Beyond the Landau-Ginzburg-Wilson paradigm: Analytical studies of non-standard quantum criticality

Dimitri Pimenov

Munich 2019

Beyond the Landau-Ginzburg-Wilson paradigm: Analytical studies of non-standard quantum criticality

Dimitri Pimenov

A dissertation submitted

to the Faculty of Physics at the

Ludwig–Maximilians–Universität München

for the degree of

DOCTOR RERUM NATURALIUM

Munich, July 29, 2019

Summary

The Landau-Ginzburg-Wilson (LGW) paradigm is the backbone of the modern understanding of critical phenoma. It rests on the assumption that a continuous phase transition can be described solely in terms of a fluctuating order parameter. In this sandwich thesis, we analyze several quantum phase transitions in two-dimensional systems where this basic assumption is violated. Our analysis is structured in three parts:

In the first part, two ordering transitions in metals are studied, where the presence of gapless fermionic modes invalidates the pure order parameter description. Instead, we apply a renormalization group approach which retains both fermionic and bosonic (order paramater) degrees of freedom, using the deviation from an upper critical dimension of $d = 5/2$ as a control parameter. We first apply this technique to describe the quantum phase transition between a normal metal and an inhomogeneous (FFLO) superconductor. Our analysis rigorously confirms the mean field expectation that this transition is continuous, and shows that interesting non-Fermi liquid physics can arise at the critical point, manifesting itself in unusual scaling of various observables. The second case study is the onset of incommensurate $2k_F$ charge density wave order. While non-Fermi liquid features are less pronounced at this transition, we find a strong dynamical nesting of the Fermi surface, which stabilizes the density wave formation and results in a continuous transition, opposed to early theoretical claims.

In the second part, we present a novel case study of "deconfined criticality" for a quantum magnet: Here, symmetries are broken on both sides of the phase transition, which would be first order within a conventional two-order-parameter LGW description. By contrast, the scenario of deconfined criticality predicts a continuous transition, driven by condensation of topological defects which carry quantum numbers of the opposite phase. Its paradigmatic application is the transition from a $SU(2)$-Neél state to a valence-bond solid (a singlet covering configuration that breaks spatial symmetries) on the square lattice. Here, we propose a novel extension and study a $SU(3)$ antiferromagnet on a triangular lattice, which supports a transition between a magnetic, three-sublattice color-ordered phase and a trimerized $SU(3)$ singlet phase. We provide a critical theory in terms of fractional bosonic degrees of freedom (and with a topological defect interpretation), and study its fixed point properties with functional renormalization group. This yields a critical fixed point in a suitable large-N-limit, implying a continuous transition in accordance with the deconfined criticality framework.

In the third and final part, we study the "polaron"-problem of a single impurity coupled to a majority Fermi sea, which is particularly interesting when formation of a bound state ("molecule") between impurity and majority is allowed for. Starting from the exactly solvable limit of infinite impurity mass, we present a controlled computation of impurity spectra (both single- and two-particle) for heavy impurities based on Feynman diagram techniques, with an eye for experiments on doped semiconductors and ultracold gases. Furthermore, we discuss various aspects of the "molecule-to-polaron" transition that occurs in these systems, which also defies a simple LGW description in the single-impurity limit.

Zusammenfassung

Das Landau-Ginzburg-Wilson (LGW) Paradigma bildet den Grundpfeiler des modernen Verständnisses kritischer Phänomene. Es beruht auf der Annahme, dass ein kontinuierlicher Phasenübergang erschöpfend mittels eines fluktuierenden Ordnungsparameters beschrieben werden kann. In dieser kumulativen Arbeit analysieren wir mehrere Quantenphasenübergänge in zweidimensionalen Systemen, bei denen diese Grundannahme verletzt ist. Unsere Analyse ist in drei Teile gegliedert:

Im ersten Teil werden zwei Phasenübergänge in Metallen betrachtet, bei denen eine reine Ordnungsparameterbeschreibung aufgrund von fermionischen Moden ohne Energielücke fehlschlägt. Stattdessen verwenden wir eine kontrollierte Renormierungsgruppenmethode, in der sowohl fermionische als auch bosonische (Ordungsparameter-) Freiheitsgrade beibehalten werden. Der Kontrollparameter ist die Abweichung von der oberen kritischen Dimension bei $d = 5/2$. Zunächst wenden wir diese Methode auf den Quantenphasenübergang zwischen einem Normalleiter und einem inhomogenen (FFLO) Supraleiter an. Unsere Analyse bestätigt auf rigorose Weise, dass es sich um einen kontinuierlichen Phasenübergang handelt (wie auch von der Molekularfeldtheorie vorhergesagt), und zeigt interessante Nicht-Fermiflüssigkeits-Physik am kritischen Punkt auf. Diese äußert sich im ungewöhnlichen Skalierungsverhalten diverser Observablen. Weiterhin untersuchen wir das Auftreten von nicht-kommensurablen $2k_F-$Ladungsdichtewellen. An diesem Phasenübergang ist die Nicht-Fermiflüssigkeits-Physik weniger stark ausgeprägt. Stattdessen finden wir eine starke dynamische Abflachung der Fermiflächen, welche die Dichtewellenformation stabilisiert und zu einem kontinuierlichen Phasenübergang führt, in Widerspruch zu frühen theoretischen Arbeiten.

Im zweiten Teil stellen wir einen neuartigen Fall von "deconfined criticality" in Quantenmagneten vor: Hier werden Symmetrien auf beiden Seiten des Phasenübergangs gebrochen, was im Rahmen einer konventionellen LGW Beschreibung (unter Verwendung zweier Ordnungsparameter) zu einem unstetigen Phasenübergang führen würde. Im Gegensatz dazu sagt das "deconfined criticality"-Szenario einen kontinuierlichen Phasenübergang voraus. Dieser wird durch die Kondensation von topologischen Defekten bewirkt, die Quantenzahlen der gegenteiligen Phase tragen. Das Paradebeispiel ist der Übergang zwischen einem $SU(2)$-Neél Zustand und einem valence-bond-solid (einem Singlet-Bedeckungsmuster, welches räumliche Symmetrien bricht) auf dem Quadratgitter. Als Erweiterung betrachten wir hier einen $SU(3)$-Antiferromagneten auf einem Dreiecksgitter, wo ein Phasenübergang zwischen einer magnetischen, Farben-geordneten Phase mit einer dreifachen Untergitterstruktur, sowie einer trimerisierten $SU(3)$-Singlet-Phase stattfinden kann. Wir stellen eine kritische Theorie auf Basis von fraktionalisierten bosonischen Freiheitsgraden auf, für die auch eine Interpretation im Sinne topologischer Defekte möglich ist. Weiterhin untersuchen wir die Fixpunkteigenschaften dieser Theorie mit Hilfe der funktionellen Renormierungsgruppe. In einem geeigneten Grenzfall grosser N ergibt dies einen kritischen Fixpunkt und entsprechend einen kontinuierlichen Phasenübergang, in Einklang mit dem "deconfined criticality"-Konzept.

Im dritten und letzten Teil untersuchen wir das "Polaron"-Problem, in dem ein einzelnes Fremd-Teilchen an ein Fermi-Ensemble von Mehrheitsfermionen gekoppelt ist. Dieses Problem ist besonders interessant, wenn ein gebundener Zustand zwischen Fremd- und Mehrheitsteilchen ("Molekül") gebildet werden kann. Mittels Feynmandiagrammmethoden berechnen wir hier auf kontrollierte Weise Fremdteilchen-Spektra (sowohl Einteilchen- als

auch Zweiteilchenspektra) für schwere Fremdteilchen, indem wir vom exakt lösbaren Grenzfall unendlich schwerer Fremdteilchen ausgehen. Dies ist für die Beschreibung von Experimenten in Halbleitern und kalten Gasen relevant. Weiterhin diskutieren wir verschiedene Aspekte des "Molekül-nach-Polaron"-Übergangs in diesen Systemen, für den eine einfache LGW-Beschreibung aufgrund des Ein-Fremdteilchen-Grenzfalls unmöglich ist.

List of Publications

This dissertation is based on the following journal articles, listed in chronological order:

P1 *Deconfined Criticality in SU(3) antiferromagnets on the triangular lattice*
D. Pimenov and M. Punk
Sec. 4.2.2 / arXiv:1703.01308   Phys. Rev. B 95, 184427 (2017)

P2 *Fermi-edge exciton-polaritons in doped semiconductor microcavities with finite hole mass*
D. Pimenov, J. von Delft, L. Glazman, and M. Goldstein
Sec. 5.2.2 / arXiv:1707.08613   Phys. Rev. B 96, 155310 (2017)

P3 *Non-Fermi liquid at the FFLO quantum critical point*
D. Pimenov, I. Mandal, F. Piazza, and M. Punk
Sec. 3.2.2 / arXiv:1711.10514   Phys. Rev. B 98, 024510 (2018)

P4 *Spectra of heavy polarons and molecules coupled to a Fermi sea*
D. Pimenov and M. Goldstein
Sec. 5.3.2 / arXiv:1809.02577   Phys. Rev. B 98, 220302(R) (2018)

P5 *Incommensurate 2kF density wave quantum criticality in two dimensional metals*
J. Halbinger, **D. Pimenov**, and M. Punk
Sec. 3.3.2 / arXiv:1902.05084   Phys. Rev. B 99, 195102 (2019)

Contents

1 Introduction

Oftentimes, at the heart of a successful physical theory there is an ingenious and concise set of ideas, forming a "story" we can understand intuitively. These stories help us to navigate through the complexity of nature, and scientific progress is commonly made by spinning new stories from old ones.

The story told in the Landau theory of phase transitions [Lan37](1937) is an especially compelling and beautiful one, with symmetry and universality as its main protagonists. In short, it goes like this: When a continuous phase transition occurs in a physical system by cooling below a critical temperature T_c, the internal symmetry of the system is reduced (broken). Knowledge of the symmetries above and below T_c, encapsulated in a bosonic order parameter ϕ, uniquely specifies the properties of the phase transition, in particular the critical exponents. As a result, transitions in very diverse physical systems (e.g., the transition from normal to superfluid helium and from a para- to a ferromagnet) are captured by a single universal theory if their symmetries match.

In this simplicity, the story sounds almost too good to be true, and in fact some refinements have to be made. In the Landau theory the critical properties are independent of the dimensionality, which matches poorly e.g. to what is known from the 1D and 2D Ising models – the former shows no phase transition, while the latter does. This oversimplification of Landau theory is due to the neglect of local fluctuations of the order parameter, and include them we must replace $\phi \to \phi(\mathbf{x})$. Thus, Landau theory amounts to a mean field treatment, which nevertheless can yield accurate results depending on microscopic parameters, e.g. for the normal-metal-to-superconductor transition. A simple way to estimate the strength of the fluctuations and thus to assess the validity of the Landau approach is given by the Ginzburg criterion [Lev59, Gin61], formulated in the late fifties.

A systematic study of these fluctuations required a new framework, the renormalization group (RG), whose development was spearheaded by Wilson [WK74] in the seventies, building mainly on prior work of Kadanoff [Kad66]. The conglomerate of the Landau-Ginzburg order parameter description and the Wilsonian RG treatment is known as the **Landau-Ginzburg-Wilson** (LGW) paradigm, which is the backbone of the modern understanding of critical phenomena.

Soon after the RG development, a new testing ground for the LGW paradigm came about through the advent of quantum phase transitions (QPTs) [Sac11]: these occur at $T = 0$ upon variation of a non-thermal control parameter (like pressure or magnetic field), and are dominated by quantum fluctuations. To describe QPTs, one must allow for dynamical order parameters, $\phi(\mathbf{x}) \to \phi(\tau, \mathbf{x})$, which brings about a new critical exponent z. Modified in this way, the LGW paradigm often still upholds, e.g. when dealing with "quantum-rotor models" which provide a description of ordering transitions in insulators (see [Sac11] for a detailed discussion, including references to experiments).

Alas, sometimes even the best story comes to an end, and so various exceptions to the LGW paradigm have been discovered. The exceptions mostly occur in 1D/2D, where quantum fluctuations are especially severe. Focusing on 2D, an example for a continuous non-LGW transition in the classical realm is the Kosterlitz-Thouless transition [Kos74] in the 2D XY model. Here, different phases are not distinguished by an order-parameter, but rather by qualitatively distinct correlations. As we will discuss in this thesis, for many relevant QPTs

the basic assumptions of the LGW paradigm can fail as well. Based on five publications, we will shed light on three instances of LGW violation in two-dimensional systems:[1]

First, in ordering transitions of two-dimensional metals the presence of gapless fermions invalidates the pure order parameter description, and one must retain both fermionic and bosonic degrees of freedom. The associated QPTs are of relevance for many correlated materials – doped cuprates, heavy fermion materials, organic superconductors, and transition metal dichalcogenides, just to name a few. Consequently, their proper theoretical description could contribute to the solution of long-standing theoretical mysteries such as high-T_c superconductivity and strange metallicity.

Another instance of LGW violation arises in insulating quantum magnets, which can be described by antiferromagnetic Heisenberg-type Hamiltonians (and their $SU(N)$ cousins): Here, symmetries are broken on both sides of the transition. In the LGW framework, this requires a two-order-parameter description, and generically implies a first order transition. However, a continuous transition is possible if it is driven by condensation of topological defects, which carry quantum numbers related to the opposite phase. This scenario is known as "deconfined criticality", and a proper low-energy description rests on fractional degrees of freedom coupled to gauge fields.

Finally, we will analyze a Fermi impurity (polaron problem) which also gives rise to an interesting quantum phase transition related to the formation of a bound state, dubbed "molecule-to-polaron transition". This transition occurs in a zero-dimensional subsystem only, and is potentially first order – both these facts inhibit a straightforward LGW-type analysis.

The QPTs under study are either newly proposed by us, or analyzed rigorously with RG methods for the first time, along with various experimental predictions. Thus, our results contribute to a better understanding of **non-LGW criticality**, which is about to grow into a firm story of its own [SBS$^+$04, Xu12]. Most certainly, many exciting plot-twists of this new story still await us in the future.

Outline

The remainder of this thesis has the following detailed structure:

○ In the short chapter 2 we will flesh out the LGW paradigm with some more details, and give a quick introduction to quantum critical points.

○ The next chapter 3 contains our results on quantum critical metals. We give a systematic introduction to this broad field in Sec. 3.1, discussing Fermi liquids, their instabilities, and the standard RG diagnosis thereof. Furthermore, we introduce the commonly used hot spot description of these instabilites. Next, we review the early, LGW-type approach to the hot spot theories, which turned out to be incorrect in retrospect. By contrast, a perturbatively controlled approach to metallic criticality is provided by the ϵ-expansion. We describe it along with a short juxtaposition between the field-theoretic RG used in this context and the usual Wilsonian RG. Our novel application of the ϵ-expansion for so-called FFLO quantum criticality (related to inhomogeneous superconductivity) is detailed in Sec. 3.2: we start with an introduction to FFLO physics and a list of our main results on the FFLO-normal metal transition, reprinting the article thereafter. Then, in Sec. 3.3 we move to our work on $2k_F$-criticality, introducing the reader to charge density waves and highlighting our results, again followed by the paper. All subsequent "paper-sections" follow this structure as well.

---

1 This does not exhaust all LGW-violating scenarios, for instance we do not discuss transitions involving topological order [Xu12].

○ Our results for two-dimensional quantum magnets are discussed in chapter 4. To begin with, in Sec. 4.1 we introduce the reader to transitions involving two symmetry-broken phases, which arise as consequence of the Lieb-Schultz-Mattis-theorem. We also sketch the LGW view on these transitions. Next, we provide a completely different view, in terms of "deconfined criticality" which involves a gauge theory and fractionalized degrees of freedom. For the RG-description of deconfined criticality, our method of choice is the functional renormalization group, and we give a lightning recap along with its gauge theory version. Finally, in Sec. 4.2, we present our results on a $SU(3)$-generalization of deconfined criticality, which arises naturally for triangular lattices.

○ The next chapter 5 is devoted the so-called "Fermi polaron problem" of a single impurity interacting with a Fermi sea, which is introduced in Sec. 5.1, including a discussion of its experimental realizations. Our main attention in the two publications of this chapter actually lies on the computation of impurity spectra far from any phase transition. The starting point for our analysis is always the limit of infinite mass impurities, and we give a phenomenological overview of the resulting spectra, and the leading finite mass modifications thereof. In the first paper (Sec. 5.2) we then zoom in on two-particle spectra, mainly in the context of optical experiments involving semiconductor-microcavities. The second paper (Sec. 5.3) moves to single-particle spectra, with a focus on ultracold gas experiments. Last, in Sec. 5.4 we discuss our ongoing (and, we caution, preliminary!) work on a potential (non-LGW) transition in the impurity model ("molecule-to-polaron transition"), which is related to bound state formation.

○ Finally, we conclude and present an outlook on selected topics in Sec. 6.

2 The LGW paradigm of critical phenomena

To set the stage, let us flesh out the introductory discussion of the LGW paradigm and QPTs with some technical details. We begin with the Landau mean field theory (see, e.g., [TT87, Xu12]), which is formulated in terms of a (multi-component) order parameter $\boldsymbol{\phi} = (\phi_1, \ldots \phi_N)$, $i = 1, \ldots N$. $\boldsymbol{\phi}$ carries an irreducible representation of the high-temperature symmetry group G. The thermal expectation of $\boldsymbol{\phi}$ should fulfill $\langle \boldsymbol{\phi} \rangle = 0$ for $T > T_c$, and $\langle \boldsymbol{\phi} \rangle \neq 0$ for $T < T_c$. This can be achieved by introducing a Landau free energy of the form

$$F_{\mathrm{L}}(\boldsymbol{\phi}) = r\boldsymbol{\phi}^2 + \sum_{ijkl} u_{ijkl}\phi_i\phi_j\phi_k\phi_l \ , \quad r = r_0(T - T_c) \ , \tag{2.1}$$

with r_0 a positive constant. Here, F_{L} has to be invariant under all symmetry transformations $g \in G$, and must also be bounded from below, which restricts the set of (constant) coefficients u in the fourth order term. We have also assumed that a third-order term is forbidden by symmetry (which e.g. holds in the simplest case of Ising symmetry, $G = Z_2$); if it is not, the Landau theory predicts a first order transition. Now, depending on the precise form of the fourth order term, for $T < T_c$ minimization of F_{L} with respect to ϕ_i leads to a manifold $\mathcal{M}$ (or a discrete set) of degenerate ground state configurations $\boldsymbol{\phi}_\alpha = \langle \boldsymbol{\phi} \rangle$, which can be individually invariant under a subgroup $G' \subsetneq G$ (i.e. $g' \circ \boldsymbol{\phi}_\alpha = \boldsymbol{\phi}_\alpha$ for $g' \in G'$). The relation between G, G', and $\mathcal{M}$ is $G/G' = \mathcal{M}$; for continuous symmetries, this is nothing but the Goldstone theorem counting the number of gapless modes.

