \nu & (C;S) & (C';S') & (C''_z;S'') & \|\hat{\psi}_{\mathrm{CS}}\| & \text{dimension} & C & S & & \text{dimension} \\
\hline
1. & 0;\tfrac{1}{2} & \tfrac{1}{2};0 & \tfrac{1}{2};\tfrac{1}{2} & s\sqrt{2} & 1\times1\times1 & \tfrac{1}{\sqrt{2}}\{[0\;1];[-1\;0]\} & \{[1;0];[0;1]\} & & 2\times2\times4 \\
2. & \tfrac{1}{2};0 & 0;\tfrac{1}{2} & \tfrac{1}{2};\tfrac{1}{2} & \sqrt{2} & 1\times1\times1 & \{[1;0];[0;1]\} & \tfrac{1}{\sqrt{2}}\{[0\;1];[-1\;0]\} & & 2\times2\times4
\end{array}
\right\} \quad \text{(C5)}
$$

QSpace of spinor $\hat{\psi}_{\mathrm{CS}}$ defined in Eq. (A59) with the reduced matrix elements for given symmetries already calculated in Eq. (A63). The operator index within the IREP is listed third, as usual. The alternating sign s required with particle-hole symmetry appears with the reduced matrix elements in the first record only [*cf.* Eq. (A63); note that the same sign s is also picked up by the double occupied state, *cf.* Eq. (A62) or (C7)].

$$
A \equiv \left\{
\begin{array}{c|ccc|cc|ccc}
\text{record} & \text{site 1} & \text{site 2} & \text{combined} & \text{multiplet space} & & \text{CGC spaces} & & \text{combined} \\
\text{index } \nu & (C;S) & (C';S') & (C'';S'') & \|A\| & \text{dimension} & C & S & \text{dimension} \\
\hline
1. & 0;\tfrac{1}{2} & 0;\tfrac{1}{2} & 0;0 & \{[1],[0]\} & 1\times1\times2 & 1. & \tfrac{1}{\sqrt{2}}[0\;-1;\;0\;1] & 2\times2\times1 \\
2. & \tfrac{1}{2};0 & \tfrac{1}{2};0 & 0;0 & \{[0],[1]\} & 1\times1\times2 & \tfrac{1}{\sqrt{2}}[0\;-1;\;0\;1] & 1. & 2\times2\times1 \\
\hline
3. & 0;\tfrac{1}{2} & 0;\tfrac{1}{2} & 0;1 & 1. & 1\times1\times1 & 1. & \begin{array}{c}\{[1\;0;\;0\;0],\\ \tfrac{1}{\sqrt{2}}[0\;1;\;0\;1],\\ [0\;0;\;0\;1]\}\end{array} & 2\times2\times3 \\
\hline
4. & 0;\tfrac{1}{2} & \tfrac{1}{2};0 & \tfrac{1}{2};\tfrac{1}{2} & \{[1],[0]\} & 1\times1\times2 & \{[1\;0],[0\;1]\} & \{[1;0],[0;1]\} & 2\times2\times4 \\
5. & \tfrac{1}{2};0 & 0;\tfrac{1}{2} & \tfrac{1}{2};\tfrac{1}{2} & \{[0],[1]\} & 1\times1\times2 & \{[1;0],[0;1]\} & \{[1\;0],[0\;1]\} & 2\times2\times4 \\
\hline
6. & \tfrac{1}{2};0 & \tfrac{1}{2};0 & 1;0 & 1. & 1\times1\times1 & \begin{array}{c}\{[1\;0;\;0\;0],\\ \tfrac{1}{\sqrt{2}}[0\;1;\;0\;1],\\ [0\;0;\;0\;1]\}\end{array} & 1. & 2\times2\times3
\end{array}
\right\} \quad \text{(C6)}
$$