Now, to describe thermal fluctuations, we must allow for a local order parameter $\boldsymbol{\phi}(\mathbf{x})$. For brevity, let us focus on a scalar parameter $\phi(\mathbf{x})$ (related to the breaking of an Ising symmetry) in the following. In the usual statistical mechanics formulation, we can then introduce a partition function for the critical degrees of freedom as [vLRVW07]

$$Z = \int \mathcal{D}\left(\phi(\mathbf{x})\right) \exp(-\beta E_{\mathrm{LG}}[\phi]) \ , \tag{2.2}$$

with inverse temperature β (setting $k_B = \hbar = 1$ troughout), and a Landau-Ginzburg energy functional

$$E_{\mathrm{LG}}[\phi] = \int d^d x \ r\phi(\mathbf{x}) + (\nabla_{\mathbf{x}}\phi(\mathbf{x}))^2 + u\phi^4(\mathbf{x}) \ , \tag{2.3}$$

corresponding to the celebrated ϕ^4 model. In this formulation, we can associate the prefactor of the quadratic term $(r + \mathbf{k}^2)$ in Fourier space with the inverse order parameter susceptibility χ^{-1}(modulo a prefactor). $\chi(\mathbf{k} = 0)$ diverges at the critical point where $r \to 0$. r is also proportional to the inverse correlation length, $r \propto \xi^{-2}$, which accordingly scales as $\xi \propto (T - T_c)^{-1/2}$ when the interaction term is neglected. Of course, the divergence of the correlation length at $T = T_c$ is the pinnnacle of criticality.

The general formulation of Eq. (2.3) allows us to assess the validity of the Landau approach [Eq. (2.1)] by comptuting the expectation value of fluctuations. On a qualitative level, the mean field treatment is legitimate as long as [KSF01]

$$\mathcal{G} \equiv \frac{\langle (\phi(\mathbf{x}) - \langle \phi(\mathbf{x}) \rangle)^2 \rangle}{\langle \phi(\mathbf{x}) \rangle^2} \lesssim 1, \quad \langle \phi \rangle = \frac{1}{Z} \int \mathcal{D}(\phi(\mathbf{x})) \phi(\mathbf{x}) \exp(-\beta E_{\mathrm{LG}}[\phi]), \qquad (2.4)$$

which is known as Ginzburg criterion. For the ϕ^4-model, one finds [KSF01] $\mathcal{G} \propto (T - T_c)^{\frac{D-4}{2}}$; this means that, close to the critical temperature, the mean field treatment is valid for $D > 4$, but invalid for smaller dimensions; in reality, however, the proportionality factor can be small, and sometimes the mean field exponents are still observed fairly close to the critical point (this is the case for 3D superconductors).

If the mean field treatment is invalid, one must apply RG methods to treat Eq. (2.3). As realized by Wilson and Fisher [WF72], one way to get the RG under control is to expand around $d = 4$; the special role of four dimensions is already seen in the Ginzburg reasoning above. We will not discuss the (standard) RG further here (for an introduction see, e.g. [Sac11]). Some thoughts on the relation between the field-theoretical and usual Wilsonian (momentum cutoff) RG for the ϕ^4-model are presented in Sec. 3.1.5. Instead, let us directly make the leap to the quantum case, paraphrasing Ref. [vLRVW07].

Quantum critical points

To incorporate quantum fluctuations, we must allow for an imaginary-time dependence of the order parameter fields, and lift the classical partition function of Eq. (2.2) to a quantum one:

$$Z_{\mathrm{q}} = \int \mathcal{D}(\phi(\tau, \mathbf{x})) \exp\left(-S[\phi]\right) , \qquad (2.5)$$

with a quantum action

$$S[\phi] = \int_0^\beta d\tau \int d^d x \ \mathcal{L}_{\mathrm{kin}}[\phi(\tau, \mathbf{x})] + r\phi^2(\tau, \mathbf{x}) + (\nabla_{\mathbf{x}} \phi(\tau, \mathbf{x}))^2 + u\phi^4(\tau, \mathbf{x}) . \qquad (2.6)$$

Here, we view r as a renormalized parameter which parametrizes the distance from the phase transition. The kinetic (or Berry-phase) contribution $\mathcal{L}_{\mathrm{kin}}$ determines the quantum critical behaviour. For phase transitions at at finite temperature T, $\mathcal{L}_{\mathrm{kin}}$ vanishes since ϕ is essentially time-independent. In short, the reason is that the typical energy scale near criticality is $r \to 0$, while the energy scale of modes at a finite Matsubara frequency ω_n is $\omega_n \propto T$.[1] Thus, all modes $\phi(\omega_n, \mathbf{k})$ are gapped and can be integrated out (yielding an unimportant prefactor in the partition function), except for the one with vanishing Matsubara frequency $\phi(0, \mathbf{k})$ corresponding to classical fluctuations $\phi(\mathbf{x})$. Thus, finite temperature phase transitions are essentially classical.

By contrast, for a true quantum phase transition at $T = 0$ the dynamical fluctuations contained in $\mathcal{L}_{\mathrm{kin}}$ are crucial. In the simplest (relativistic-like) case, $\mathcal{L}_{\mathrm{kin}} \propto \omega_n^2 \phi^2(\omega_n, \mathbf{k})$, which implies that the $T \to 0$ limit of the quantum action is simply a $(d+1)$-dimensional version of the Landau-Ginzburg energy functional (2.3). But $\mathcal{L}_{\mathrm{kin}}$ can also involve different powers of ω_n, implying an "anisotropy" between time and space directions. Including $\mathcal{L}_{\mathrm{kin}}$ and the effect of interactions (in some approximate fashion) then gives rise to a generalized order parameter susceptibility χ, which is proportional to a scaling factor

$$\chi(\omega_n, \mathbf{k}) \propto \tilde{\chi}(\mathbf{k}^2 \xi^2, \omega_n^2 \xi_\tau^2), \quad \xi \propto |r|^{-\nu}, \ \xi_\tau \propto |r|^{-z\nu} . \qquad (2.7)$$

---

1 We will move freely between real and Fourier space in the following.

Here, the correlation length ξ is the characteristic scale for spatial fluctuations, while ξ_τ can be seen as a correlation "length" in the imaginary time direction. Both these scales diverge as the critical point $r \to 0$ is approached; the divergence of ξ is given by the correlation length exponent ($\nu = 1/2$ within the mean field treatment discussed before), while the divergence of ξ_τ involves the additional "dynamical critical exponent" z – for the relativistic-like case, $z = 1$.

A prototypical phase diagram which contains a quantum-critical point (and where r now denotes the distance from that point) is shown in Fig. 2.1.
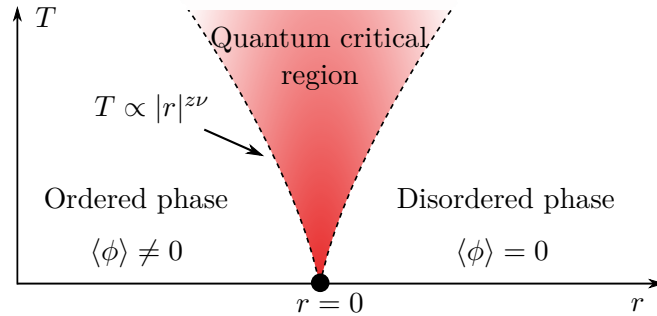

Figure 2.1 Schematic phase diagram near a metallic quantum critical point, see main text.

The quantum critical point may or may not be the endpoint of a line of finite-temperature transitions.[2] The region of this phase diagram of most interest to us is the "scaling fan" at finite temperature which emanates from the quantum critical point, surrounded by the ordered phase with $r < 0, \langle \phi \rangle \neq 0$, and the disordered symmetric phase with $r > 0, \langle \phi \rangle = 0$. The scaling fan is delimited by the crossover lines $T \simeq |r|^{z\nu}$. For $T > |r|^{z\nu}$, the temporal correlation length $\xi_\tau \propto |r|^{-z\nu}$ is larger than the typical inverse energies $1/\omega_n \propto 1/T$, and the system is essentially critical. This critical behaviour also imprints itself on typical obervables (as e.g. discussed in the next section), until our low-energy description eventually looses its validity at temperatures that are too high.

As advertised in the introduction, in the forthcoming sections we will discuss situations where the above order-parameter description is incomplete, starting with quantum critical metals. Still, we will mostly assume that the schematic phase diagram of Fig. 2.1 is valid, although the critical exponents will change once we deviate from the pure order-parameter theory.

---

2 In a strict 2D situation, this is forbidden by the Mermin-Wagner theorem for order parameters related to continuous symmetries, unless the phase-transition is of Kosterlitz-Thouless type.

3 Quantum critical points in two-dimensional metals

3.1 Introduction

3.1.1 Phenomenology of metallic criticality

Landau's Fermi liquid theory (see, e.g., [Mah00, vLRVW07]) forms the basis for a phenomenological description of any "good" metal. At its heart lies the idea that particle-hole excitations in a system of interacting fermions can be seen as quasiparticles which inherit the quantum numbers of the non-interacting fermions. At low energies $\omega \propto |k - k_F|$ (with k_F the Fermi momentum) and small temperatures T, the quasiparticles become long lived, with diverging lifetime $\propto \omega^{-2}$ (self-energy ω^2). Interactions between the quasi-particles can be parametrized by a few dimensionless "Landau-parameters", which imprint themselves in the quasiparticle effective mass and density of states, and appear in the prefactors of thermodynamic observables. However, within Fermi liquid theory, the interactions do not compromise the one-to-one correspondence with the free fermions. Typical signatures of a Fermi liquid are the resistivity (from electron-electron scattering in a system without translational invariance) $\rho \propto T^2$, the specific heat $C_V(T) \propto T$, or the constant spin susceptibility χ.

The correspondence with the free Fermi gas, and hence the validity of Fermi liquid theory, breaks down if electronic instabilities occur. As a result, the system spontaneously develops a new kind of order, with order parameter $\langle \phi \rangle \neq 0$. Within Fermi liquid theory, this is signaled by large negative values of the Landau parameters, which lead to divergent susceptibilities $\propto \langle \phi\phi \rangle$. Paradigmatic examples of ordered phases relevant to us here are those associated with a breaking of rotational symmetries (nematic order arising from "Pomeranchuk" instabilities [DM06]) or translational invariance (density-waves), which can both be accompanied with magnetic order. Further, the ordered phase may or may not have a residual metallic character.

In this work, we will mostly be interested in the quantum critical point associated with the instability. According to our discussion in Sec. 2, this critical point gives rise to a "scaling fan" at finite temperatures. In this fan, physics is dominated by critical correlations, which can be in marked contrast to Fermi liquid expectations. In particular, interactions with the critical order parameter fluctuations can fully decohere the Landau quasiparticles, such that the quasi-particle weight Z vanishes.[1] As a result, the physical observables show non-Fermi liquid, or "strange metal" behaviour. Let us list some of its hallmark manifestations, with the obvious caveat that often deviating origin stories for these phenomena can be found.

○ **linear-in-T resistivity**: The typical low-temperature resistivity of a metal scales as $\rho = \rho_0 + AT^\alpha$ with temperature T, where ρ_0 is a residual impurity contribution. For a Fermi liquid, $\alpha = 2$ if phononic contributions can be disregarded (they lead to $\alpha = 1$). Genuine Non-Fermi liquid behaviour with exponent $\alpha = 1$ for quasi two-dimensional systems is e.g. frequently found in d- and f- electron metals [Ste01], cuprates at optimal doping [Hus08], as well as iron-pnictide superconductors [INH09].

○ **no resistivity saturation**: In a good metal, the resistivity eventually saturates at high temperatures. As a rule of thumb, the upper bound is reached when the mean free path

---

[1] Note that the Fermi surface may still be well-defined, for instance via a kink singularity in the momentum distribution function [Sen08].

ℓ (with $\rho \propto 1/\ell$) is of the order of the interatomic distance ("Mott-Ioffe-Regel limit"). The cuprates are a prominent exception to this rule, with the resistivity overshooting the Mott-Ioffe-Regel limit already at moderate temperatures [GCH03].

○ **non-linear specific heat**: Another important non-Fermi liquid signature is a non-linear T-dependence of the specific heat. Often, a logarithmic enhancement $C_V \propto T \log(1/T)$ is observed, for example again in the d- and f- electron metals [Ste01], or very recently in the cuprates [MGB$^+$19] upon magnetic field-induced suppression of superconductivity. Furthermore, $C_V(T)/T$ also shows an upswing in organic superconductors [LWD$^+$07] which may be indicative of an FFLO quantum critical point (see [PZS16] and Section 3.2).

○ **high-T_c-superconductivity**: In principle, the non-Fermi liquid behaviour should be observable down to the lowest temperatures. While this is the case for certain heavy-fermion metals [vLRVW07], and possibly the FFLO case, in the high-T_c materials the putative quantum critical point is hidden beneath a superconducting "dome". In fact, the presence of the critical point constitutes a possible explanation for high-T_c: broadly speaking, the associated critical order parameter fluctuations act as a "pairing glue", which induces Cooper pair formation, and yields a critical temperature T_c well above the ordinary phonon-mediated one (see, e.g. [LSBK15]). On the other hand, the decoherence of the fermions at the critical point counteracts pairing. The competition of these two effects is a very active, and controversial, field of research [MMSS15, WAA$^+$16].

3.1.2 RG diagnosis of Fermi liquid instabilities

After the phenomenological overview of the last section, let us have a deeper microscopical look into non-Fermi liquid criticality, by first understanding how instabilities in a Fermi liquid arise. A particularly clear diagnosis is obtained with RG methods. To illustrate how this works, let us sketch the standard ($T = 0$) RG treatment of 2D Fermions with a spherical Fermi surface, following [Sha94]. We start from an action

$$\mathcal{S} =$$
$$\sum_{k,\sigma} \bar{\psi}_\sigma(k)(-i\omega + \mathbf{k}^2/2m - \mu)\psi_\sigma(k) + \sum_{k_1,k_2 k_1',k_2';\sigma_1,\sigma_2} u(k_1, k_2; k_1', k_2')\bar{\psi}_{\sigma_1}(k_1')\bar{\psi}_{\sigma_2}(k_2')\psi_{\sigma_2}(k_2)\psi_{\sigma_1}(k_1),$$
$$k = (\omega, \mathbf{k}), \quad \sum_k = \sum_{\omega, \mathbf{k}}, \tag{3.1}$$

where $u \propto \delta^{(2)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2') \cdot \delta(\omega_1 + \omega_2 - \omega_1' - \omega_2')$ is a generic short-ranged interaction, without further frequency dependence beyond energy conservation. $\mu = k_F^2/2m$ is the Fermi energy. To parametrize momentum, we use spherical coordinates, writing $\mathbf{k} = (k_F + q) \cdot \mathbf{e}_\theta$. In a low-energy limit, we may then expand $\mathbf{k}^2/2m - \mu \simeq v_F q$, with Fermi velocity $v_F = k_F/m$; this linearization is approximately valid up to a cutoff Λ, see Fig. 3.1. Thus, our momentum scales read $|q| < \Lambda \ll k_F$.

To run the RG program (see, e.g., [Sac11]), the standard way is to integrate out momenta $|q| \in [\Lambda/b, \Lambda]$, with a scale factor $b \gtrsim 1$, and subsequently rescale $\Lambda/b \to \Lambda$. Frequencies are left untouched. This decimation procedure can be done loop order by loop order.[2] At tree level (no loops), it just corresponds to a scaling analysis, and leads to the following conclusions for the interaction u: First, the dependence of u on radial momenta k is irrelevant, as found by Taylor expansion. Thus, one just has to analyze $u(\theta_1, \theta_2; \theta_1', \theta_2')$ – we'll suppress spins which don't do much here. From this remaining coupling function, only a one-dimensional

---

2 We implicitly assume that u is small to render the loop expansions controlled; this will be discussed at length below for a related problem.

set of coupling survives (is marginal), the rest abruptly renormalizes to zero. The surviving couplings are, first, the couplings related to forward scattering, for which $\theta_1' = \theta_1, \theta_2' = \theta_2$ (or permuted). Second, the pairing interaction is of importance, with $\theta_2 = -\theta_1, \theta_2' = -\theta_1'$.
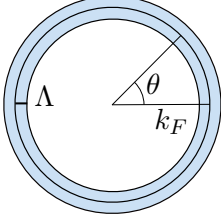

Figure 3.1 Spherical Fermi-surface with radial cutoff Λ

To infer the ultimate fate of the couplings, one must go to one-loop order. In effect, one finds that the pairing interaction is relevant (i.e., diverges during the RG flow), if it starts out attractive (or at least one angular momentum component of it). By contrast, forward scattering is exactly marginal, and does not renormalize at all. Thus, pairing is the only potential weak-coupling instability of fermions with a spherical Fermi surface.

For a realistic material with a more complicated Fermi-surface geometry, additional instabilities can occur. For example, nested Fermi surfaces, for which the fermion dispersions fulfill $\epsilon(\mathbf{k}) = \epsilon(\mathbf{k} + \mathbf{Q})$ with a nesting wave-vector $\mathbf{Q}$, are generically prone to density wave instabilities [Sha94] (in the charge or spin channel, see below).

3.1.3 Bose-Fermi hot spot theories

In principle, a purely fermionic description is all we need to correctly identify the critical behavior of the system. However, since an action of the form (3.1) is hard to treat in all but the simplest cases, in practice often two deformations are applied. Let us discuss them in term.

Order parameter theory

One way to simplify the RG analysis is to focus on a single instability channel, and to perform a Hubbard-Stratonovich decoupling in it. In a very schematic fashion (see, e.g., [Rei06] for a detailed account), this works as follows: We start from the fermionic partition function in a functional integral formulation $Z = \int \mathcal{D}(\psi, \bar{\psi}) \exp(-\mathcal{S})$, and multiply it by a unity of the form $\mathbb{1} = \int \mathcal{D}(\phi, \phi) \exp(-\bar{\phi}(-u)\phi)$. Here, we have assumed that u is attractive (which generates an relevant RG flow according to our discussion above), and suppressed indices. Then, we shift ϕ by a fermion bilinear $X \sim \psi\psi$ or $\bar{\psi}\psi$, related to the decoupling channel in question. This eliminates the four-fermion interaction term, trading it for a Yukawa-type coupling of the form $u\bar{X}\phi + u\bar{\phi}X$. ϕ can also be seen as **order parameter field** indicating a non-trivial ordering of the fermions, since a non-zero expectation value $\langle \phi \rangle \neq 0$ implies $\langle X \rangle \neq 0$ on mean-field level.

For concreteness, let's consider a decoupling in the direct density channel without spins and focus on an interaction that depends on the transferred momentum $\mathbf{q}$ only, $u = u(\mathbf{q})$. The decoupling then produces a term of the form

$$\mathcal{S}_{\text{int}} = \sum_{\substack{\omega,\nu \\ \mathbf{k},\mathbf{q}}} u(\mathbf{q})\bar{\psi}(\omega + \nu, \mathbf{k} + \mathbf{q})\psi(\omega, \mathbf{k})\phi(\nu, \mathbf{q}) \,, \tag{3.2}$$

with a real field ϕ. Furthermore, a quadratic term for ϕ will be generated, of the form $\frac{1}{2} \sum_{\nu,\mathbf{q}} \phi(\nu, \mathbf{q})u(\mathbf{q})\phi(-\nu, -\mathbf{q})$. We can also rescale the boson $\phi \to \phi/u$, which will lead to a quadratic term $\propto \phi^2/u$, and remove u from the interaction term. In any case, the coefficients of the interaction term and the quadratic term for the bosons are linked. However, once we start the renormalization process, i.e. integrate out high-energy degrees of freedom, the terms will renormalize independently. For the boson quadratic term, we may assume that this renormalization will initially produce analytic terms (as is always the case in a Wilsonian

RG scheme), and the boson quadratic term should read

$$\mathcal{S}_\phi = \tfrac{1}{2} \sum_{\nu, \mathbf{q}} \phi(\nu, \mathbf{q}) \chi^{-1}(\nu, \mathbf{q}) \phi(-\nu, -\mathbf{q}), \tag{3.3}$$

where the order parameter susceptibility χ is of *Ornstein-Zernicke* form [ACS03]

$$\chi^{-1}(\nu, \mathbf{q}) = \nu^2 + \mathbf{p}^2 + r \ , \quad \mathbf{q} = \mathbf{Q} + \mathbf{p} \ . \tag{3.4}$$

In the latter expression, we have assumed that the susceptibility is peaked around a momentum $\mathbf{Q}$; a nonzero value of $\mathbf{Q}$ corresponds to the breaking of translational symmetries. The mass/gap $r \propto \xi^{-2}$ is proportional to the squared inverse correlation length, and $r \to 0$ indicates an instability.[3] It is clear that we can associate the boson mass r with our tuning parameter from Sec. 2.

Finally, it is common to relabel the Yukawa coupling as $\lambda(\mathbf{q})$, and assume that it can be expanded as $\lambda(\mathbf{q}) \simeq \lambda(\mathbf{Q}) + \mathcal{O}(\mathbf{q} - \mathbf{Q}) \equiv \lambda$, dropping the momentum dependent terms because of RG irrelevance.

For our derivation of the bosonic action, we have relied on a microscopic "parent" theory at high energies which was purely fermionic. However, this is not really necessary, since after all we do not know the correct high-energy fermionic theory for a realistic problem either. So, if one is interested in universal properties of a phase transition, one can directly write down a Bose-Fermi theory from symmetry arguments, similar to the original Landau approach (Sec. 2), but retaining the fermions. This is exemplified for an Ising-type order parameter ϕ in Ref. [MS10b].

Restriction to a hot spot

Let's return to the fermions. So far, we assumed them to live near the whole Fermi surface. For the textbook problem of the superconducting instability of fermions on a spherical Fermi surface sketched in Sec. 3.1.2, this is indeed unavoidable, since fermions with arbitrary momenta $(\mathbf{k}, -\mathbf{k})$ can form a Cooper pair. For other instabilities, and more realistic Fermi surfaces, it is often permissible to restrict the fermions to a set of **hot spots**. That is, we expand the Fermion dispersion around certain fixed momenta $\mathbf{q}_\alpha$ on the Fermi surface. In the two papers discussed in this chapter, these points are determined from a "pseudo-nesting" condition, as illustrated in Fig. 3.2:

Figure 3.2 Hot spot description of a density wave instability.
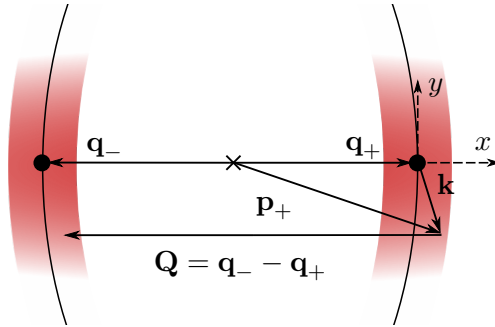
E.g. the dispersion of fermions with momenta $\mathbf{p}_+ = (\mathbf{q}_+ + \mathbf{k})$, $|\mathbf{k}| \ll |\mathbf{q}_+|$, can be approximated by $\epsilon_{\mathbf{p}_+} \simeq k_x + k_y^2$, with k_x and k_y measured in the dashed coordinate system indicated in

---

3 which matches with the fermionic description of Sec. 3.1.2 if we imagine that $r \propto 1/u$.

the figure. The prefactors of the dispersion are removed by rescaling, and the sign of the linear term is determined by the Fermi surface orientation. If this fermion absorbs a "$2k_F$-wavevector" $\mathbf{Q} = \mathbf{q}_- - \mathbf{q}_+$ in a scattering process, it ends up close to the hot spot $\mathbf{q}_-$, with an energy $-k_x + k_y^2$. In particular, a fermion with a momentum tangent to the Fermi surface at $\mathbf{q}_+$ still has a small energy $\sim k_y^2$ after scattering, since the tangents at $\mathbf{q}_-, \mathbf{q}_+$ are parallel; this constitutes the "pseudo-nesting".[4] Due to the large scattering phase space, $\mathbf{Q}$ is a preferential density wave-vector (as well as $-\mathbf{Q}$ for inversion symmetric problems), and the fermions in the "hot regions" close to the hot spots $\mathbf{q}_-, \mathbf{q}_+$ are prone to density wave formation, while the rest of the Fermi surface remains cold, i.e. Fermi liquid like.

It should be noted that for the fine-tuned case of a perfectly circular Fermi surface the density wave susceptibility will actually be peaked for all momenta with $|\mathbf{Q}| = 2k_F$. However, for a generic Fermi surface, this degeneracy will be lifted and the hot spot description is valid, as e.g. exemplified for a lattice model exhibiting a CDW instability with d-wave symmetry in Ref. [HM12].

Let us formulate the hot spot theory by expanding around the relevant momenta. We denote $\phi(\nu, \pm\mathbf{Q} + \mathbf{p}) \equiv \phi^\pm(\nu, \mathbf{p}) \equiv \phi^\pm(p)$ and $\psi(\omega, \mathbf{q}_\pm + \mathbf{k}) \equiv \psi_\pm(k)$. In these variables, a prototypical low-energy action that describes a density wave transition reads

$$
\mathcal{S}_{\mathrm{DW}} = \sum_{s=\pm;\sigma} \int_k \bar{\psi}_{s,\sigma}(k) \left(-i\omega + sk_x + k_y^2\right) \psi_{s,\sigma}(k) + \int_k \phi^+(p) \left(\nu^2 + p_x^2 + p_y^2 + r\right) \phi^-(-p)
$$

$$
+ \lambda \sum_\sigma \int_{k,p} \left[\phi^+(p)\bar{\psi}_{+,\sigma}(k+p)\psi_{-,\sigma}(k) + \phi^-(-p)\bar{\psi}_{-,\sigma}(k-p)\psi_{+,\sigma}(k)\right] ,
$$

$$
\int_k = \int \frac{d\omega}{(2\pi)^3} dk_x dk_y . \tag{3.5}
$$

Here, we have taken a continuum limit to transform sums into integrals, which have an implicit UV cutoff Λ similar to Sec. 3.1.2. In the expansion of the Bose fields, we have retained analytical terms only, as in Eq. (3.4); as we will see below, non-analytical terms will eventually be generated due to the coupling to the fermions. Furthermore, we have reintroduced spins, and assumed a spin-independent order parameter ϕ and a Yukawa coupling that is diagonal is spin space. Thus, ϕ describes a charge density wave (see Sec. 3.3), i.e. a periodic spatial variation of the fermion density without spin structure.

At this point, let us give an (incomplete) overview over related hot spot theories which frequently appear in the literature. First, in the above we have implicitly assumed that $\mathbf{Q}$ is incommensurate, i.e. it is not a rational multiple of a reciprocal lattice vector. In the **commensurate** case, in particular if $\mathbf{Q}$ is a reciprocal lattice vector by itself, one must identify $\phi \equiv \phi^+ = \phi^-$. This drastically changes the renormalization group analysis, as already discussed in one of the first works on the subject by Altshuler et al. [AIM95].

In addition to charge density waves, one can also decouple in the **spin density wave** (SDW) channel, corresponding to antiferromagnetic order. This requires inclusion of a vector order parameter $\vec{\phi} = (\phi_x, \phi_y, \phi_z)$ which schematically couples to the Fermions as $\vec{\phi}\bar{\psi}_{+,\sigma}\vec{\tau}_{\sigma\sigma'}\psi_{-,\sigma'}$, where $\vec{\tau}$ is a vector of Pauli matrices. Typically one considers antiferromagnetism on a square lattice (Néel order), associated with a commensurate wave-vector of $\mathbf{Q} = (\pi, \pi)$. The SDW critical point is e.g. relevant for the description of the hole-doped cuprates [ACS03], where the wave-vector couples four pairs of hot spots on an open Fermi surface centered around (π, π) as shown in Fig. 3.3(a). In this case, the tangents at a pair of hot spots are not parallel.

---

4 "perfect nesting" would mean that the Fermi surface at $\mathbf{q}_-, \mathbf{q}_+$ is completely flat.

(a)

(b)

(c)

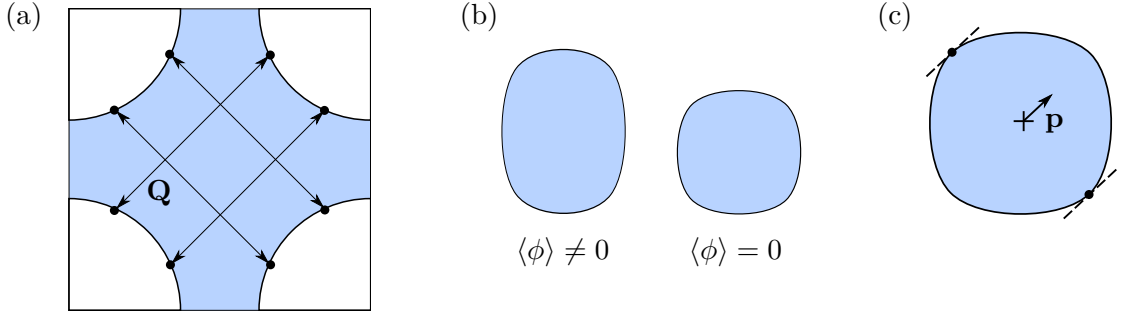$\langle \phi \rangle \neq 0$ $\langle \phi \rangle = 0$

Figure 3.3 Fermi surface geometries for paradigmatic hot spot theories. (a) Typical Fermi surface and density wave vector at the SDW critical point. (b) Change of the Fermi surface across an Ising nematic transition. (c) Hot spots in the Ising nematic case. The Figs. are adapted from [Sac11].

Another class of instabilities at finite wave-vector is related to superconductivity of 2D electrons in an in-plane magnetic field. Here, spin polarization of the Fermi surface leads to formation of Cooper pairs with non-zero net momentum $\mathbf{Q}$, with associated order parameter (superconducting gap) $\Delta(\mathbf{Q})$. A detailed analysis of this case (the first proper RG treatment to date) is performed in this thesis, see Sec. 3.2.

Let us turn to instabilities with $\mathbf{Q} = 0$. The onset of ferromagnetic ordering [CPR04] is an example in the spin channel. A ($\mathbf{Q} = 0$)-instability in the charge channel corresponds to deformations of the Fermi surface, the paradigmatic example being the **Ising nematic** phase transition. As sketched in Fig. 3.3(b), it corresponds to a reduction of the Fermi surface symmetry from fourfold to twofold rotational invariance, and can be captured by a real Ising order parameter ϕ coupled to fermions. In a purely fermionic language, the Ising nematic transition is the result of a "Pomeranchuk instability" in a higher angular momentum component of forward scattering [DM06].

An order parameter boson peaked at $\mathbf{Q} = 0$ can couple to the full Fermi surface. Nevertheless, a hot spot description is legitimate: One can show that bosons with small fixed momenta $\mathbf{p}$ can only efficiently scatter fermions around points on the Fermi surface to which $\mathbf{p}$ is tangent [MS10a], see Fig. 3.3(c). Thus, although the full Fermi surface is "hot", fermions at points with different tangents effectively do not mix, and it is enough to restrict a low-energy description of the transition to an (arbitrary chosen) pair of hot spots. An appropriate action reads:

$$\mathcal{S}_{\text{Ising}} = \sum_{s=\pm;\sigma} \int_k \bar{\psi}_{s,\sigma}(k) \left(-i\omega + sk_x + k_y^2 \right) \psi_{s,\sigma}(k) + \frac{1}{2} \int_k \phi(p) \left(\nu^2 + p_x^2 + p_y^2 + r \right) \phi(-p)$$

$$+ \lambda \sum_{s=\pm;\sigma} \int_{k,p} \phi(p) \bar{\psi}_{s,\sigma}(k+p) \psi_{s,\sigma}(k) \, , \tag{3.6}$$

which looks rather similar to the density wave case (but the small formal differences result in sizeable physical changes).

3.1.4 RG description of hot spot criticality - early approaches

Let us take a first swing at an RG analysis of hot spot theories by performing a power counting analysis. If we require that the fermionic kinetic part of Eqs. (3.5), (3.6) is scale

invariant, this leads to the following (Gaussian) scaling:

$$\omega' = \omega \cdot b, \quad k'_x = k_x \cdot b, \quad k'_y = k_y \cdot \sqrt{b}, \quad \psi'(k') = \psi(k) \cdot b^{-7/4}, \quad \phi'(k') = \phi(k) \cdot b^{-7/4},$$
$$r' = r \cdot b^2, \quad \lambda' = \lambda \cdot b^{1/4} \ . \tag{3.7}$$

As per usual, the mass term is strongly relevant. More importantly, the Yukawa coupling λ is relevant as well, because λ increases under successive rescaling transformations. This means that the theory will flow to strong coupling in the low-energy limit, and any perturbative expansion in λ is bound to fail – a simple fact that makes the analysis of metallic criticality both challenging and interesting.

Hertz-Millis approach

An influential historical attempt to get the RG behaviour under control is due to Hertz [Her76] (later refined by Millis [Mil93]). Hertz took the Hubbard-Stratonovich idea one step further, and integrated out the fermions completely (see also [Sac11]). The result is a purely bosonic, **Landau-Ginzburg**-type theory similar to the one we discussed in Sec. 2, but "microscopically derived" (see also Appendix A of 3.2.2). For the Ising nematic, it is of the form [MS10a]:

$$\mathcal{S}_{\text{Hertz}} = \sum_{n \geq 2,\text{even}}^{\infty} \int_{p_1} \ldots \int_{p_n} \Gamma_n(p_1, \ldots p_n) \phi(p_1) \ldots \phi(p_n) \ . \tag{3.8}$$

Terms with an odd number of fields can be neglected on symmetry grounds. The even vertices Γ_n are determined by the Feynman diagrams shown in Fig. 3.4
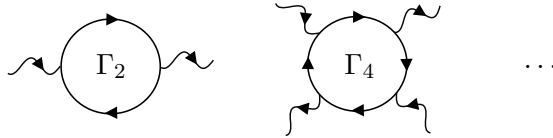

Figure 3.4 Feynman diagrams that determine the coefficents Γ_n. Straight lines denote fermions, (amputated) wavy lines bosons.

Evaluation of Γ_2 using the hot spot fermion propagators of Eq. (3.6) leads to $\Gamma_2(p_1, p_2) = \delta_{p_1, -p_2} \cdot \gamma |\nu| / |p_y|$, with a numerical coefficient γ. This term, commonly denoted **Landau damping** by analogy to Landau's theory of damping of charged waves in a plasma, is of crucial importance. It reflects the fact that the bosonic excitations are completely overdamped due to decay into particle-hole pairs of the Fermi sea. In the diagram for Γ_2, one must integrate over all internal fermion momenta, and not just over large ones as in the RG. As a result, Landau damping predominantly comes from low-energy fermions, and is non-analytic, whereas the (Wilsonian) RG exclusively produces analytic terms (as well as the phenomenological Landau-Ginzburg ansatz). While the hot spot electrons yield the Landau damping only, integration over electrons from the full Fermi surface (which we neglected) will yield analytic terms similar to the Ornstein-Zernicke result (3.4). In total, an appropriate quadratic term reads

$$\mathcal{S}_{\text{Hertz},2} = \int_p \phi(p) \left(p_x^2 + p_y^2 + \gamma \frac{|\nu|}{|p_y|} + r \right) \phi(-p) \ , \tag{3.9}$$

where we have dropped the terms ν^2 which are strongly subleading compared to the Landau damping term in the low-energy limit. Let us perform a scaling analysis of (3.9), similar

to Eq. (3.7). Since the reference to the Fermi surface is lost, we can scale p_x, p_y equally. Ensuring the scale-invariance of the Landau damping term, the total scaling reads

$$\nu' = \nu \cdot b^3, \quad p_x' = p_x \cdot b, \quad p_y' = p_y \cdot b, \quad \phi'(p') = \phi(p) \cdot b^{-7/2}, \quad r' = r \cdot b^2 . \qquad (3.10)$$

The modified scaling of frequencies implies a dynamical critical exponent $z = 3$, which is crucial for the subsequent analysis.

What about the quartic terms? Somewhat ad hoc, Hertz assumed that Γ_4 possesses an analytic expansion, and the quartic term can be written to leading order as

$$\mathcal{S}_{\text{Hertz},4} = u \int_{p_1, p_2, p_3, p_4} \delta \left(\sum_{i=1}^{4} p_i \right) \phi(p_1)\phi(p_2)\phi(p_3)\phi(p_4) , \qquad (3.11)$$

with a constant factor u. Under the scaling (3.10), u transforms as $u' = u/b$, i.e. it is strongly irrelevant. This comes about due to the effective increased dimensionality $D = d + z$ for $z = 3$. As a result, boson interactions can naively be disregarded, and the scaling (3.10) is exact.

Given our hot spot scaling in (3.7), this seems weird, since we clearly found that (Bose-Fermi) interactions are important. Furthermore, a non-analytic term was already obtained for Γ_2. And indeed, in later works [MS10a, ACS03] it was shown that the basic assumption of Hertz is incorrect, and a careful evaluation of the term Γ_4 in Fig. 3.4 does yield singular non-analytic terms with a complicated momentum and frequency dependence. These terms can be relevant. Similar conclusions hold for theories with a $\mathbf{Q} \neq 0$ order parameter [ACS03]. Therefore, integrating out the gapless fermions does not allow for an economic treatment. In relation to the LGW paradigm, we can thus say the following:

> A basic assumption of the LGW paradigm is that the Landau functional possesses an analytic expansion in powers and gradients of the order parameter. However, in problems with a Fermi surface where the order parameter couples to gapless fermionic excitations, non-analytic contributions can arise! To correctly find the critical behaviour, one must therefore treat bosons and fermions on the same footing.