QSpace of identity A-tensor combining two fermionic sites – the symmetry records are sorted with respect the combined symmetry $q'' \equiv (C'',S'')$, where groups with the same (C'',S'') are separated by horizontal lines for clarity. Considering the tensor-product of multiplet and CGC spaces, the combined space has total dimension of $2\cdot1 + 1\cdot3 + 2\cdot4 + 1\cdot3 = 16$ as expected for two spinful fermionic levels. Compared to the A-tensor in (C4) with abelian charge conservation, the number of combined symmetry sectors has been further reduced from 6 to 4 [*i.e.* number of horizontally separated groups sharing the same q''], with an overall reduction in the number of multiplets present in the QSpace reduced from 10 to 6 [having $1+2+3+1+2+1 = 10$ in (C4), and here $2+1+2+1 = 6$].

is given by the QSpace,

$$
H \equiv \left\{
\begin{array}{cc|c|cc}
(C;S) & (C';S') & \|H\| & \multicolumn{2}{c}{\text{CGC spaces}} \\
\hline
\tfrac{1}{2};0 & \tfrac{1}{2};0 & h_{\frac{1}{2},0} & \mathbf{1}^{(2)} & 1. \\
0;\tfrac{1}{2} & 0;\tfrac{1}{2} & h_{0,\frac{1}{2}} & 1. & \mathbf{1}^{(2)}
\end{array}
\right\}. \quad \text{(C8)}
$$

Thus only the two reduced matrix elements $h_{1/2,0}$ and $h_{0,1/2}$ are left free to choose without compromising $\mathrm{SU(2)}_{\mathrm{SC}}^{\otimes 2}$ symmetry.

A non-scalar IROP is given in terms of the 4-element spinor $\hat{\psi}_{\mathrm{CS}}$ in Eq. (A59), which combines two creation and two annihilation operators. Its symmetries have been identified in Eq. (A61), leading to the IROP $\hat{\psi}_{\mathrm{CS}}$ with

the QSpace presented in (C5). Note that the size of the third dimension is 4 [see combined CGC dimension in the last column], consistent with the four operators that constitute the IROP.

The QSpace representation for the A-tensor combining two fermionic levels is given in (C6). The standard SU(2) addition rules are quickly confirmed. For example, the combined symmetry $q'' = (0,0)$ [records 1-2] can result from two combinations, namely $(\tfrac{1}{2},0)\otimes(\tfrac{1}{2},0)$, or $(0,\tfrac{1}{2})\otimes(0,\tfrac{1}{2})$, leading to a two dimensional multiplet space. All of this is transparently encoded in the QSpace. Here, QSpace (C6) again shows an identity A-tensor, as seen, for example, in the combined reduced multiplet space of

P7.

	multiplet space				dim.	CG dimension	
index	$\lvert S,$	$C_1,$	$C_2,$	$C_3 \rangle$		$d_S d_{C_1} d_{C_2} d_{C_3} =$	d_{tot}
1.	$\lvert 0;$	$0;$	$0;$	$\frac{1}{2} \rangle$	1	$1 \times 1 \times 1 \times 2 =$	2
2.	$\lvert 0;$	$0;$	$\frac{1}{2};$	$0 \rangle$	1	$1 \times 1 \times 2 \times 1 =$	2
3.	$\lvert 0;$	$\frac{1}{2};$	$0;$	$0 \rangle$	1	$1 \times 2 \times 1 \times 1 =$	2
4.	$\lvert 0;$	$\frac{1}{2};$	$\frac{1}{2};$	$\frac{1}{2} \rangle$	1	$1 \times 2 \times 2 \times 2 =$	8
5.	$\lvert \frac{1}{2};$	$0;$	$0;$	$0 \rangle$	2	$2 \times 1 \times 1 \times 1 =$	2
6.	$\lvert \frac{1}{2};$	$0;$	$\frac{1}{2};$	$\frac{1}{2} \rangle$	1	$2 \times 1 \times 2 \times 2 =$	8
7.	$\lvert \frac{1}{2};$	$\frac{1}{2};$	$0;$	$\frac{1}{2} \rangle$	1	$2 \times 2 \times 1 \times 2 =$	8
8.	$\lvert \frac{1}{2};$	$\frac{1}{2};$	$\frac{1}{2};$	$0 \rangle$	1	$2 \times 2 \times 2 \times 1 =$	8
9.	$\lvert 1;$	$0;$	$0;$	$\frac{1}{2} \rangle$	1	$3 \times 1 \times 1 \times 2 =$	6
10.	$\lvert 1;$	$0;$	$\frac{1}{2};$	$0 \rangle$	1	$3 \times 1 \times 2 \times 1 =$	6
11.	$\lvert 1;$	$\frac{1}{2};$	$0;$	$0 \rangle$	1	$3 \times 2 \times 1 \times 1 =$	6
12.	$\lvert \frac{3}{2};$	$0;$	$0;$	$0 \rangle$	1	$4 \times 1 \times 1 \times 1 =$	4