Large N expansion

Let us therefore return to the Bose-Fermi hot spot theories (3.5), (3.6). A priori, these theories do not possess any small parameter. The basic idea of a **perturbative RG** is to introduce a small parameter α artificially by deforming the action in a non-invasive fashion, and then extrapolate to the physical limit $\alpha \to \mathcal{O}(1)$.

One very simple way to do so is a $1/N$ expansion: One extends the number of fermionic flavors (spins), or equivalently the number of hot spots [ACS03] to N, at the same time rescaling $\lambda \to \lambda/\sqrt{N}$. The standard RPA diagram for the Bose self-energy (equivalent to the the diagram for Γ_2 in Fig. 3.4) incures a factor of N due to the fermion loop and is $\mathcal{O}(1)$. Other diagrams are suppressed by factors of $1/N$, enabling a systematic expansion in $1/N$.

But yet again, hot spot theories cannot be treated with such a simple expansion [Lee09, MS10b]: Essentially due to the anisotropic scaling (3.7), diagrams that contain bare boson lines are IR divergent. As a result, one must reorder the perturbation series, and work with RPA dressed bosons instead (see also Sec. 3.1.5). This reorganization spoils the naive $1/N$ power counting; there are important diagrams at every loop order (so-called "planar" diagrams), the evaluation of which is very difficult.

3.1.5 RG description of metallic criticality - ϵ-expansion

Dimensional extension

Another useful tool which yields perturbatively controlled and even numerically accurate results e.g. for the ϕ^4 theory is ϵ-expansion. In the ϕ^4 case, one allows the space dimension to take continuous values, and expands around $d = 3 + 1$ where the four-boson coupling becomes marginal by power counting, as already evidenced by the Ginzburg-criterion of Eq. (2.4). When adapting this idea to the hot spot case, one encounters the following obstacle: now there are momentum directions tangential and perpendicular to the Fermi surface, thus the dimensional extension is not unique. How to increase the dimensions precisely?

Probably the simplest choice would be to increase the dimensionality of space and of the Fermi surface alike ("fixed co-dimension"), by adding terms of the form $\bar{\psi}(k)k_z^2\psi(k)$ to the action. However, this has a downside: as indicated in Fig. 3.5, on a Fermi surface with dimensionality higher than 1, any two points share a common tangent vector, in contrast to the standard 2D theory where this is true for the hot spots only (see Fig. 3.2). This leads to the explicit appearance of non-universal UV scales $\sim k_F$ in the theory, also known as UV-IR mixing, and spoils locality in momentum space [ML15].
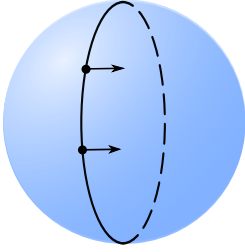

Figure 3.5 Higher-dimensional Fermi surface. Fig. adapted from [Lee18].

A particularly elegant prescription to circumvent this problem was given by Dalidovich and Lee (DL) [DL13], building on previous work by Senthil and Shankar [SS09]. DL argued that one should perform the extension in such a way that the Fermi surface (a line in two dimensions) stays one-dimensional. In other words, one adds dimensions perpendicular to the Fermi surface, increasing its co-dimension.

For the Ising nematic case this is schematically done as follows: one introduces two-component spinors of Dirac type,

$$\Psi_\sigma(k) = \begin{pmatrix} \psi_{+,\sigma}(k) \\ \bar{\psi}_{-,\sigma}(-k) \end{pmatrix} , \quad \bar{\Psi}_\sigma(k) = \Psi_\sigma^\dagger(k)\sigma_y , \tag{3.12}$$

and rewrites the kinetic term for the electrons as

$$S_\psi = \sum_\sigma \int_k \bar{\Psi}_\sigma(k)\left(-ik_0\sigma_y + i\sigma_x\delta_k\right)\Psi_\sigma(k), \quad \delta_k = k_x + k_y^2 , \tag{3.13}$$

where we have relabeled $\omega \to k_0$. Gapless fermions, i.e. those on the Fermi surface, are determined by $\delta_k = 0$. Now, it is well known (e.g. from the Haldane model of Graphene) that a Hamiltonian of the from $H = a \cdot \sigma_x + b \cdot \sigma_y + c \cdot \sigma_z$ with $a, b, c \in \mathbb{R}$ has a spectrum $E = \pm\sqrt{a^2 + b^2 + c^2}$. Accounting for the factors of i in (3.13), we can thus add a term

$$\Psi_\sigma(k)(-ik_z\sigma_z)\Psi_\sigma(k) = -k_z \sum_\sigma \bar{\psi}_{+,\sigma}(k)\bar{\psi}_{-,\sigma}(-k) + \text{h.c.} , \tag{3.14}$$

leading to a dispersion $E_k = \pm\sqrt{k_z^2 + \delta_k^2}$. As a result, the spinor Fermi surface is given by $k_z = 0, \delta_k = 0$, i.e. it is still a one-dimensional line embedded in a three-dimensional space.

Similar to the Haldane model, the price we have to pay for introducing these terms is a reduction of symmetry: as seen in (3.14), the added term is of triplet pairing form, and is neither invariant under $U(1)$ phase rotations related to charge conservation, nor under

$SU(2)$ spin rotations. We will operate under the assumption that this symmetry breaking is harmless (see also below), and the modified model correctly reproduces critical exponents.

To allow for general continuous dimensions, we again relabel $k_x \to k_{d-1}, k_y \to k_d$, and introduce a vector $\boldsymbol{\Gamma} = (\gamma_0, \gamma_1, \ldots \gamma_{d-2})$ of "Gamma matrices" fulfilling the Clifford algebra $\{\gamma_\alpha, \gamma_\beta\} = 2\delta_{\alpha\beta}$. More pedestrian, in our cases of interest we can simply choose $\boldsymbol{\Gamma} = (\sigma_y, \sigma_z \ldots \sigma_z)$. This allows us to completely generalize the Ising nematic action of Eq. (3.6) to d dimensions:

$$\mathcal{S}_{\text{Ising},d} = \sum_\sigma \int_{k^{d+1}} \bar{\Psi}_\sigma(k) \left[-i\boldsymbol{\Gamma} \cdot \boldsymbol{K} + i\sigma_x \delta_k \right] \Psi_\sigma(k) + \frac{1}{2} \int_{k^{d+1}} \phi(k) \left[|\boldsymbol{K}|^2 + k_{d-1}^2 + k_d^2 \right] \phi(-k)$$

$$+ \lambda \sum_\sigma \int_{k^{d+1}, q^{d+1}} \phi(q) \bar{\Psi}_\sigma(k+q) i\sigma_x \Psi_\sigma(k) , \quad \boldsymbol{K} = (k_0, k_1, \ldots k_{d-2}) , \quad \int_{k^{d+1}} = \int \frac{d^{d+1}k}{(2\pi)^{d+1}} .$$

$$(3.15)$$

Here, we have tacitly erased the boson mass r, i.e. we tune to the critical point. We can now repeat our hot spot scaling analysis of (3.7), and find

$$\boldsymbol{K}' = \boldsymbol{K} \cdot b , \quad k'_{d-1} = k_{d-1} \cdot b , \quad k'_d = k_d \cdot \sqrt{b} ,$$

$$\Psi'(k') = b^{-\left(\frac{d}{2} + \frac{3}{4}\right)} \Psi(k) , \quad \phi'(k') = b^{-\left(\frac{d}{2} + \frac{3}{4}\right)} \phi(k) , \quad \boxed{\lambda' = \lambda b^{\frac{5}{4} - \frac{d}{2}}} . \quad (3.16)$$

The scaling equation for the Yukawa coupling λ is of particular interest: as λ becomes marginal in $d = 5/2$, we can perform a controlled expansion in $d = \frac{5}{2} - \epsilon$.

Field-theoretical renormalization group

To extract the renormalization group behaviour, the field-theoretical method of choice is "dimensional regularization with minimal subtraction". A detailed account can be found elsewhere [KSF01, Vas04]; here, we just pinpoint the relation to the usual Wilsonian RG.

As the simplest case of illustration, we again consider d-dimensional Euclidean ϕ^4 theory:

$$S_\phi = \int d^d x \frac{1}{2} \left[(\nabla_x \phi(x))^2 + r\phi(x)^2 \right] + \frac{u}{4!} \phi(x)^4 , \quad (3.17)$$

where we introduced some unimportant multiplicity factors. Let us recall the Wilsonian renormalization of the interaction u [Sac11]. One RG step proceeds by integrating out UV scales $k \in [\Lambda/b, \Lambda]$. Perturbatively in u, the leading contribution comes from the standard one-loop diagram L shown below, with external momenta set to zero. After this decimation step, the modified value of u reads

$$L = \begin{array}{c} \times\!\!\!\!\bigcirc\!\!\!\!\times \\ u \quad u \end{array} \quad \Rightarrow \quad \tilde{u} = u + \left(-u^2 \frac{3}{2} \int_{\Lambda/b}^{\Lambda} \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 + r)^2} \right) , \quad (3.18)$$

with a combinatorial factor $\frac{3}{2}$. The new theory effectively has a reduced cutoff Λ/b. In the standard RG, we must then rescale $k' = k \cdot b$ to restore the original cutoff, which amounts to setting $u' = b^{d-4}\tilde{u}$. We then let $b = 1 + dl$ and expand in dl. Furthermore letting $r \to 0$ for brevity, this leads to

$$\frac{u' - u}{dl} = \frac{du}{dl} = (4 - d)u - u^2 \frac{3}{2} S_d \Lambda^{d-4} , \quad (3.19)$$

with S_d a numerical phase factor. Finally, we introduce a dimensionless variable u_r, which gives the ultimate IR beta function

$$u_r = u/\Lambda^{4-d} \ , \quad \frac{du_r}{dl} = (4-d)u_r - u_r^2 \frac{3}{2}S_d \ , \tag{3.20}$$

which has the stable Wilson-Fisher fixed point $u_r^\star \propto (4-d)$; for $4-d = \epsilon \ll 1$, the expansion thus becomes perturbatively controlled.

The field-theoretical renormalization group arrives at the same result [KSF01], but from a different ideological viewpoint. The basic principle is to analyze the theory in the limit $\Lambda \to \infty$. This is physically motivated in fundamental high-energy theories such as QED, since there is no underlying lattice which provides a natural short distance cutoff; in the condensed matter context we can simply view this as a technical trick to derive the flow equations.

In Eq. (3.18), we never used explicitly that Λ is large (but we are certainly free to do so), and the derivation worked for any dimension. If we now let $\Lambda \to \infty$, and want to mimic the decimation of UV scales, we must let $d \nearrow 4$: for dimensions strictly smaller than four, the integral is convergent, and the contribution from $k \sim \Lambda$ is suppressed with powers of Λ. $d > 4$ might also work, but will not be used since we are ultimately interested in $d = (2+1)$. With this in mind, let us look at the interaction part

$$\mathcal{S}_{I,\text{bare}}(u) = \int d^d x \frac{1}{4!} u \phi^4(x) \ . \tag{3.21}$$

for $d = 4 - \epsilon$ from a field-theoretical perspective, and delay the RG for a moment. This term is a function of a **bare** coupling u. As soon as perturbative corrections, such as L in Eq. (3.18), are added, u will change. However, if we evaluate the associated integral without any restriction of momenta, $k \in \mathbb{R}$, this will clearly yield a divergent result as $d \nearrow 4$, such that the modified u will be divergent as well.

To get rid of this unphysical divergence, we make a change of variable, and introduce a renormalized theory by writing

$$S_{I,\text{bare}}(u) = \mathcal{S}_{I,\text{ren}}(\mu, u_r) = \int d^d x \frac{1}{4!} \mu^\epsilon \cdot u_r Z_u \phi^4(x) \ \Rightarrow u = \mu^\epsilon u_r Z_u \ . \tag{3.22}$$

We have done two things: first, we inserted an arbitrary momentum scale μ[5] and boldly called the remaining dimensionless coupling u_r, which parallels the rescaling in Eq. (3.20). μ effectively takes the place of Λ in the Wilsonian approach. Second, we have introduced a renormalization constant Z_u, which should absorb all divergences arising from perturbative corrections (order by order in u_r). Thus, u_r is a finite, physical, and dimensionless coupling. In principle, the fields ϕ will renormalize as well, but we can neglect this renormalization to lowest order in u_r.

To find Z_u, we have to evaluate the diagram L with u replaced by $u_r \mu^\epsilon$; as shown below $Z_u \simeq 1 + u_r$, and thus it does not have to be included in the coupling to leading order in u_r when computing the diagram.

For evaluation of L we will employ the standard extrapolation to continuous dimensions via Gamma functions (dimensional regularization). We can then assign the relevant integral

---

5 Of course, this has nothing to do with the chemical potential μ used before.

a value in any dimension except $d = 4, 6, 8 \ldots$. One obtains [KSF01]:

$$- \frac{2}{3} u_r^2 \mu^{2\epsilon} \int \frac{d^d k}{(2\pi)^d} \frac{1}{(p^2 + m)^2} = \tag{3.23}$$

$$- \frac{2}{3} u_r^2 \mu^{2\epsilon} \frac{\Gamma(\epsilon/2)}{(4\pi)^2} \left(\frac{1}{m^2} \right)^{\epsilon/2} = - \frac{2}{3} \frac{u_r^2 \mu^\epsilon}{(4\pi)^2} \left[\frac{2}{\epsilon} + \mathcal{O}\left(1, \frac{\mu^2}{m}\right) + \mathcal{O}(\epsilon) \right] .$$

Let us call the divergent part (as $\epsilon \to 0$)

$$- \frac{4}{3\epsilon} \frac{u_r^2 \mu^\epsilon}{(4\pi)^2} \equiv u_r \mu^\epsilon \left(-Z_{u,1} \right) \tag{3.24}$$

Without the term Z_u in Eq. (3.22), u_r would pick up a divergent one-loop correction $u_r \to u_r - u_r \mu^\epsilon Z_{u,1}$. To erase this divergence, we effectively subtract it from the action, by finally defining

$$Z_u = (1 + Z_{u,1}) = 1 + \frac{4}{3\epsilon} \frac{u_r}{(4\pi)^2} . \tag{3.25}$$

Note that the divergent part $-Z_{u,1}$ we have subtracted only depends on u_r and not on m, μ, which is known as "minimal subtraction". Such a convenient subtraction method can always be applied in renormalizable theories [KSF01].

So how about the RG? Our goal must be to find the scale dependence of u_r. To this end, we go back to Eq. (3.22), and make the trivial observation that the bare coupling u does not depend on μ. This immediately implies

$$0 = \mu \frac{du}{d\mu} = \epsilon \mu^\epsilon u_r Z_u + \mu^\epsilon \beta_u Z_u + \mu^\epsilon u_r \frac{dZ_u}{du_r} \beta_u , \quad \beta_u = \mu \frac{du_r}{d\mu} . \tag{3.26}$$

Note the following: in the Wilsonian scheme, the contribution of L was added to u; in the field-theoretic scheme, it is effectively subtracted in Z_u. However, in the above we see that $\mu du/d\mu$ should compensate $\mu dZ_u/d\mu$, such that the signs of the flow agree in the Wilsonian and field-theoretical cases.

We can solve Eq. (3.26) for β_u by making the ansatz

$$\beta_u(u_r, \epsilon) = \beta_0(u_r) + \epsilon \beta_1(u_r) , \tag{3.27}$$

and comparing the coefficients $\mathcal{O}(1), \mathcal{O}(\epsilon)$. With this ansatz, we miss contributions to the beta function of order u_r^3/ϵ, but this is fine since our computation is controlled to order u_r^2 only. In this way, we find

$$\beta_u = -\epsilon u_r + u_r^2 \frac{3}{(4\pi)^2} . \tag{3.28}$$

The betafunction so obtained is of UV type, measuring the change of u_r as the momentum scale is increased; to find the IR beta function, we define $\mu = \mu_0 \exp(-l) \Rightarrow \mu d/d\mu = -d/dl$, and thus

$$\frac{du_r}{dl} = \epsilon u_r - u_r^2 \frac{3}{(4\pi)^2} . \tag{3.29}$$

This result almost agrees with our previous one of Eq. (3.20), apart from a different prefactor of the quadratic term. However, this is completely immaterial, as we can always rescale $u_r \to u_r/\alpha$ with an arbitrary number α, and thus change the offending prefactor at will.

While the field-theoretical RG is somewhat less intuitive than the Wilsonian one, it yields the same results, and turns out to be technically much more convenient for the evaluation of hot spot theories. In particular, it spares us from complicated d-dimensional angular integrals that are unavoidable in the Wilsonian scheme when external momenta are retained in the Feynman diagrams.

Validity and key results of the ϵ-expansion

Can we trust the ϵ-expansion? In the classical ϕ^4 problem, it is known to work extraordinary well [KSF01]: up to five-loop results are available, which nicely agree with space-shuttle experiments with superfluid helium. For the computation to yield numerically accurate critical exponents, Borel-type resummation techniques have to be applied, since the ϵ-expansion is an asymptotic series only.

We cannot be so ambitious here, but rather assume that the leading loop ϵ expansion correctly reproduces the order of the phase transitions and the order of magnitude of the critical exponents. On the one hand, this is suggested by the fact that our physical parameter $\epsilon = 5/2 - d$ is indeed somewhat small in the physical case $d = 2$. Further, the validity of the ϵ-expansion can be checked in the SDW case by comparison to an accidental exact solution [SLL17]. One finds that the critical exponents are indeed smooth functions of the space dimension and thus the basic results of the ϵ-expansion should hold, although some subtle problems might inhibit the full knowledge of subleading scaling of the correlation functions [SLL18].

Let us list a few key results in metallic criticality obtained via ϵ-expansion.

○ For most hot spot theories studied so far, the ϵ-expansion produces a stable fixed point at a coupling $\lambda^\star = \mathcal{O}(\epsilon)$, with the mass term being the only relevant perturbation. This is a strong indication that the corresponding phase transition is continuous. To my knowledge, the possibility of a first order transition is only left open for the CDW case in the geometry of Fig. 3.3 (a), where a two-loop analysis indicates a breakdown of the ϵ-expansion [SL16].

○ The RG also allows to obtain the scaling form of bosonic or fermionic correlation functions which determine observable quantities, e.g. in scattering experiments with light or neutrons. The fermionic correlations G_ψ are proportional to scaling functions, with a most general form $f(\omega^{1/z}/k_x, \omega^{1/z}/|k_y|^{2z_y})$. Here, z_y parametrizes the reshaping of the Fermi surface during the RG flow – in the Ising nematic case, $z_y = 1$ is fixed by a "sliding symmetry", but in other problems (such as the CDW case discussed in Sec. 3.3) a different value can arise. The dynamical critical exponent z is typically found to be larger than one, e.g., $z = 3/2$ when extrapolated to $d = 2$ in the Ising nematic case [DL13]. After analytical continuation, a legitimate form of G_ψ is

$$G_\psi(\omega + i0^+) = \frac{1}{\omega - k_x - |k_y|^{2z_y} + i|\omega|^{1/z}} \ . \tag{3.30}$$

This implies a large decay rate $\Gamma \propto |\omega|^{1/z}$, and a breakdown of the quasiparticle picture (quasiparticle weight $Z = 0$). Thus, the RG analysis yields the anticipated non-Fermi liquid fixed point.

○ The ϵ-expansion is also a suitable framework for studying "competing instabilities", in particular superconductivity [SL15, Man16], and the usual finding is that superconductivity is indeed enhanced by the bosonic fluctuations.

○ Last, the ϵ-expansion can also be used to control finite T computations, and resolve thermodynamic observables. A typical scaling result for the specific heat is $C_V(T) \propto T^{(d-\theta)/z}$. Here, θ is an exponent which indicates whether hyperscaling (a simple relation between thermodynamic and correlation function exponents) is obeyed or not. For the SDW, where the critical degrees of freedom are located near isolated hot spots, hyperscaling is obeyed [PSS15] ($\theta = 0$). In the Ising nematic case, where the full Fermi surface is hot, it is violated with $\theta = 1$ [EMS16].

3.2 FFLO quantum criticality

3.2.1 Overview

In the previous discussion of hot spot physics, we mostly avoided the superconducting channel, since it usually does not comply with a hot spot description. As elucidated in [PZS16], there is one exception: superconductivity at a finite wave vector, the so-called Fulde-Ferrell-Larkin-Ovchinnikov (FFLO) order, which arises in strong magnetic fields h.

h couples to the orbital and spin degrees of freedom of the electrons forming the Cooper pairs in a superconductor, and increasing h eventually induces pair breaking. In type-II superconductors, the orbital coupling leads to formation of an Abrikosov vortex lattice, and superconductivity is lost at a critical field h_{orb} where the vortices overlap. The spin coupling favours a spin polarization of the Fermi surface due to the Zeeman effect (Pauli paramagnetism). At a field h_P, the Zeeman energy equals the superconducting condensation energy, destroying superconductivity. The relative strength of these effects is described by the Maki parameter [MT64] $\alpha_M = \sqrt{2}h_{orb}/h_P$, which is typically of order Δ/ϵ_F [MS07], with Δ the superconducting gap and ϵ_F the Fermi energy. Thus, in most cases the orbital effect dominates (i.e., destroys superconductivity at smaller fields), and $\alpha_M \ll 1$. The opposite limit $\alpha_M \gg 1$ can arise in two types of systems: First, α_M can be large in heavy-fermion compounds with a small effective Fermi energy [MS07]. Second, if the magnetic field is applied in-plane in 2D systems, the orbital motion of the electrons is inhibited, and the Zeeman effect wins. We will be interested in the latter systems in particular.

If one just considers the competition of BSC superconductivity and the Zeeman effect, at low temperatures the transition towards the normal state will be first order (as usual in transitions involving two competing energy scales). However, as independently found in the sixties by Fulde and Ferrell [FF64] and Larkin and Ovchinnikov [LO65], at high enough h the energy of the superconducting state can actually be lowered by allowing for Cooper pairs with a non-zero total momentum. For an anistropic Fermi surface, a sketch is shown in Fig. 3.6(a):
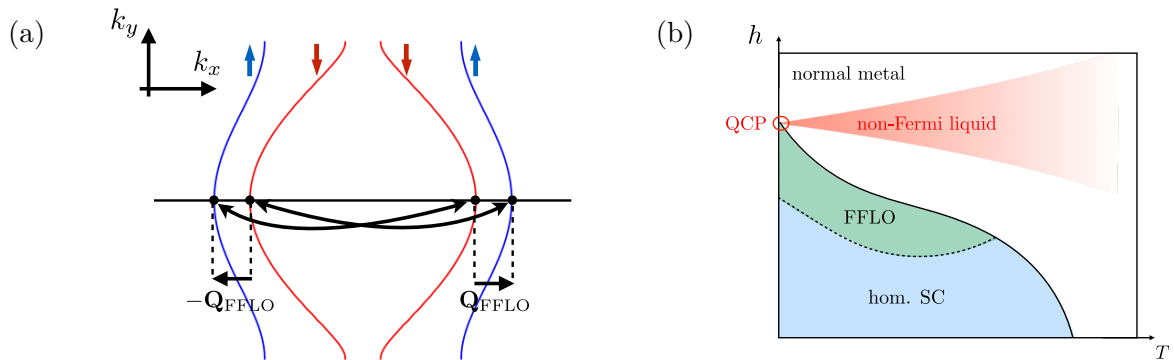

Figure 3.6 FFLO order. (a) Spin polarization and FFLO wave vectors on a anisotropic Fermi surface (b) Phase diagram of an FFLO superconductor with a quantum critical point (QCP) at $T = 0$. The dashed line indicates a first order transition. Fig. is adapted from our publication [PMPP18] and Ref. [PZS16].

Pseudo-nested portions of the spin polarized Fermi surfaces (with Fermi momenta $k_{F,\downarrow}$ and $k_{F,\uparrow}$) are separated by wave vectors $\pm\mathbf{Q}_{FFLO}$. Cooper pairs with momenta $\pm\mathbf{Q}_{FFLO}$ are energetically favoured, leading to an inhomogeneous superconducting gap, $\Delta(\mathbf{x}) \simeq \Delta_0 \cos(\mathbf{Q}_{FFLO} \cdot \mathbf{x})$. The corresponding $h - T$ phase diagram is shown in Fig. 3.6(b): at large h, on mean field level there is a first order transition (dashed line) from the homogeneous BSC superconductor towards the FFLO phase [MS07]. At even higher fields, the normal metal appears via a continuous transition.

Although the FFLO phase has been predicted decades ago, its experimental detection has been hindered by several obstacles for a long time. To begin with, appearance of FFLO order requires very strong ($\simeq 20T$) and perfectly in-plane magnetic fields. FFLO order is also fragile with respect to disorder [Asl69], thus demanding very clean samples. Finally, in isotropic systems the simple phase diagram above is replaced by a cascade of phases with more and more complex order parameters [MC04], and so a sizeable anisotropy is favourable.

These strict requirements are met in organic superconductors [Wos18], which are cleanly grown, layered 2D materials with strongly elliptic Fermi surfaces. In 2007, FFLO was first observed in the organic superconductor $\kappa - (\mathrm{ET})_2\mathrm{Cu(NCS)}_2$ via specific heat measurements [LWD$^+$07]. The experiment nicely reproduced the normal-metal-superconductor transition line of Fig. 3.6(b), in particular its increased slope at small temperatures, which is a clear indication for the FFLO phase. Successful observations in other materials, and with other methods (such as NMR) followed suit [Wos18].

We have only discussed the finite T properties so far. In addition, tuning h at $T = 0$ gives rise to an interesting quantum critical point with an associated non-Fermi liquid scaling fan [see Fig. 3.6(b)], both for the isotropic [SM06] and anisotropic [PZS16] cases. In particular Ref. [PZS16] made interesting predictions for organic superconductors departing from a Bose-Fermi type theory, with the critical degrees centered around the hot spots indicated in Fig. 3.6(a). The role of the order parameter bosons is played by the fluctuating superconducting gap (in Fourier space) $\Delta(\mathbf{q})$. The hot spot action was then analyzed with RPA-like perturbation theory. Similar to previous works, Ref. [PZS16] also contains a Hertz-Millis type argument (see Sec. 3.1.4) for the continuity of the phase transition.

In the article below, our goal was to understand the anisotropic case from a rigorous RG perspective. In particular, as explained before, the mean-field or Hertz-Millis description of FFLO criticality is questionable, and it is essential to retain the fermionic degrees of freedom. The main findings of this analysis are as follows:

○ Via a controlled ϵ-expansion (see Sec. 3.1.5) we have obtained a fixed point at a finite Yukawa-coupling $\lambda^\star$, with only one relevant RG direction $\propto r|\Delta|^2$. This strongly supports the scenario of a continuous phase transition.

○ The computed critical correlations confirm the notion of a non-Fermi-liquid fixed point. For instance, the quasiparticle decay rate has a non-Fermi-liquid power law dependence $\Gamma \propto \omega^{2/3}$. Thermodynamic quantities like specific heat are also estimated to be of non-Fermi liquid type, with $C_V(T) \propto T^{2/3}$ if hyperscaling is violated (although this is a somewhat subtle point).

○ One particularly interesting aspect of FFLO criticality is that the critical point is potentially "naked", and not covered by a (BCS) superconducting dome. To show this, and also scan through other competing instabilities, we evaluated the one-loop anomalous dimension of a generic fermion bilinear $\propto \bar{\psi}\psi, \psi\psi$. Indeed, we do not find an enhancement of ordinary superconductivity by the FFLO fluctuations. The only possible competitor revealed by this analysis is a charge-density wave (CDW) with momentum $2k_{F,\downarrow}$ or $2k_{F,\downarrow}$ (see Sec. 3.3).

○ From a technical point of view, we presented the first analysis of a hot spot theory involving the superconducting channel. One difficulty encountered therein is the appearance of a complicated "$\sqrt{}$"- form in the one-loop Bose propagator. To simplify the calculations, we performed an expansion in the relative Fermi velocity δv of $\downarrow -$ and $\uparrow -$ fermions (see publication), which maps the "$\sqrt{}$" into a Landau damping form. The expansion in small δv seems justified for physical systems, but a better understanding of its implications is desirable and left for future work.

Non-Fermi liquid at the FFLO quantum critical point

by

D. Pimenov,[1] I. Mandal,[2] F. Piazza,[2] and M. Punk[1]

[1] Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for

NanoScience, and Munich Center for Quantum Science and Technology (MCQST),

Ludwig-Maximilians-Universität München, 80333 München, Germany

[2]Max-Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany

Non-Fermi liquid at the FFLO quantum critical point

Dimitri Pimenov,[1,][*] Ipsita Mandal,[2] Francesco Piazza,[2] and Matthias Punk[1]

[1]*Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience, Ludwig-Maximilians-University Munich, 80333 Munich, Germany*
[2]*Max-Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*

When a 2D superconductor is subjected to a strong in-plane magnetic field, Zeeman polarization of the Fermi surface can give rise to inhomogeneous FFLO order with a spatially modulated gap. Further increase of the magnetic field eventually drives the system into a normal metal state. Here, we perform a renormalization group analysis of this quantum phase transition, starting from an appropriate low-energy theory recently introduced in Phys. Rev. B **93**, 085112 (2016). We compute one-loop flow equations within the controlled dimensional regularization scheme with fixed dimension of Fermi surface, expanding in $\epsilon = 5/2 - d$. We find a new stable non-Fermi-liquid fixed point and discuss its critical properties. One of the most interesting aspects of the FFLO non-Fermi-liquid scenario is that the quantum critical point is potentially naked, with the scaling regime observable down to arbitrary low temperatures. In order to study this possibility, we perform a general analysis of competing instabilities, which suggests that only charge density wave order is enhanced in the vicinity of the quantum critical point.

I. INTRODUCTION

A variety of strongly correlated electron materials show unusual metallic behavior, which cannot be described within Landau's Fermi liquid theory. In many cases, this non-Fermi-liquid regime seems to be tied to the presence of a quantum critical point (QCP) between a normal metal and a different symmetry broken phase [1]. One paradigmatic example are certain heavy Fermion materials, where the non-Fermi-liquid regime seems to extend out of a QCP related to the onset of antiferromagnetic order [2].

Of special interest and practical relevance are quasi-two-dimensional systems, where the coupling between electrons and order parameter fluctuations in the vicinity of the QCP is particularly strong. This leads to a loss of electronic quasiparticle coherence due to an intricate interplay between electronic degrees of freedom and the order-parameter dynamics [3–8]. The fact that no well-defined quasiparticle excitations exist in such strongly coupled systems makes the theoretical description of these non-Fermi liquids especially challenging.

Two notable theoretical developments added considerably to our understanding of such non-Fermi liquids. First, it was realized that models of fermions coupled to order parameter fluctuations can be numerically simulated using quantum Monte Carlo techniques avoiding the infamous sign problem under certain conditions [9]. Second, it was shown that field-theoretical approaches can be controlled by increasing the co-dimension of the Fermi surface, which allows for the computation of critical exponents in a systematic epsilon expansion [10,11]. In this work, we will make use of the latter ideas in particular.

So far, most of the theoretical works focused on the experimentally relevant cases of spin-density wave or Ising-nematic critical points in metals. Here, we consider a different problem instead and study the quantum critical point between a normal metal and an inhomogeneous Fulde-Ferell-Larkin-Ovchinnikov (FFLO) superconductor [12,13] in two dimensions. This scenario was put forward by Piazza *et al.* [14], who showed that, for appreciable in-plane anisotropy of the Fermi surface, there is a strong coupling between electrons and FFLO fluctuations in the vicinity of hot spots on the Fermi surface, potentially giving rise to non-Fermi-liquid behavior in the quantum critical regime extending from the QCP at finite temperature, see Fig. 1. A similar treatment of the isotropic case can be found in Ref. [15].

The stabilization of FFLO phases requires clean superconducting materials with suppressed orbital pair breaking effects plus highly anisotropic Fermi surfaces, such as the ones shown by layered materials [16]. Several strong indications of such phases are found in an increasing number of experimental cases, involving organic superconductors [17–20], heavy-fermion systems [21,22], iron-based superconductors [23,24], Al films [25], as well as superconductor-ferromagnet bilayers [26,27].

While the previous study [14] of FFLO non-Fermi liquid criticality was based on a perturbative, RPA-type approach, we will employ the epsilon expansion by Dalidovich and Lee [11] in this work. This allows us to compute critical exponents in a systematic expansion around $d = 5/2$ dimensions, similar to the Ising-nematic problem.

One intriguing aspect of non-Fermi liquids in the vicinity of FFLO critical points is that the QCP is potentially "naked" and not masked by a competing order. Indeed, in the Ising-nematic as well as the SDW scenarios, the order parameter fluctuations give rise to an effective attraction between the electrons,
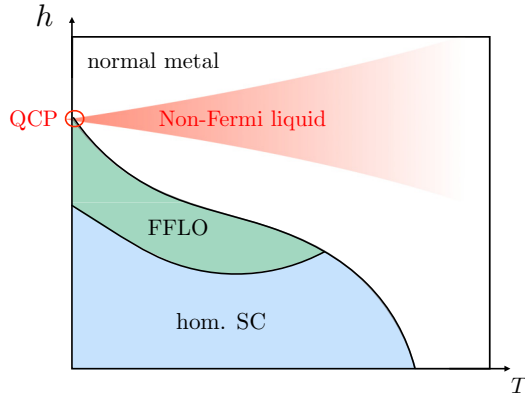
---
[*]d.pimenov@physik.lmu.de

FIG. 1. Typical temperature-magnetic field phase diagram of a superconductor susceptible to FFLO pairing. This picture was adapted from Ref. [14].

burying the QCP deep underneath a superconducting phase [28–35]. One consequence of this competing superconductivity is that the scaling regime of the QCP might be hardly accessible in experiments. By contrast, there is no obvious superconducting order parameter with a different symmetry competing with FFLO superconductivity, which could potentially mask the FFLO QCP. It might be possible, however, that other types of competing orders, such as charge density waves, are enhanced by fluctuations of the FFLO order parameter. We will discuss this issue in detail later in this work.

The rest of this paper is outlined as follows. First, we will give a nontechnical overview of our main results and their physical consequences in Sec. II. Detailed computations are presented in the subsequent sections. In Sec. III, the system under consideration is introduced, studied on mean-field level, and lifted to higher dimensions. In Sec. IV, we discuss one-loop quantum corrections, from which the renormalization group flow and critical properties are derived in Sec. V. Possible competing instabilities are analyzed in Sec. VII. Finally, a conclusion is presented in Sec. VIII. Technical details of the computations are carried out in Appendices.

II. SUMMARY OF RESULTS

An appropriate field-theoretical description of the FFLO-normal metal quantum phase transition has to include dynamics of a bosonic FFLO order parameter Δ (a spatially modulated gap) coupled to the relevant "slow" electronic degrees of freedom ψ. As we show in Sec. III, such a description is accomplished by a low-energy action which contains three parameters $\{m, g, \delta v\}$. Here, m is the "boson mass," respectively, inverse correlation length, which is proportional to the deviation from the critical magnetic field h_c and allows us to tune through the phase transition, g is the strength of the electron-boson coupling (which is proportional to the microscopic electron attractive interaction), and δv is a parameter, which describes the relative spin velocities of the electrons perpendicular to the Fermi surface (which we call the k_x direction).

An RG analysis of this low-energy action, which treats fermions and bosons on equal footing, is the only rigorous way to gain insight into the critical features of the transition,

TABLE I. Critical exponents at the FFLO fixed point, $g = g^\star$. Here, z is the dynamical critical exponent, $\eta_\psi = \eta_\Delta$ are the anomalous dimension of fermions and bosons (which coincide in $\mathcal{O}(\epsilon)$, and ν is the correlation length exponent.

Critical Exponent		Value in $d = 2$ at $\mathcal{O}(\epsilon)$
z	dyn. crit. exponent	$3/2$
$\eta_\psi = \eta_\Delta$	anomalous dim.	$-1/4$
ν	corr. length. exp.	1

see, e.g., Chap. 18 of Ref. [36] for an introduction. In the RG, the parameters of the low-energy action will flow as a function of the energy/length scale. In this work, we study the simplified flow of the interaction parameter g at the quantum critical point ($m = 0$), and also set $\delta v = 0$ for technical reasons.

The first goal of the RG analysis is to locate a fixed point $g = g^\star$, which gives access to critical exponents and correlations. To our knowledge, this was not yet accomplished in the study of FFLO criticality. Using an epsilon-expansion method introduced in the context of metallic quantum critical points [11], we find a stable fixed point corresponding to a continuous transition at $g^\star \propto \epsilon^{3/4}$, where $\epsilon = 5/2 - d = 1/2$.

The critical exponents obtained in our analysis of this new fixed point are presented in Table I. In this table, z is the dynamical critical exponent, which determines how the timelike direction scales compared to the spacelike directions. η_ψ, η_Δ are the anomalous dimensions of the fermions and bosons (which coincide at one-loop level), i.e., the deviation from the scaling determined by power counting for the free theory. ν is the correlation length exponent, given by the inverse RG eigenvalue of the mass term m.