Table III. State space of 3-channel site with $\mathrm{SU}(2)_{\mathrm{SC}}^{\otimes 4} \equiv \mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)_{\mathrm{charge}}^{\otimes 3}$ symmetry.

	multiplet space			dim.	CG dimension	
index	$\lvert S$	C_z	$T \rangle$		$d_S \cdot d_{C_z} \cdot d_T =$	d_{tot}
1.	$\lvert 0;$	$-\frac{3}{2};$	$0\ 0 \rangle$	1	$1 \times 1 \times 1 =$	1
2.	$\lvert 0;$	$-\frac{1}{2};$	$2\ 0 \rangle$	1	$1 \times 1 \times 6 =$	6
3.	$\lvert 0;$	$\frac{1}{2};$	$0\ 2 \rangle$	1	$1 \times 1 \times 6 =$	6
4.	$\lvert 0;$	$\frac{3}{2};$	$0\ 0 \rangle$	1	$1 \times 1 \times 1 =$	1
5.	$\lvert \frac{1}{2};$	$-1;$	$1\ 0 \rangle$	1	$2 \times 1 \times 3 =$	6
6.	$\lvert \frac{1}{2};$	$0;$	$1\ 1 \rangle$	1	$2 \times 1 \times 8 =$	16
7.	$\lvert \frac{1}{2};$	$1;$	$0\ 1 \rangle$	1	$2 \times 1 \times 3 =$	6
8.	$\lvert 1;$	$-\frac{1}{2};$	$0\ 1 \rangle$	1	$3 \times 1 \times 3 =$	9
9.	$\lvert 1;$	$\frac{1}{2};$	$1\ 0 \rangle$	1	$3 \times 1 \times 3 =$	9
10.	$\lvert \frac{3}{2};$	$0;$	$0\ 0 \rangle$	1	$4 \times 1 \times 1 =$	4

Table IV. State space of 3-channel site with $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ symmetry.

	multiplet space		dim.	CG dimension	
index	$\lvert S$	$\mathrm{Sp}(6) \rangle$		$d_S \cdot d_{C_z} \cdot d_T =$	d_{tot}
1.	$\lvert 0;$	$0\ 0\ 1 \rangle$	1	$1 \times 1 \times 14 =$	14
2.	$\lvert \frac{1}{2};$	$0\ 1\ 0 \rangle$	1	$2 \times 1 \times 14 =$	28
3.	$\lvert 1;$	$1\ 0\ 0 \rangle$	1	$3 \times 1 \times 6 =$	18
4.	$\lvert \frac{3}{2};$	$0\ 0\ 0 \rangle$	1	$4 \times 1 \times 1 =$	4

Table V. State space of 3-channel site with $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{Sp}(6)$ symmetry.

records $1 - 2$, or $4 - 5$).

3. Three channels with $\mathrm{SU}(3)$ channel symmetry

Consider a system with three spinful particle-hole symmetric channels, as introduced in Sec. A 9 d. A single site then has a full Hilbert space of dimension $4^3 = 64$. Three

symmetry settings are analyzed: a set of plain $\mathrm{SU}(2)$ symmetries, a combination with the $\mathrm{SU}(3)$ channel symmetry, and finally the largest symmetry present which includes the enveloping symplectic symmetry $\mathrm{Sp}(6)$. Since the QSpaces in given context are extensive, a more compact comparison of these symmetry settings in the numerical context is given, instead.