The main value of these critical exponents lies in the fact that they determine the critical correlations, i.e., the electron and boson propagators. In accordance with the RPA-type treatment of Ref. [14] (which is thereby set on solid ground), the scaling forms of the two-point correlators in 2D agree with

$$G(\omega, k_x, k_y) = \frac{1}{i\omega - \delta_k - \Sigma(\omega)},$$
$$\delta_k \propto k_x + k_y^2, \quad \mathrm{Im}[\Sigma(\omega)] \propto g^{4/3}\omega^{2/3} \quad (1)$$

for electrons. For bosons, one obtains

$$D(\omega, k_x, k_y) = \frac{1}{k_y^2 - \Pi(\omega, k_y)}, \quad \Pi(\omega, k_y) \propto -g^2 \frac{|\omega|}{|k_y|}, \quad (2)$$

where Π is the inverse pair propagator. The k_x dependence of the boson propagator is irrelevant in the RG sense. The nonanalytic behavior of the self-energies supports our claim that the quantum critical point is of non-Fermi liquid type. Under assumption of ω/T scaling, signatures of these critical correlations are measurable in the non-Fermi liquid region indicated in Fig. 1. This region is delimited by the two crossover lines satisfying $k_B T \sim |h - h_c|^{z\nu}$ with $z\nu = 3/2$ according to our results. Examples for physical observables include the following.

(1) *Magnetic susceptibility χ*: a simple computation (see Appendix F and Refs. [15,37]) shows that the fluctuation
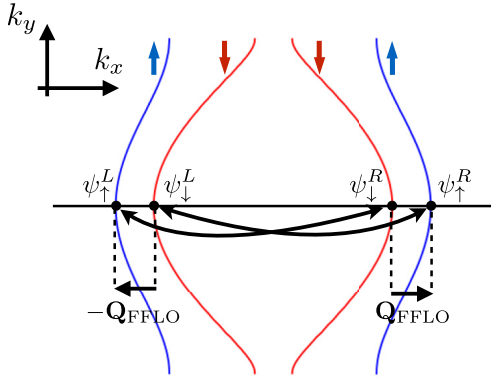
FIG. 2. Typical Fermi surface of an anisotropic metal susceptible to FFLO pairing. Fluctuations of the pairing amplitude Δ strongly couple left branch fermions with right branch fermions with opposite spin at the hot spots. This picture was adapted from [14].

contribution to the magnetic susceptibility χ_Δ scales as $\chi_\Delta \propto \ln(h - h_c)$.

(2) *Fermionic decay rate Γ and density of states $\rho(\omega)$*: from (1) one immediately sees that the quasiparticle decay rate has a non-Fermi-liquid-like power law dependence $\Gamma \propto \omega^{2/3}$, while by integrating the spectral function over momenta [14], one finds $\rho(\omega) \propto \omega^{1/3}$.

(3) *Specific heat capacity*: although the determination of thermodynamic quantities is a somewhat subtle issue (see Sec. VIII), we expect that $C(T) \propto T^{\frac{4}{3} - \frac{2}{3}\theta}$. Here, $\theta = \theta(\delta v)$ is a hyperscaling violation exponent, which should fulfill $\theta(\delta v = 0) = 1$.

Finally, our RG analysis also identifies possible competing orders which may preempt the FFLO transition and lead to a "competing order dome" around the FFLO critical point. We find a charge density wave (CDW) peaked at $2k_F$ to be the most promising candidate. Since $2k_F$ is much larger than Q_{FFLO}, an experiment sensitive to momentum (e.g., using x-ray scattering techniques) could serve to distinguish between the FFLO and CDW orders, although in practice difficulties may arise due to the required low temperatures and high magnetic fields [38].

III. CRITICAL THEORY

A. Critical theory in 2 + 1 dimensions

When an anisotropic 2D metal at $T = 0$ is subjected to a strong in-plane magnetic field h, and orbital effects can be neglected, the electron Fermi surface is spin-polarized. A typical sketch is shown in Fig. 2. Let us now assume that the electrons interact with some generic short-range interaction

$$H_{\text{int}} = -g \int d^2\mathbf{r} \, \psi_\uparrow^\dagger(\mathbf{r}) \psi_\downarrow^\dagger(\mathbf{r}) \psi_\downarrow(\mathbf{r}) \psi_\uparrow(\mathbf{r}), \qquad (3)$$

resulting in Cooper pairing. To derive a low-energy effective action, which makes this pairing explicit, one can perform an exact Hubbard-Stratonovich decoupling of the interaction term (3) in the Cooper channel; thereby, one introduces bosonic fields $\Delta(\mathbf{r})$ and $\bar{\Delta}(\mathbf{r})$ with a free action $\int d\tau d^2\mathbf{r} \, g|\Delta(\mathbf{r})|^2$, which couple to the Fermions in Yukawa-like manner, $\propto g\Delta(\mathbf{r})\bar{\psi}_\uparrow(\mathbf{r})\bar{\psi}_\downarrow(\mathbf{r})$. Due to the spin polarization and the anisotropy of the Fermion dispersion, the bosonic

fields (which correspond to the pairing amplitude) are peaked at momenta $\pm Q_{FFLO} \neq 0$, which is the very definition of the FFLO state. Due to the electron fluctuations, the bosonic mass term gets renormalized, $g \to m \equiv g - \Pi(0,0;h)$, where Π is the inverse pair propagator at vanishing energy-momentum, and we explicitly denoted its magnetic field dependence.[1] As h is increased above the Pauli upper critical field $h = h_c$, the renormalized mass changes from negative to positive values, and the system crosses from the FFLO phase to the normal metal phase along the ($T = 0$) line in the phase diagram of Fig. 1. Accordingly, m is proportional to the reduced magnetic field, $m \propto (h - h_c)/h_c$, in precise analogy to Ginzburg-Landau theory. Further details on the procedure described above are presented in Appendix A, illustrated by a mean-field discussion of the phase transition for a specific microscopic model.

By phase-space considerations, the low-energy fermions at the four hot spots with vanishing curvature in the k_y direction shown in Fig. 2 are most strongly susceptible to pairing, with Cooper pair wave vectors $\pm Q_{FFLO}$. Following the above rationale, a zero temperature action which captures the phase transition between the FFLO and normal metal phases can be readily derived along the lines of Ref. [14] [see Eq. (4) therein]:

$$\mathcal{S} = \int_{k^{2+1}} \sum_{\substack{i = L, R \\ \alpha = \uparrow, \downarrow}} \bar{\psi}_\alpha^i(k) \big(-ik_0 + v_\alpha^i k_x + k_y^2 \big) \psi_\alpha^i(k)$$

$$+ \int_{k^{2+1}} \big(m + k_0^2 + k_x^2 + k_y^2 \big) |\Delta(k)|^2$$

$$- g \int_{k^{2+1}, p^{2+1}} [\bar{\Delta}(k)\psi_\downarrow^L(p)\psi_\uparrow^R(k - p)$$

$$+ \bar{\Delta}(k)\psi_\downarrow^R(p)\psi_\uparrow^L(k - p) + \text{H.c.}], \qquad (4)$$

where $k_0 = \omega$, and

$$v_\alpha^i = \begin{cases} -v_\alpha, & i = L \\ +v_\alpha, & i = R \end{cases}, \quad v_\alpha > 0, \quad \text{and} \quad \int_{k^{d+1}} \equiv \int \frac{d^{d+1}k}{(2\pi)^{d+1}}. \qquad (5)$$

Here, the fermion fields ψ_α^i are expanded around the respective hot spots (see Fig. 2), while the boson fields Δ are expanded around $\pm Q_{FFLO}$. For simplicity, we assume that the pairing is of Larkin-Ovchinnikov type [13], $\Delta(\mathbf{r}) \propto \cos(Q_{FFLO} \cdot \mathbf{r})$, peaked around $\pm Q_{FFLO}$ with equal amplitude.

By the Hubbard-Stratonovich procedure sketched above, the bosons Δ originally just have a mass term $m \propto g$ and no dispersion. However, the kinetic terms and the renormalized mass will be automatically generated during the RG procedure, when high-energy degrees of freedom are integrated out (or, equivalently, arise from the leading analytical boson self-energy corrections involving fermions [14]). Since an action which is appropriate for RG analysis should contain all analytical RG-relevant terms (nonanalytical terms do not renormalize), we include these additional boson terms here from the start. Note that terms $(k_0^2 + k_x^2)|\Delta|^2$ are actually RG-irrelevant by

_____

[1]We perform the Hubbard-Stratonovich decoupling in such a way that $\Pi \propto g^2$, which is why our bare boson mass is g instead of $1/g$.
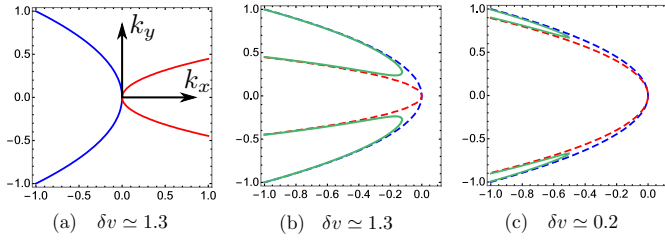
FIG. 3. Fermi surfaces in the FFLO phase at mean-field level. (a) Fermi surface without pairing, i.e., $g\Delta_0 = 0$. (b) Fermi surface for electron/hole operators (see main text), at $\delta v \simeq 1.3$. Dashed lines: $g\Delta_0 = 0$. Full lines: $g\Delta_0 = 0.05$. (c) Same as (b), but $\delta v \simeq 0.2$.

tree-level power counting (see below), which is why we do not need curvature coefficients for them. Alternatively, one can just view the boson terms as expansion in powers and gradients of an FFLO pairing order parameter Δ, as familiar from other non-Fermi-liquid scenarios like Ising-nematic [7] or SDW order [8].

B. Mean-field analysis of superconducting phase

As a first step, let us recall the mean-field level treatment of the action (4) (compare, e.g., Refs. [39–41]) in the superconducting phase, which amounts to the replacement $\Delta(k) \to \Delta_0 \delta^{(3)}(k), |\Delta_0| > 0$. For clarity, we focus on the fermionic branches $\psi_\uparrow^R, \psi_\downarrow^L$, with dispersions

$$\xi_\uparrow^R(k) = v_\uparrow k_x + k_y^2,$$
$$\xi_\downarrow^L(k) = -v_\downarrow k_x + k_y^2. \tag{6}$$

A zoom-in on the respective Fermi surfaces (compared to Fig. 2, momenta are shifted towards a common origin) is shown in Fig. 3(a). The parameter that determines the Fermi surface shapes is the velocity detuning δv:

$$\delta v \equiv 2(v_\uparrow - v_\downarrow)/(v_\uparrow + v_\downarrow). \tag{7}$$

We now introduce Nambu spinors in the standard fashion:

$$\Phi(k) = (\psi_\uparrow^R(k), \overline{\psi}_\downarrow^L(-k))^T. \tag{8}$$

This means that we perform a particle-hole transformation for the spin-down electrons; the Fermi surface of the new fermionic degrees of freedom without pairing is shown in Fig. 3(b) (dashed lines). The mean-field pairing Hamiltonian derived from Eq. (4) is then readily diagonalized by Bogoliubov transformation, with rotated degrees of freedom:

$$\gamma_+(k) = u_k \psi_\uparrow^R(k) - v_k \overline{\psi}_\downarrow^L(-k),$$
$$\overline{\gamma}_-(k) = v_k \psi_\uparrow^R(k) + u_k \overline{\psi}_\downarrow^L(-k). \tag{9}$$

where u_k, v_k are some weights. The corresponding dispersions read

$$E_\pm = \pm \tfrac{1}{2}(\xi_\uparrow^R(k) - \xi_\downarrow^L(-k)$$
$$\pm \sqrt{(\xi_\uparrow^R(k) + \xi_\downarrow^L(-k))^2 + 4g^2|\Delta_0|^2)}. \tag{10}$$

Unlike in the BCS problem, gapless fermionic degrees of freedom remain; the ground state of the system is a condensate

of Cooper pairs with a Fermi sea of $\gamma_\pm$ on top. A plot of the corresponding γ_+ Fermi surface for $\delta v \simeq 1.3$ is shown in Fig. 3(b) (full green line); γ_- fermions are gapped for $\delta v > 0$.

Microscopically, the parameter δv grows monotonously for increasing magnetic fields. This parameter also controls the effectiveness of pairing. Indeed, for $\delta v \to 0$, the full Fermi surface gaps out; the problem becomes BCS-like. This trend is demonstrated in Fig. 3(c), which shows the same quantities as Fig. 3(b), but for a significantly smaller value $\delta v \simeq 0.2$.

As already seen in Fig. 3, the limit $\delta v \to 0$ is rather peculiar. Still, taking this limit will be required later on to gain analytical control over the problem. The implications of this procedure will be discussed in more detail below [(see Sec. IV and Appendix D)].

C. Critical theory in $d + 1$ dimensions

Let us now focus on the phase transition from the FFLO to the normal metal phase, which can be driven by tuning the boson mass m in Eq. (4) from negative to positive values. Going beyond a Landau-Ginzburg type analysis of the phase transition (as found, e.g., in Refs. [42,43]), we will treat both bosons and fermions as dynamical degrees of freedom, and look for the critical RG fixed point of the action (4) in the IR. However, this fixed point is located at strong coupling; to access it perturbatively, we must introduce a small parameter ϵ into the action which suppresses quantum fluctuations. A convenient way of doing so is to increase the space dimension d, thereby successively tuning the Yukawa interaction between bosons and fermions marginal as d approaches the critical dimension d_c. For $d = d_c$, the interacting critical fixed point then collapses with the noninteracting Gaussian one, and we can therefore derive RG flow equations perturbatively in $\epsilon = d_c - d$.

In the presence of a Fermi surface, one may increase the number of dimensions tangential or perpendicular to it [44,45]. Some aspects of the scheme with increased tangential dimensions (or fixed codimension), where $d_c = 3$, are outlined in Appendix E; in short, this extension is problematic because it leads to a breakdown of the hot spot theory in the parameter regime where the computations are analytically tractable. Let us therefore follow [10,11] and increase the perpendicular dimensions. That is, the Fermi surface is always one-dimensional, and the fermionic density of states is successively reduced. This amounts to an expansion around $d_c = 5/2$.

To implement this dimensional extension in practice, we employ the formalism and techniques introduced in Ref. [11], where renormalization group equations are computed within the dimensional regularization (called DIMREG henceforth) and minimal subtraction schemes (see Refs. [46,47] for an introduction). We will work at $T = 0$; thermal fluctuations on a different, isotropic model for the FFLO transition were recently studied in Ref. [48] with functional RG methods.

For shorter notation, we define fermionic "spinors" Ψ:

$$\Psi_\alpha(k) = \begin{pmatrix} \psi_\alpha^R(k) \\ \overline{\psi}_\alpha^L(-k) \end{pmatrix}, \quad \overline{\Psi}_\alpha(k) = (\overline{\psi}_\alpha^R(k), \psi_\alpha^L(-k)) \cdot \sigma_y, \tag{11}$$

where σ_y is a Pauli matrix. The kinetic term for the fermions can then be generalized to $d + 1$ dimensions as

$$\sum_{\alpha=\downarrow,\uparrow} \int_{k^{d+1}} \bar{\Psi}_\alpha(k)(-i\mathbf{\Gamma} \cdot \mathbf{K} + i\sigma_x \delta_\alpha(k))\Psi_\alpha(k). \quad (12)$$

Here, $\mathbf{K} = (k_0, k_1, \ldots, k_{d-2})$, and the momenta k_x, k_y are relabeled as $k_x \to k_{d-1}, k_y \to k_d$. δ_α is the right branch fermion dispersion, $\delta_\alpha = v_\alpha k_{d-1} + k_d^2$. $\mathbf{\Gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_{d-2})$ is a vector of two-dimensional Gamma matrices, which fulfills the Clifford algebra, $\{\gamma_\alpha, \gamma_\beta\} = 2\delta_{\alpha\beta}$. In the integer cases of interest:

$$d = 2: \quad \mathbf{K} = k_0, \quad \mathbf{\Gamma} = \sigma_y, \quad (13)$$

$$d = 3: \quad \mathbf{K} = (k_0, k_1), \quad \mathbf{\Gamma} = (\sigma_y, \sigma_z). \quad (14)$$

To uniquely specify the Gamma-matrix structure, in general dimensions, we choose the continuation

$$\mathbf{\Gamma} = (\sigma_y, \vec{\Gamma}) = (\sigma_y, \sigma_z, \ldots \sigma_z), \quad (15)$$

where the "vector" $\vec{\Gamma}$ has $(d - 2)$ entries.

The introduction of generalized Gamma matrices is a standard tool in DIMREG of fermionic theories, see, e.g., Ref. [49]. In the condensed matter context, an alternative point of view is the following: we add *one* extra dimension perpendicular to the Fermi surface, extending the action with terms of triplet-pairing form [11]:

$$\bar{\Psi}_\alpha(k)(-ik_1\sigma_z)\Psi_\alpha(k) = k_1 \left(\sum_{\alpha=\downarrow,\uparrow} \bar{\psi}_\alpha^R(k)\bar{\psi}_\alpha^L(-k) + \text{H.c.} \right). \quad (16)$$

These terms gap out the Fermi surface except for the one-dimensional branches of Fig. 2. In all computations, we then continuously tune the "weight" of this extra dimension, by using a radial integral measure $\int dk_1 k_1^{d-3}/(2\pi)^{(d-2)}$. It should be noted that by introducing these extra terms we have broken the spin-rotation symmetry in the xy plane of the original action (4).

The kinetic term for the bosons in the $(d + 1)$-dimensional action generalizes to

$$\int_{k^{d+1}} \left(|\mathbf{K}|^2 + k_{d-1}^2 + k_d^2 + m \right) |\Delta(k)|^2. \quad (17)$$

The terms in the noninteracting parts of the action (12) and (17) are invariant under the scaling transformations:

$$\mathbf{K} = \frac{\mathbf{K}'}{b}, \quad k_{d-1} = \frac{k'_{d-1}}{b}, \quad k_d = \frac{k'_d}{\sqrt{b}},$$
$$\Psi(k) = b^{\frac{d}{2}+\frac{3}{4}} \Psi'(k'), \quad \Delta(k) = b^{\frac{d}{2}+\frac{3}{4}} \Delta'(k'). \quad (18)$$

At tree level, the terms $(|\mathbf{K}|^2 + k_{d-1}^2)|\Delta|^2$ are irrelevant, and will stay so in ϵ expansion as long as ϵ is small. Let us therefore erase these terms from the action. Furthermore, as we are mostly interested in the quantum critical point, we will set the renormalized mass $m = 0$ in the following. The IR divergences resulting from these two steps can be regularized by using dressed boson propagators in all computations [11].

Inserting the spinor definitions (11), the interaction term is easily rewritten in higher dimensions. In total, the critical
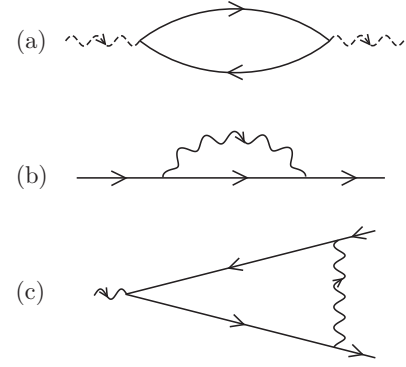

FIG. 4. One-loop diagrams. Dashed wavy lines (a) indicate bare boson propagators, while straight lines indicate electron propagators. Full wavy lines [(b) and (c)] represent bosons dressed with the self-energy of (a). External lines are amputated.

action in $d + 1$ dimensions then reads

$$S = \int_{k^{d+1}} \bar{\Psi}_\alpha(k)\left(-i\mathbf{\Gamma} \cdot \mathbf{K} + i\sigma_x(v_\alpha k_{d-1} + k_d^2)\right)\Psi_\alpha(k)$$
$$+ \int_{k^{d+1}} k_d^2 |\Delta(k)|^2 - g\mu^{\epsilon/2} \int_{k^{d+1}, p^{d+1}} \left[\bar{\Delta}(k)\sigma_y^{\alpha\alpha'}\bar{\Psi}_\alpha(-p)\right.$$
$$\times M_1\Psi_{\alpha'}(k - p) + \Delta(k)\sigma_y^{\alpha\alpha'}\bar{\Psi}_\alpha(k - p)M_2\Psi_{\alpha'}(-p)\Big], \quad (19)$$

where we introduced matrices acting in spinor space

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (20)$$

and employed a summation convention for spin indices. Note that the pairing terms of the original action (4) have the form of a standard density term in the spinor language. We have made the tree-level scaling dimension of the interaction explicit by replacing $g \to g\mu^{\epsilon/2}$, where μ is an arbitrary mass scale, and

$$\epsilon = 5/2 - d. \quad (21)$$

In the standard logic of ϵ-expansion, we will work in the limit $\epsilon \to 0$, where the interaction term becomes marginal, and determine the critical exponents at the interacting fixed point to order ϵ. Extrapolating to the physically relevant value $\epsilon = 1/2$, we can then make a controlled qualitative estimate of critical exponents and the universality class of the problem.

IV. ONE-LOOP DIAGRAMS

To compute the flow equations in DIMREG, one needs to evaluate the possible one-loop corrections to the action (19), whose diagrammatic representations are shown in Fig. 4. Note that tadpole contributions to the fermion self-energy are disregarded since they can renormalize the chemical potential only. Higher loop diagrams are multiplied with a higher power of the coupling g. Below, we will show that $g \propto \epsilon^{3/4}$ at the critical point, thus higher-loop diagrams are suppressed for $\epsilon \to 0$. In this work, we will disregard them altogether. To evaluate these diagrams analytically, we need to make one important approximation: we consider the limit of *vanishing velocity detuning*, $\delta v \to 0$ [cf. Eq. (7)].
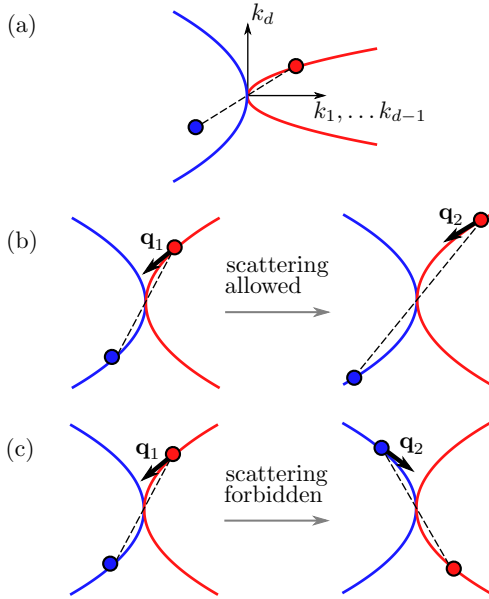
FIG. 5. Zoom-in of Fig. 2, showing the two Fermi surface branches of fermions coupled by pairing fluctuations at $+\mathbf{Q}_{\mathrm{FFLO}}$, shifted towards a common origin in momentum space. (a) Case of nonzero velocity detuning $\delta v \neq 0$, where only fermions with momenta close to $k = 0$ are strongly entangled. (b) Case of vanishing velocity detuning $\delta v \to 0$, where the electrons with momenta $\pm \mathbf{k}_1$ and tangent vector $\mathbf{q}_1$ can be scattered to close-by momenta $\pm \mathbf{k}_2$ with similar tangent vector $\mathbf{q}_2$. (c) Same initial configuration as Fig. 5(b), but with different final momenta $\pm \mathbf{k}_2$; now, the final tangent vector $\mathbf{q}_2$ differs strongly from the initial one, and the phase space for the scattering is negligible.

In a realistic experimental setup, $\delta v = \mathcal{O}(0.1)$ [14] is indeed small. However, the limit $\delta v \to 0$, while being computationally convenient, is somewhat singular, as already indicated in Sec. III B. This can be seen pictorially in Fig. 5: for nonvanishing velocity detuning [Fig. 5(a)], two Fermi surface branches interacting with each other have different curvatures. Thus only electrons with momenta close to the hot spot at $\mathbf{k} = 0$ (the branches are shifted towards a common origin) scatter strongly with FFLO fluctuations. For any electron close to the Fermi surface with large momentum $\mathbf{k}$ away from the hot spot [red dot in Fig. 5(a)], the corresponding electron with momentum $-\mathbf{k}$ (indicated by a dashed line and a blue dot), which would be most susceptible to FFLO pairing, has momentum far from the Fermi surface, and thus pairing is suppressed.

On the other hand, if the two spin velocities are equal [Figs. 5(b) and 5(c)], an arbitrary electron on the Fermi surface with momentum $\mathbf{k}_1$ can scatter against its counterpart with momentum $-\mathbf{k}_1$, as also demonstrated in Sec. III B. However, the FFLO fluctuations can only scatter these electrons efficiently into a pair of electrons with momenta $\pm \mathbf{k}_2$, such that $\mathbf{k}_2 \simeq \mathbf{k}_1$. The tangent vector to the Fermi surface of the initial pair $\mathbf{q}_1$ must almost coincide with the final tangent vector $\mathbf{q}_2$, as shown in Fig. 5(b). If $\mathbf{q}_2 \not\simeq \mathbf{q}_1$, as shown in Fig. 5(c), the scattering process is energetically suppressed. The fact that scattering processes are only local in momentum space prevents the explicit appearance of UV scales and thereby

justifies application of the hot spot theory. Note that this argument remains true only as long the Fermi surface is strictly 1D; for higher-dimensional Fermi surfaces, which arise in the RG scheme with fixed codimension, the limit $\delta v \to 0$ is even more singular and results in UV-IR mixing [44], eventually leading to a breakdown of the hot spot expansion; see Appendix E for further details.

Despite its smallness, in a fully fledged RG analysis of the problem, δv should be treated as a running coupling. We will leave this involved task for future (numerical) work, and focus on $\delta v \to 0$ from now on, which should be qualitatively correct as long as δv does not exhibit a runaway flow in the full RG procedure.

Let us now evaluate the boson self-energy Π of Fig. 4(a). This diagram dresses the bare boson Green's function

$$D_0(k) \equiv \langle \Delta(k) \bar{\Delta}(k) \rangle_0 = 1/k_d^2, \qquad (22)$$

where the subscript 0 indicates that averages are taken with respect to the noninteracting action and reads

$$\Pi(k) = -g^2 \mu^\epsilon \int_{p^{d+1}} \sum_{\alpha \neq \alpha'} \mathrm{Tr}[G_\alpha(-p) M_1 G_{\alpha'}(k-p) M_2]. \qquad (23)$$

Here, the electron Green's function is defined by

$$G_\alpha(k) \equiv \langle \Psi_\alpha(k) \bar{\Psi}_\alpha(k) \rangle_0 = -i \frac{-\mathbf{\Gamma} \cdot \mathbf{K} + \sigma_x \delta_k}{\mathbf{K}^2 + (\delta_k)^2}, \qquad (24)$$

where $\delta_k = k_{d-1} + k_d^2$, i.e., we have scaled out the equal velocities. Evaluation of (23) is done in Appendix B 1 and yields

$$\Pi(k) = \chi_d \frac{g^2 \mu^\epsilon}{|k_d|} \left(d\, k_0^2 + |\vec{k}|^2 \right) \left(k_0^2 + |\vec{k}|^2 \right)^{\frac{d-3}{2}}, \qquad (25)$$

with

$$\chi_d = \frac{\Gamma((1-d)/2)}{2^{d+2} \pi^{(d+1)/2}} \frac{\Gamma(d/2)^2}{\Gamma(d)},$$
$$\chi_{5/2} \simeq -0.0178. \qquad (26)$$

In Eq. (25), $\vec{k}$ are the extra dimensions inserted in the DIMREG scheme, i.e., $\mathbf{K} = (k_0, \vec{k})$. The fact that we have an anisotropy in $\mathbf{K}$ space is a peculiarity of the original pairing vertex, leading to a matrix structure in spinor space with matrices M_1 and M_2 [see Eqs. (19) and (20)], which are not Gamma matrices. This anisotropy can be easiest understood taking the fermion self-energy as an example, see below. For $d = 2$, there are no extra dimensions, and Eq. (25) simplifies to the 2D result found in Ref. [14].

Two further comments on the result (25) are in order. First, to arrive at (25), we had to make a trivial regularization by subtracting $\Pi(0, k_d)$ (in any dimension). The residual momentum dependence of this subtraction is an artefact of the $\delta v \to 0$ limit; for $\delta v \neq 0$, at least in the physical case $d = 2$, one obtains a finite result for the self-energy by subtracting $\Pi(\delta v \neq 0, k = 0)$. If we could take $\delta v \to 0$ in the last step of the computation, i.e., before dropping momentum cutoffs, this trivial mass renormalization (which is perfectly legitimate as we focus on the critical point where the boson is massless) would always suffice. However, in practice we have to take
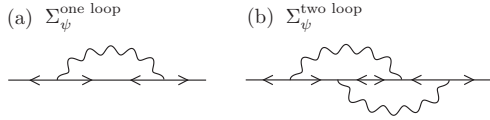
FIG. 6. Anomalous contributions to the fermion self-energy (in terms of the original fermion fields ψ). At one-loop, such contributions are impossible (a), but can arise at two loop (b).

the limit $\delta v \to 0$ first, and subtract $\Pi(0, k_d)$ (which amounts to a "superconducting logarithm") in effect. A more detailed justification of this step is presented in Appendix D. Second, although at first glance of the Fermi surface of Fig. 2 one could expect Π to have a SDW-type behavior $\Pi(k) \sim |\mathbf{K}|$ [8], our result (25) is a standard Landau damping term familiar from the Ising-nematic case [11], apart from the anisotropy discussed above. This is again a consequence of the pairing structure of the original vertex.

As in the Ising-nematic case, the boson self-energy is UV finite as $d \to 5/2$. Still, this contribution is crucial, as the further loop corrections of Figs. 4(b) and 4(c) are only IR finite if the boson lines are taken to be dressed, which we will do in the following, compare Ref. [11].

Let us now evaluate the fermion self-energy of Fig. 4(b). For a fermion of spin κ, there are two contributions,

$$\Sigma_1^\kappa(k) = g^2 \mu^\epsilon \int_{p^{d+1}} D(p) M_1 G_\beta(k+p) M_2 \sigma_y^{\kappa\beta} \sigma_y^{\beta\kappa}, \quad (27)$$

$$\Sigma_2^\kappa(k) = g^2 \mu^\epsilon \int_{p^{d+1}} D(-p) M_2 G_\beta(k+p) M_1 \sigma_y^{\kappa\beta} \sigma_y^{\beta\kappa}, \quad (28)$$

representing the two ways to draw the arrow on the boson line. Evaluating these integrals in leading order in ϵ (see Appendix B 2), we obtain

$$\Sigma(k) = \Sigma_1^\kappa(k) + \Sigma_2^\kappa(k)$$
$$= \frac{u_g g^{4/3}}{\epsilon} \sigma_y (-i k_0) + \text{finite terms},$$
$$u_g \simeq -0.0813. \quad (29)$$

Thus we find that the fermion self-energy only depends on the frequency, and not on the extra momenta $\vec{k}$ as for the Ising-nematic [11]. This is easily understood as follows: as discussed before, see Eq. (16), insertion of extra dimensions $\vec{k}$ gives rise to triplet pairing terms already at the noninteracting level, or, in other words, to anomalous terms in the bare fermion Green's function $\propto \vec{k}$, when expressed in terms of the original fermion fields ψ (see, e.g., Ref. [50]). Therefore, to obtain a contribution to $\Sigma(k) \propto \vec{k}$, there must be an anomalous contribution to the self-energy. However, this is not possible at one loop. This is seen pictorially in Fig. 6(a), which shows an impossible diagram (since four fermions are annihilated at the vertices) in terms of original fermion fields. Note that at higher loop level such contributions can arise, see Fig. 6(b).

Last, we need to compute the vertex correction of Fig. 4(c). In $d = 2$, this diagram is trivially absent, but not in $d > 2$ (due to the anomalous terms). However, we still find that there is no ϵ-divergent vertex correction; further details are relocated to Sec. VII, where we discuss general vertex corrections that reflect possible competing orders.

V. RENORMALIZATION

A. Flow equation

To obtain a UV finite renormalized action, we have to add the fermion self-energy as a counter-term, employing the minimal subtraction scheme where the counterterm depends on g only:

$$S_{\text{CT}} = \sum_{\alpha=\downarrow,\uparrow} \int_{k^{d+1}} \frac{Z_{1,1}(g)}{\epsilon} \bar{\Psi}_\alpha(k)(-i\sigma_y k_0)\Psi_\alpha(k),$$

$$Z_{1,1}(g) = u_g \, g^{4/3}. \quad (30)$$

Then, the renormalized action is obtained as $\mathcal{S}_{\text{ren}} = \mathcal{S} + \mathcal{S}_{\text{CT}}$. We define a renormalization constant $Z_1 = 1 + Z_{1,1}/\epsilon$ and introduce unrenormalized (bare) fields and couplings as

$$k_0^b = k_0 Z_1, \quad \vec{k}^b = \vec{k}, \quad k_{d-1}^b = k_{d-1}, \quad k_d^b = k_d,$$

$$\Psi_b(k_b) = Z_1^{-1/2}\Psi(k), \quad \Delta_b(k_b) = Z_1^{-1/2}\Delta(k),$$

$$g^b = Z_1^{-1/2}\mu^{\epsilon/2} g. \quad (31)$$

These relations bring the renormalized action back in the form of the initial bare action (19) except for the dimensionful coupling g^b:

$\mathcal{S}_{\text{ren}}$

$$= \int_{(k^b)^{d+1}} \bar{\Psi}_\alpha^b(k^b)\left(-i\mathbf{\Gamma} \cdot \mathbf{K}^b + i\sigma_x(k_{d-1}^b + (k_d^b)^2)\right)\Psi_\alpha^b(k^b)$$

$$+ \int_{(k^b)^{d+1}} (k_d^b)^2 |\Delta^b(k^b)|^2$$

$$- g^b \int_{(k^b)^{d+1},(p^b)^{d+1}} \left[\bar{\Delta}^b(k^b)\sigma_y^{\alpha\alpha'}\bar{\Psi}_\alpha^b(-p^b)M_1\Psi_{\alpha'}^b(k^b - p^b)\right.$$

$$\left. + \Delta^b(k^b)\sigma_y^{\alpha\alpha'}\bar{\Psi}_\alpha^b(k^b - p^b)M_2\Psi_{\alpha'}^b(-p^b)\right]. \quad (32)$$

Let us determine the flow of the renormalized coupling g at a fixed UV value of the bare coupling g^b as the mass scale μ is decreased. It is described by the beta function

$$\beta = \frac{dg}{d\ln(\mu)}, \quad (33)$$

which fulfills the equation

$$\beta\left(\frac{g}{2}Z_1' - Z_1\right) - \frac{\epsilon}{2}g Z_1 = 0. \quad (34)$$

We may solve it making the standard ansatz $\beta = \beta_0 + \epsilon\beta_1$, where $\beta_{0,1}$ depend on g only. Comparing the coefficients of the parts regular in ϵ of Eq. (34) yields[2]

$$\beta = -\frac{u_g}{3}g^{7/3} - \frac{\epsilon}{2}g. \quad (35)$$

The beta function has a fixed point at

$$g^\star = \left(\frac{3\epsilon}{-2u_g}\right)^{3/4}, \quad u_g \simeq -0.0813. \quad (36)$$

---

[2] Note that the solution (35) violates Eq. (34) at order $g^{11/3}/\epsilon$. This is a standard artefact of approximating the renormalization constant $Z_1 \simeq 1 + Z_{1,1}/\epsilon$ at one-loop level, and should be successively improved by higher loop contributions.

Writing $\mu = \mu_0 e^{-\ell}$, the RG eigenvalue of g at $g = g^\star$ in the IR ($\ell \to \infty$) is $-\frac{2}{3}\epsilon$, i.e., the fixed point is stable (respectively, critical, as we have dropped the RG relevant mass term from the action). This indicates a second-order phase transition between the FFLO and normal metal phases. A continuous transition was also found in the mean-field study of our precursor work [14] and other 2D studies [51,52].