The first setting, $\mathrm{SU}(2)_{\mathrm{SC}}^{\otimes 4} \equiv \mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)_{\mathrm{charge}}^{\otimes 3}$, is based on four independent $\mathrm{SU}(2)$ symmetries. The 64-dimensional state space of single site decomposes into 13 multiplets in 12 symmetry sectors, as listed in Tbl. III. All of these contain a single representative multiplet, except for the space $\lvert \frac{1}{2}; 0; 0; 0 \rangle$ in row 5, which contains two multiplets. In contrast, using the $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ symmetry, instead, the 64 states of the three fermionic levels decomposes into the 10 multiplet spaces listed in Tbl. IV. Thus compared to the $\mathrm{SU}(2)_{\mathrm{CS}}^{\otimes 4}$ symmetry setting in Tbl. III, the number of multiplet spaces is further reduced with all multiplet spaces containing a single multiplet only. This suggests that the latter symmetry including the channel $\mathrm{SU}(3)$ is somewhat more efficient as it allows to compactify multiplet spaces more strongly. Given the multiplet space in Tbl. IV, for example, the most general Hamiltonian in the 64×64 dimensional Hilbert space compatible with given symmetry consists of the 10 reduced matrix elements appearing in the multiplet space only.

A further strong boost in numerical efficiency can be obtained, if the Hamiltonian supports it, by combining the particle-hole symmetry of $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{SU}(2)_{\mathrm{charge}}^{\otimes 3}$ with the channel symmetry in $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ to their enveloping $\mathrm{Sp}(6)$ symmetry. The resulting state space for the state space of a single 3-channel site is given in Tbl. V. The 64×64 dimensional Hilbert space has thus been reduced to a total of four multiplets only.

All three symmetry settings have been successfully implemented within the NRG framework. By starting with a single site [i.e. the basic fermionic three-level unit as introduced in Sec. A 9 d], and iteratively adding a site within the NRG, new multiplet spaces are quickly explored and built up within the first few NRG iterations. In practice, the CGC spaces of newly generated multiplets are also stored for latter retrieval. Once truncation of the state space within NRG sets in, however, the generation of IREPs eventually saturates to within a finite range of multiplet spaces.

The resulting space for adding up to three further sites without truncation is indicated in Tbl. VI. With reasonable numerical resources, it is feasible within the $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{U}(1)_{\mathrm{charge}} \otimes \mathrm{SU}(3)_{\mathrm{channel}}$ (second) or $\mathrm{SU}(2)_{\mathrm{spin}} \otimes \mathrm{Sp}(6)$ (third) setting, to keep all states up to three sites total within the NRG, first truncating only when a fourth site is added. This leads to $N_{\mathrm{K}} = 9,086$ [$N_{\mathrm{K}} = 1,232$] kept multiplets for the second and third symmetry setting, respectively. The corresponding memory requirements for a general basis transformation for adding another site (A-tensor) then amounts to about

sites	abelian dim.	SU(2)$_\text{spin}$ ⊗ SU(2)$^{\otimes 3}_\text{charge}$					SU(2)$_\text{spin}$ ⊗ U(1)$_\text{charge}$ ⊗ SU(3)$_\text{channel}$					SU(2)$_\text{spin}$ ⊗ Sp(6)				
n	$D^* = 64^n$	N_S	D	D^*/D	memory	MEM*	N_S	D	D^*/D	memory	MEM*	N_S	D	D^*/D	memory	MEM*
1	64	12	13	4.9	< 18 K		10	10	6.4	< 13 K		4	4	16	< 6 K	
2	4,096	61	388	10.6	528 K	>8.7 M	69	260	15.8	359 K	>12 M	23	61	67.1	162 K	>34 M
3	262,144	192	14,229	18.4	27 M	>23 G	226	9,086	28.9	11 M	>31 G	60	1,232	213	7 M	>112 G
4	16,777,216	469	590,856	28.4	24 G	>65 T	565	366,744	45.7	6.8 G	>85 T	132	31,640	530	334 M	>355 T

Table VI. Comparison of different symmetry scenarios for the same underlying physical system of a symmetric 3-channel setup, analyzing the product spaces of up to $n = 4$ sites. Each site represents a Hilbert space of dimension 64, thus n sites amounts to an overall Hilbert space of dimension $D^* = 64^n$ [second column]. This state space can be decomposed into D multiplets in N_S symmetry sectors using an A-tensor for the addition of every new site. These A-tensors are encoded in terms of QSpaces. The total memory requirement for each A-tensor is listed, given sparse CGC representation. In addition, as a comparison to a fully abelian setting, MEM* indicates the memory that had been required if the tensor products between reduced multiplets and CGC spaces was carried out explicitly [K,M,G,T for kilo-, mega-, giga-, and tera-bytes, respectively].

sites	SU(2)$_\text{spin}$ ⊗ SU(2)$^{\otimes 3}_\text{charge}$						SU(2)$_\text{spin}$ ⊗ U(1)$_\text{charge}$ ⊗ SU(3)$_\text{channel}$					SU(2)$_\text{spin}$ ⊗ Sp(6)			
n	d_S	d_C	d_C	d_C	sparsity	CGS/A	d_S	$d_{U(1)}$	d_C	sparsity	CGS/A	d_S	$d_\text{Sp(6)}$	sparsity	CGS/A
1	4	2	2	2	0.36	0.8	4	1	8	0.12	0.8	4	14	0.027	0.90
2	7	3	3	3	0.36	0.8	7	1	27	0.12	0.8	7	126	0.027	0.90
3	10	4	4	4	0.28	0.11	10	1	64	0.064	0.33	10	616	0.011	0.94
4	13	5	5	5	0.22	$< 10^{-3}$	13	1	125	0.039	0.003	13	2457	0.006	0.55

Table VII. Comparison of different symmetry scenarios as in Tbl. VI in terms of (i) largest multiplet dimension d for each individual symmetry, and (ii) overall average sparsity of the CGC spaces, *i.e.* the number of non-zero elements divided by the total number of matrix elements. The last columns for each symmetry (CGS/A) shows the memory requirement of all sparse CGC spaces in a given QSpace A_n relative to the entire QSpace.

7 G [0.3 G]. The corresponding full NRG iteration for adding another site then takes several hours on a state-of-the-art 8-core workstation.

The same calculation, however, gets quickly impossible as fewer symmetries are available or used in the actual computation as can be seen from Tbl. VI. For example, if only the abelian part of the symmetry had been accounted for in the computation, the corresponding memory requirement can be estimated by considering the explicit tensor product of the multiplet space with the CGC spaces, leading in terms of the SU(2)$^{\otimes 4}_\text{SC}$ setting to about 23 G and 65 T(!) for $n = 3$ and $n = 4$, respectively, the latter being completely hopeless in practice. The explicit treatment of non-abelian symmetries, however, clearly makes the latter case feasible with a reasonable amount of numerical resources.

a. SU(3) *symmetry*

The irreducible SU(3) multiplets generated in the actual NRG run using SU(2)$_\text{spin}$ ⊗ U(1)$_\text{charge}$ ⊗ SU(3)$_\text{channel}$ symmetry as presented in the main text [*cf.* Sec. V], are shown in terms of their weight diagrams in Fig. 12. For comparison, the SU(3) IREPs present in the description of a single site are $q^\text{SU(3)} \equiv (q_1, q_2) \in \{(0,0),\ (0,1),\ (0,2),\ (1,0),\ (2,0)\}$, *cf.* Tbl. IV. The apparent symmetry *w.r.t.* to flipping the quantum numbers in (q_1, q_2) is also reflected in the overall set of multiplets

generated within the NRG. As seen in Fig. 12, all IREPs (q_1, q_2) with $q_1 + q_2 \leq 8$ are present, except for $(0, 8)$ and $(8, 0)$.

Inner multiplicity, as expected for SU(3), is clearly present and depicted in the weight diagrams of Fig. 12 by the encircled set of points. There the number of points inside a circle stands for the multiplicity of the corresponding z-labels in the multiplet. Inner multiplicity decreases in shells as one moves outward the multiplet, which is seen particularly well for the multiplets $q_1 = q_2$. The states on the outer circumference have no multiplicity, *i.e.* have unique z-labels, as expected. This demonstrates the uniqueness of the maximum weight state [cf. Sec. A 3 c], which was required for the numerical state space decomposition in Sec. B 1.

Due to the two-dimensional label structure of the SU(3) multiplets together with inner multiplicity, their internal dimensions can get significantly larger as compared to SU(2) multiplets. The largest SU(3) multiplet $(4, 4)$ encountered in the actual NRG run presented in Fig. 12, for example, has an internal irreducible dimension of $d = 125$ [see also Tbl. VII]. This implies, for example, that with respect to the diagonalization of a Hamiltonian, a 125-fold degeneracy has been reduced to a *single* multiplet. In contrast, the multiplet structure for SU(2)$^{\otimes 4}_\text{SC}$ is clearly weaker as it only includes SU(2) symmetries. There the largest quantum numbers encountered in an NRG run with comparable number of kept states includes $S \leq 6$ in the spin sector, leading to an
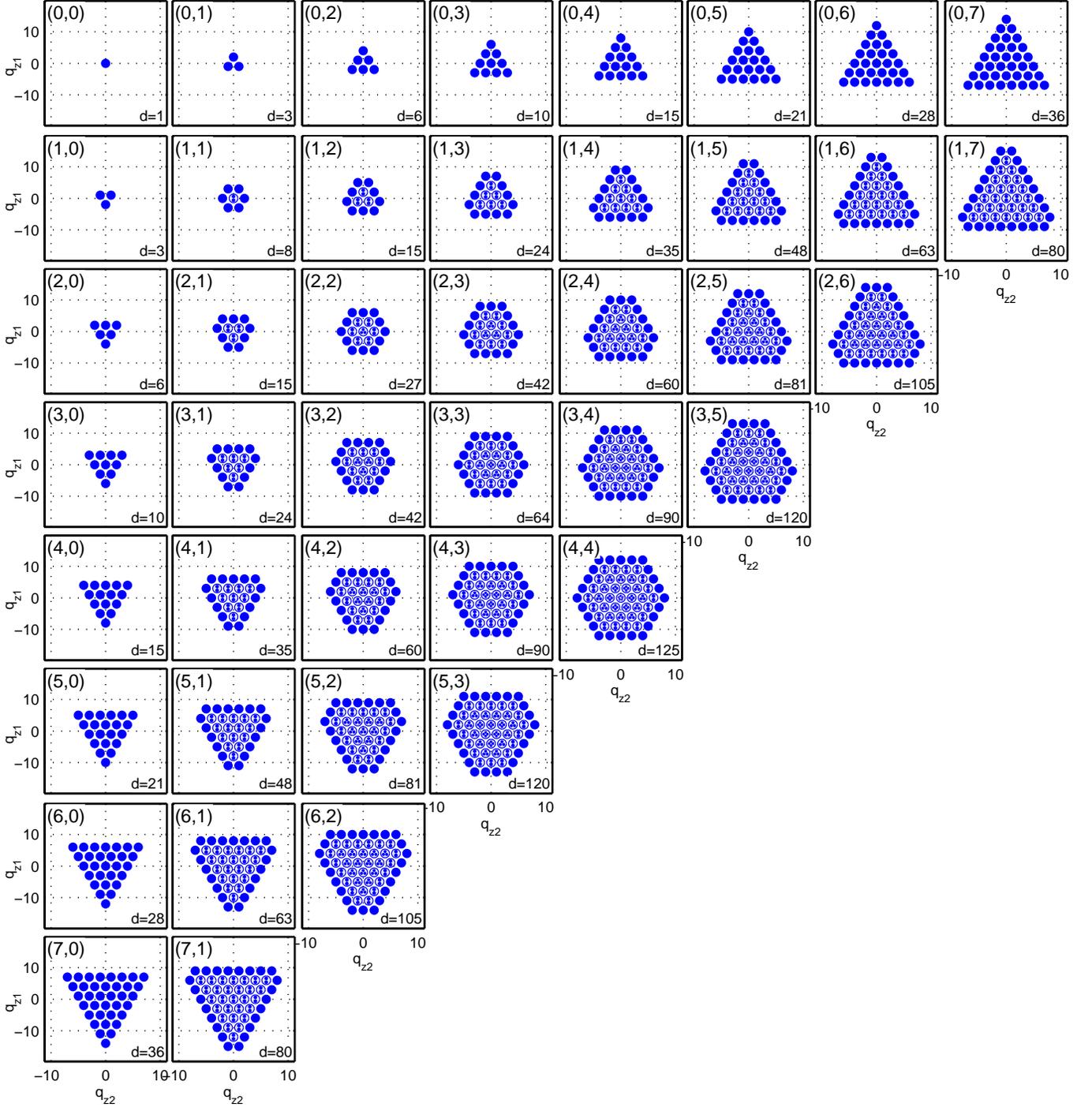
Figure 12. Weight diagrams of SU(3) multiplets generated in a typical NRG run for the symmetric 3-channel system including $SU(2)_{\text{spin}} \otimes U(1)_{\text{charge}} \otimes SU(3)_{\text{channel}}$ symmetries ($\Lambda = 4$, $E_{\text{trunc}} = 7$). The multiplet label (q_1, q_2), as defined in Eq. (A34), is specified with each multiplet in the upper left corner of its panel. For a weight diagram of a specific IREP (q_1, q_2), the corresponding z-labels (q_{z1}, q_{z2}) of SU(3) for each individual state within the multiplet are depicted as points in a two-dimensional plot. In case of inner multiplicity, *i.e.* that several states within the same IREP share exactly the same z-labels, these states are shown as an encircled group of smaller points. The dimension for every multiplet (number of points drawn within a panel) is indicated to the lower right of each panel. The first panel [multiplet $(0,0)$] represents the scalar representation with multiplet dimension $d = 1$. Multiplet $(1,0)$ represents the defining three-dimensional representation [*cf.* Sec. A 4 c], and $(1,1)$ the regular representation of dimension 8 equal to the dimension of SU(3), *i.e.* the number of its generators. The largest multiplet encountered in given NRG run is the multiplet $(4,4)$ with an irreducible dimension of $d = 125$.

Sp(6) multiplet	dimension	Sp(6) multiplet	dimension
q	d	[cont'd]	
(0 0 0)	1	(1 2 0)	350
(1 0 0)	6	(1 0 2)	378
(0 0 1)	14	(0 3 0)	385
(0 1 0)	14	(3 1 0)	448
(2 0 0)	21	(1 1 1)	512
(3 0 0)	56	(3 0 1)	525
(1 1 0)	64	(0 1 2)	594
(1 0 1)	70	(0 2 1)	616
(0 0 2)	84	(2 2 0)	924
(0 2 0)	90	(2 0 2)	1078
(0 1 1)	126	(1 3 0)	1344
(4 0 0)	126	(2 1 1)	1386
(2 1 0)	189	(1 2 1)	2205
(2 0 1)	216	(1 1 2)	2240
(0 0 3)	330		

Table VIII. Sp(6) multiplets generated in a fully converged NRG run for the symmetric 3-channel system using $SU(2)_{\mathrm{spin}} \otimes Sp(6)$ ($\Lambda = 4$, $E_{\mathrm{trunc}} = 7$). Multiplet $(0,0,0)$ represents the scalar representation of dimension 1, multiplet $(1,0,0)$ the defining representation of dimension 6, and multiplet $(2,0,0)$ the regular representation of dimension 21 which is also equal to the number of generators for Sp(6). The largest tensor-product decomposition was between the product spaces of IREPs of dimension 14 and 512, yielding a combined product space dimension of $7,168$. Run time of the bare NRG run was about 2 hrs on a state-of-the-art 8-core workstation with moderate memory requirements of $\lesssim 4.5$ G.

individual multiplet dimension of at most 13 [see also Tbl. VII]. In the overall combination of the symmetries, this implies that for comparable number of states, *i.e.* for a comparable accuracy within the NRG, on average about 50% more multiplets need to be kept within the $SU(2)_{\mathrm{SC}}^{\otimes 4}$ setting as compared to the case when SU(3) is included [see Fig. 9 in the main text].

Finally, the individual weight diagrams in Fig. 12 show well-known symmetries, such as a reflection symmetry of each diagram around the vertical y-axes, or the reflection symmetry between the multiplets (q_1, q_2) and (q_2, q_1) around the horizontal axis. These *Weyl symmetries* may be used to evaluate or encode CGC spaces more efficiently.[26] For the purpose of this paper, however, these symmetries were not exploited, given also that the pure numerical evaluation of the CGC spaces as outlined earlier was already sufficiently fast.

b. Sp(6) *symmetry*

The complete set of Sp(6) symmetries generated in the fully converged NRG run (using $\Lambda = 4$ and $E_{\mathrm{trunc}} = 7$ as used in the results in the main text), is listed in Tbl. VIII. All multiplets had been generated within the first four Wilson shells. The fact that the symmetry Sp(6) fully incorporates non-abelian particle-hole and channel symmetry, manifests itself by observing that all eigenenergies in the multiplet spaces are now strictly *non-degenerate* throughout an entire NRG calculation. *Huge* degeneracies of several thousands can be split off in terms of tensor products with Sp(6) multiplets.

For given model, the symmetry Sp(6) in fact also allowed to reduce the rather coarse discretization of $\Lambda = 4$ in the NRG calculation underlying Tbl. VIII. For comparison, if $\Lambda = 2$ is used, instead, while keeping the same $E_{\mathrm{trunc}} = 7$, it turns out, the largest multiplet generated is $(2,1,2)$ of dimension 5720. The largest intermediate product space to be decomposed into IREPs becomes as large as $14 \times 1386 = 19,404$. Having $\Lambda = 2$, this required twice the Wilson chain length for the same range in energy scales, leading to an overall run time of the entire NRG run of about 32 hrs with still reasonably manageable memory requirements of $\lesssim 20$ G.

As a rough general estimate, typical multiplet dimensions, as they occurred in practice, scale like 10^r where r is the rank of the symmetry. For $SU(2)_{\mathrm{spin}}$, this implies multiplets of dimension $\lesssim 10$, for the $SU(4)_{\mathrm{channel}}$ symmetry, indeed, one had multiplets of dimension of $\lesssim 100$, while now for Sp(6), a symmetry of rank 3, one easily reaches multiplet dimensions on the order of a few 1000 (*cf.* Tbl. VIII). Therefore with increasing rank of the symmetry, the numerical effort strongly shifts from the multiplet space to the CGC spaces. For sets of smaller symmetries with rank $r \leq 2$ this leads to a strong gain in numerical efficiency, while the numerical overhead for the CGC spaces remains negligible. Reaching symmetries of rank 3, such as Sp(6), the numerical effort within the CGC spaces can now become comparable to or even larger than the operations on the higher multiplet level.

Table VII summarizes the situation by comparing the maximal multiplet spaces with the corresponding sparsity and memory requirements of the CGC spaces for the first few A-tensors, when combining up to $n = 4$ sites without truncation. As the internal multiplet dimensions quickly grow for higher rank symmetries, nevertheless only an ever smaller fraction of Clebsch-Gordan coefficients are non-zero. As seen from Tbl. VII, the sparsity roughly grows exponentially with the rank of the symmetry. Nevertheless, with the memory requirement of the *sparse* CGC for Sp(6) comparable or even larger than the storage of the reduced matrix elements on the higher multiplet level (see last column in Tbl. VII), full storage including also the zero CGC spaces would have extremely inflated overall storage requirement. In this sense, here sparse storage of CGC spaces becomes mandatory.

In order to distinguish numerical noise, *i.e.* negligible CGC matrix elements, from actual matrix elements then, this requires an accurate evaluation of the CGC matrix elements. Double precision accuracy as compared to the *exact theoretical* CGC coefficients was sufficient, in practice. In particular, this implies, that also the matrix elements of the generators for given IREPs of the symmetry

are known essentially *exactly* at any step. In the iterative approach, however, when new multiplets are generated through tensor products with smaller entities, numerical error can easily pile up. In this case, the accuracy of the matrix elements of the generators must be better than double precision. For practical purposes, quad precision on matrix elements of the generators turned out sufficient.

Finally, for example, with standard CGC spaces of rank-3, the sparse storage of CGC spaces also requires the generalization of standard sparse storage and sparse operations to arbitrary rank tensors. All of these is achieved, in general, by proper efficient permutations in sparse index space which requires a fast sorting scheme, together with reshaping of higher rank-objects to standard two-dimensional sparse objects with fused indices, since this allows to employ fast routines for standard sparse multiplication.

P7.