B. Critical properties

Let us discuss critical properties of this new fixed point, which are intimately linked with experimental observables. First, we define the dynamical critical exponent z:

$$z = 1 - \frac{d\ln(1/Z_1)}{d\ln(\mu)} = 1 + \frac{1}{Z_1}Z_1'\beta. \qquad (37)$$

At the fixed point, we find

$$z^\star = 1 + \epsilon. \qquad (38)$$

From the renormalization of fields in Eq. (31), the anomalous dimensions of bosons and fermions read

$$\eta_\Psi = \eta_\Delta = \frac{1}{2}\frac{d\ln(1/Z_1)}{d\ln(\mu)} = \frac{1-z}{2} = -\frac{\epsilon}{2}\bigg|_{z=z^\star}. \qquad (39)$$

z and η feed into the scaling behavior of correlation functions, which can be determined in the standard way, defining renormalized Green's functions by

$$\langle \Psi(k_1)\dots\Psi(k_m)\bar{\Psi}(k_{m+1})\dots\bar{\Psi}(k_{2m})\Delta(k_{2m+1})\dots$$

$$\Delta(k_{2m+n})\bar{\Delta}(k_{2m+n+1})\dots\bar{\Delta}(k_{2m+2n})\rangle$$

$$= G^{(m,m,n,n)}(\{k_i\}; g, \mu)$$

$$\times \delta^{d+1}\left(\sum_{i=1}^{m}k_i + \sum_{i=2m+1}^{2m+n}k_i - \sum_{j=m+1}^{2m}k_j - \sum_{j=2m+n+1}^{2m+2n}k_j\right), \qquad (40)$$

with spin and space-time indices suppressed. These correlators are related to the bare ones derived from the bare action (32) by multiplicative renormalization and fulfill the scaling equation

$$\left\{ \sum_{i=1}^{2m+2n} z\, k_{i,0}\frac{\partial}{\partial k_{i,0}} + \vec{k}_i\nabla_{\vec{k}_i} + k_{i,d-1}\frac{\partial}{\partial k_{i,d-1}} + \frac{k_{i,d}}{2}\frac{\partial}{\partial k_{i,d}} \right.$$

$$- \beta\frac{\partial}{\partial g} - 2m\left(\eta_\Psi - \frac{4-\epsilon}{2}\right) - 2n\left(\eta_\Delta - \frac{4-\epsilon}{2}\right)$$

$$\left. + (\epsilon - z - 2) \right\} G^{(m,m,n,n)}(\{k_i\}; g, \mu) = 0. \qquad (41)$$

At the fixed point where $\beta = 0$, and the RG exponents are given in Eqs. (38) and (39), Eq. (41) implies a scaling form of the fermion two-point function:

$$G(\mathbf{k}) \propto \frac{1}{\delta_k} f\left(\frac{k_0^{1/(1+\epsilon)}}{\delta_k}, \frac{|\vec{k}|}{\delta_k}\right), \quad \delta_k = k_{d-1} + k_d^2, \qquad (42)$$

where f is a universal scaling function. In particular, in $d = 2$ ($\epsilon = 1/2$), this scaling form is consistent with the fermion self-energy $\propto k_0^{2/3}$ obtained in Ref. [14]. We therefore find, for $\epsilon > 0$, non-Fermi-liquid behavior where the quasiparticle nature of fermions is destroyed by strong order parameter fluctuations;

exactly at $\epsilon = 0$, the system is a marginal Fermi liquid. For bosons, one finds the same scaling form as in Eq. (42) with δ_k replaced by k_d^2:

$$D(\mathbf{k}) \propto \frac{1}{k_d^2} f\left(\frac{k_0^{1/(1+\epsilon)}}{k_d^2}, \frac{|\vec{k}|}{k_d^2}\right). \qquad (43)$$

Apart from the critical correlations (42), also the scaling behavior on the normal metal side is of interest, characterized by the correlation length exponent ν. To find it, we need to include a mass perturbation $m|\Delta|^2$ in the action, and ν is given by the inverse RG eigenvalue of m. Then, we need to compute the boson self-energy $\Pi(0)$—the mass will acquire an anomalous dimension if $\Pi(0)$ shows a (logarithmic) $1/\epsilon$ divergence. In our evaluation of Π in Appendix B 1, such a logarithmic divergence does not arise, at least at one-loop in the analytically controlled limit $\delta v \to 0$. By power counting, we can thus conclude

$$\nu = 1 + \mathcal{O}(\epsilon^2). \qquad (44)$$

What is more, our theory is similar to the nematic case, where the boson self-energy does not diverge up to 3-loop, [11]. So, we can expect that the estimate (44) holds to higher loop level as well.

VI. PHYSICAL OBSERVABLES

Equations (42) and (43), obtained in a controlled perturbative procedure, are the major result of this work. Equation (43) tells us the scaling form of the pair susceptibility D. For ordinary BCS [53–55] as well as unconventional high-T_c [56] superconductors, the imaginary part of this quantity is proportional to the Josephson current in a SIN junction setup for a small applied bias voltage; it remains to be seen if this idea can be carried over to FFLO superconductors. Furthermore, by integration over D^2 (see Appendix F), one can obtain the fluctuation contribution to the spin susceptibility χ_Δ in the normal state. For $d = 2$, we find a weakly divergent behavior as a function of the reduced magnetic field, $\chi_\Delta \propto \ln[(h - h_c)/h_c]$. This is in agreement with the RPA result of Ref. [15].

The correlator G in Eq. (42) describes the fate of electronic excitations. In $d = 2$, they decay in non-Fermi-liquid manner, with a large rate $\Gamma(k_0) \propto k_0^{2/3}$. The hot-spot density of states $\rho(k_0)$ of these excitations can be found by integrating the electronic spectral function over momenta [14], $\rho(k_0) \propto k_0^{1/3}$. In addition, a constant contribution to $\rho(k_0)$ from the cold, Fermi-liquid-like parts of the Fermi surface will arise.

As long as ω/T scaling is not violated [57–59], these overdamped excitations will strongly influence the temperature dependence of observables within the quantum critical region of Fig. 1. This region is delimited by the two crossover lines satisfying $k_B T \sim |h - h_{QCP}|^{z\nu}$ with $z\nu = 3/2$ according to our results. For instance, one can extract the critical contribution to the specific heat, which scales as $C \propto T^{(d-\theta)/z} = T^{\frac{4}{3} - \frac{2}{3}\theta}$. Here, θ is an exponent which describes hyperscaling violation. Usually, hyperscaling violation occurs in systems with a critical Fermi surface, where the integral of the singular part of the free energy along the entire Fermi surface alters the thermodynamic properties [60]. In the context of the FFLO critical point discussed here, hyperscaling violation is not expected to occur for a sizable velocity detuning δv, when the critical degrees

of freedom live in the vicinity of isolated hot spots. Then, $\theta = 0$ and therefore $C \propto T^{4/3}$. This is similar to the SDW hot spots studied in Refs. [61,62]. By contrast, for the case of vanishing velocity detuning to which our RG computation was restricted, the entire Fermi surface becomes hot. As a result, one expects a hyperscaling violation exponent $\theta = 1$ and therefore $C \propto T^{2/3}$. We emphasize again, however, that the hot spot theory (our field theoretical starting point) remains applicable in this limit as well: the infinite set of hot spot pairs decouple in the low energy limit, because electrons can only scatter with small momentum transfer tangential to the Fermi surface, similar to the Ising-nematic case. For this reason, we are confident that our RG computation remains valid for finite velocity detunings as well, even though thermodynamic observables may depend strongly on the velocity detuning via the hyperscaling violation exponent θ.

From the low-energy form of $\rho(k_0)$ of the hot quasiparticles, one can also make a prediction for the temperature dependence of the NMR relaxation rate, $1/(T \, T_1) \propto T^{2/3}$ [14]. Note that for strong velocity detuning, the cold electrons give an additional constant contribution to $1/(T \, T_1)$ (Korringa law).

In organic superconductors, measurements of specific heat [63,64] and NMR rates [19] within the putative quantum-critical region have been already taken. While one may see indication for non-Fermi-liquid behavior in the data (see Ref. [14]), quantitative statements and meaningful estimates on critical exponents cannot be made yet. A new round of data taking on a larger temperature interval might provide a conclusive insight.

VII. COMPETING ORDERS

Non-Fermi liquid fixed points, where the critical correlations take a form similar to Eqs. (42) and (43), arise in numerous physical contexts. As discussed above, in principle, the zero-temperature form of the correlations manifests itself in a quantum-critical region at finite temperatures, see Fig. 1. However, the critical scaling is often masked by a "dome" of a competing, mostly superconducting order [31,33,62,65], at least for conventional critical points associated with the onset of broken symmetry [32]. The FFLO-normal metal fixed point is different in this regard: since we deal with a phase transition towards superconductivity already, one can expect the fixed point to be "naked." Other superconducting orders, e.g., of triplet type, may of course occur, but seem unlikely given the Fermi surface geometry of Fig. 2, in accordance with a recent Monte-Carlo study of a Hubbard model with spin imbalance [66].

Going beyond these naive expectations, one may answer the question how competing instabilities are modified close to our new non-Fermi-liquid fixed point systematically in the DIMREG framework; following the treatment of Ref. [67], we consider the insertion of a generic fermion bilinear into the critical action (19). In the spinor language, this term can be of two types: either

$$\text{type 1}: \quad \lambda \int_{k^{d+1}} \bar{\Psi}_{\alpha}(k) A \Psi_{\beta}(k) B_{\alpha\beta} \quad \text{or} \quad (45)$$

$$\text{type 2}: \quad \lambda \int_{k^{d+1}} \Psi_{\alpha}^{T}(k) A \Psi_{\beta}(-k) B_{\alpha\beta} + \text{H.c.}, \quad (46)$$
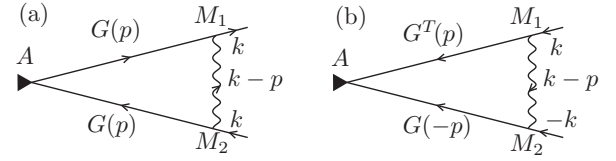

FIG. 7. Generic one-loop vertex correction in spinor space: (a) type-1 vertex (b) type-2 vertex.

where A and B are 2×2 Hermitian matrices: A acts in spinor space, while B acts in spin space. λ is a real-valued scalar, which can be viewed as an external source field coupling to the respective order parameter.

Restricting ourselves to instabilities where the bare vertex is momentum independent, a general vertex can be written as sum of such terms. As seen explicitly below, the quantum corrections do not mix at one-loop level, so it suffices to study the terms individually. We aim to classify the quantum corrections V to these operators at one-loop level. The corresponding diagrams are shown in Fig. 7.

In leading order in ϵ, these diagrams renormalize λ as $\lambda \to \lambda(1 + u_{\lambda} g^{4/3}/\epsilon)$; for $u_{\lambda} > 0$ (< 0), the instabilities are enhanced (suppressed). In RG formulation, the associated beta functions fulfill

$$\beta_{\lambda} = \frac{d\lambda}{d \ln \mu} = \lambda(-1 - \eta_{\lambda}) \quad (47)$$

with anomalous dimension η_{λ}. Proceeding as in the previous section, we find

$$\eta_{\lambda} = \frac{2}{3} u_{\lambda} g^{4/3} = \bigg|_{g = g^{\star}} - \frac{u_{\lambda}}{u_g} \epsilon. \quad (48)$$

To compute one-loop corrections V to the fermion bilinears of Eqs. (45) and (46), as a basis for the matrices A and B we choose $\mathbb{1}, \sigma_x, \sigma_y, \sigma_z$. The calculations are then fairly straightforward; technical details are presented in Appendix C. Let us sketch the results, starting with type-1 competing orders. For $A = \mathbb{1}$ or $A = \sigma_z$, the ϵ-divergent vertex corrections are proportional to

$$V \propto \int_{p_{d+1}} \vec{\Gamma} \cdot \vec{p} \cdot f(|\vec{p}|), \quad (49)$$

where $\vec{\Gamma}$ is the vector of Gamma matrices for the extra inserted dimensions (i.e., this vector has one entry in $d = 3$), and f is some function. In $d = 2$, there are no extra dimensions, and (49) vanishes trivially. Indeed, type-1 corrections with diagonal spinor matrices A correspond to superconducting instabilities; for these, the one-loop vertex correction is trivially absent as the diagram simply cannot be drawn. In higher dimensions, Eq. (49) also vanishes by antisymmetry. In particular, the FFLO boson-fermion vertex correction vanishes as already stated in Sec. IV. Thus superconducting vertices are not modified at the critical point at one-loop level. Of course, for pairing vertices, one should also take into account momentum dependent form factors, but these should only render the vertex less RG-relevant.

For $A = \sigma_x$, the corrections V are shown to vanish as well, similar to the vertex corrections in the Ising-nematic case [11]. Finally, for $A = \sigma_y$, the corrections vanish for

TABLE II. Type-1 instabilities, of the form $\bar{\Psi}_\alpha \sigma_y \Psi_\beta B_{\alpha\beta}$, modified by the FFLO at one loop.

Spin matrix B	Terms in action	Anomalous dimension η_λ
$\mathbb{1}$	$\bar{\psi}_\alpha^R \psi_\alpha^R - \bar{\psi}_\alpha^L \psi_{1\alpha}^L$	1.00ϵ, enhanced
σ_x	$\bar{\psi}_\uparrow^R \psi_\downarrow^R + \bar{\psi}_\downarrow^R \psi_\uparrow^R - \bar{\psi}_\uparrow^L \psi_\downarrow^L - \bar{\psi}_\downarrow^L \psi_\uparrow^L$	-1.00ϵ, suppressed
σ_y	$i(\bar{\psi}_\downarrow^L \psi_\uparrow^L + \bar{\psi}_\downarrow^R \psi_\uparrow^R) - i(\bar{\psi}_\uparrow^R \psi_\downarrow^R + \bar{\psi}_\uparrow^L \psi_\downarrow^L)$	1.00ϵ, enhanced
σ_z	$\bar{\psi}_\uparrow^R \psi_\uparrow^R - \bar{\psi}_\downarrow^R \psi_\downarrow^R - (\bar{\psi}_\uparrow^L \psi_\uparrow^L - \bar{\psi}_\downarrow^L \psi_\downarrow^L)$	-1.00ϵ, suppressed

$d = 2$ only (by Cauchy's integral theorem). Near $d = 5/2$, there are nonzero contributions; these lead to enhancement or suppression depending on the spin matrix B. Writing out the vertex (45) in terms of ordinary fermions ψ, the results are summarized in Table II.

Thus the type-1 vertices influenced by FFLO fluctuations correspond to density interactions between fermions with the same sheet index, with relative phases locked in various ways. Let us go over to type-2 competing orders, as these are easier to interpret and quantitatively more important. In particular, they also pick up sizable corrections for $d = 2$. For spinor matrices $A = \mathbb{1}, \sigma_z$, the quantum corrections vanish analogously to Eq. (49). For $A = \sigma_x, \sigma_y$, nontrivial corrections can arise. Evaluating all combinations $\Psi_\alpha^T(k) A \Psi_\beta(-k) B_{\alpha\beta}$ is again straightforward and shown in Appendix C; some combinations of A and B vanish trivially due to anticommutation of fermion fields. The results are summarized in Table III.

As indicated in Table III, competing orders that acquire a nontrivial one-loop correction from FFLO order correspond to the spin density-wave (SDW) or charge density-wave (CDW) channel. Only the latter order, with a wave vector peaked at $2k_{F,\downarrow}$ or $2k_{F,\uparrow}$, is enhanced. Note that this order, which is referred to as $2k_F$ scattering in Ref. [11], is suppressed in the Ising-nematic case; the change in sign can be cross-checked by integrating out bosons and noting that the resulting effective four-fermion interaction has an opposite sign when decoupled in the $2k_F$ channel in the Ising-nematic case compared to the FFLO case. In summary, our analysis of instabilities identifies the $2k_F$ CDW as the only serious competitor for FFLO criticality in $d = 2$.

Of course, this DIMREG computation can only predict how a tendency to order is enhanced, but not if there is an instability in the first place. A first indication that CDW order may indeed be important here can be obtained by straightforward evaluation of the corresponding vertex diagram with both fermions and bosons dressed by FFLO self-energies, which indeed shows a logarithmic divergence. To unambiguously answer the question which ordering tendency (FFLO or CDW) is more important, one would need to perform an RG analysis of an action, which treats both orders on the same footing, e.g., similar to Ref. [33]; we leave this task for future work.

VIII. CONCLUSION AND OUTLOOK

In this work, we have analyzed the quantum critical point between an FFLO superconductor and a normal metal phase in an anisotropic 2D system. Computing critical properties in a controlled expansion in $\epsilon = 5/2 - d$ dimensions, we have found a non-Fermi-liquid fixed point, characterized by a dynamical critical exponent $z = 1 + \epsilon$ and a correlation length exponent $\nu = 1 + \mathcal{O}(\epsilon^2)$ to leading order in ε. We derived the scaling forms of electronic and order-parameter correlations, and discussed possible physical manifestations.

One big advantage of the FFLO critical point compared to other non-Fermi liquid systems is that the scaling regime of the QCP is potentially accessible down to arbitrary low temperatures, if the quantum critical point is not masked by a competing order, such as superconductivity in heavy Fermion compounds or cuprate superconductors. In order to shed some light on this question, we also performed a general analysis of competing instabilities and found that charge density wave ordering is enhanced in the vicinity of the FFLO critical point. It is thus possible that the FFLO QCP is masked by a CDW phase in certain materials, depending on microscopic details. Extending our RG analysis to a situation where FFLO and CDW fluctuations are treated on equal footing would be an interesting problem for future study. In a similar spirit, one could attempt an RG analysis of disorder [68], which is known to destroy the FFLO state in organic superconductors [69].

Our analytical derivation relies heavily on the approximation that the spin-up and spin-down Fermi surface branches have the same curvature, respectively, vanishing velocity detuning $\delta v \to 0$. While this parameter choice is physically grounded, treating the $\delta v \neq 0$ case, e.g., numerically would be very interesting, potentially revealing a modification of the Fermi surface shape as in the SDW case [8]. In addition, one could try to start from the opposite limit $\delta v \to \infty$. A higher loop analysis of the problem would be desirable as well, but appears rather involved; alternatively, for $\delta v \neq 0$, one could

TABLE III. Type-2 instabilities, of the form $\bar{\Psi}_\alpha^T A \Psi_\beta B_{\alpha\beta} + $ H.c., modified by the FFLO at one loop.

Spinor A	Spin B	Terms in action	Anomalous dimension η_λ
σ_x	σ_y	$2i(\bar{\psi}_\downarrow^L \psi_\uparrow^R - \bar{\psi}_\uparrow^L \psi_\downarrow^R) + $ H.c.: SDW in y direction	-1.70ϵ, suppressed
σ_y	$\mathbb{1}$	$2i\bar{\psi}_\alpha^L \psi_\alpha^R + $ H.c.: CDW at $2k_{F,\downarrow}$ or $2k_{F,\uparrow}$	2.69ϵ, enhanced
σ_y	σ_x	$2i(\bar{\psi}_\downarrow^L \psi_\uparrow^R + \bar{\psi}_\uparrow^L \psi_{2\downarrow}^R) + $ Hc.: SDW in x direction	-2.69ϵ, suppressed
σ_y	σ_z	$2i(\bar{\psi}_\uparrow^L \psi_\uparrow^R - \bar{\psi}_\downarrow^L \psi_\downarrow^R) + $ H.c.: SDW in z direction	-2.69ϵ, suppressed

apply the scheme with fixed co-dimension as shortly discussed, and see if it leads to similar results.

APPENDIX A: MEAN-FIELD PHASE TRANSITION OF A MICROSCOPIC MODEL

To illustrate our field theoretic starting point, in this Appendix we recall the ordinary Ginzburg-Landau picture of the phase transition. Paraphrasing the treatment of Ref. [14], we start from a microscopic model appropriate, e.g., for the Bechgaard salt $(TMTSF)_2ClO_2$ [18,70]: we consider spinful fermions freely moving along chains oriented in x direction, with a small interchain hopping parameter t. When these electrons are Zeeman-coupled to a magnetic field h, the free fermionic Hamiltonian reads

$$H_0 = \sum_{\alpha=\uparrow,\downarrow,\mathbf{k}} \xi_\alpha(\mathbf{k})\psi_\alpha^\dagger(\mathbf{k})\psi_\alpha(\mathbf{k}),$$

$$\xi_\alpha = k_x^2/2 - 2t\cos(k_y) - \mu - s_\alpha h, \quad \{s_\uparrow, s_\downarrow\} = \{1,-1\}, \tag{A1}$$

where μ is the chemical potential, and we set the fermion mass and interchain distance to 1. Plotting the Fermi surface with parameters $\mu = 3.3$, $t = 0.5$, and $h = 1.0$ readily reproduces Fig. 2.

We now assume that the electrons interact with some short-range attractive interaction hamiltonian H_{int} (e.g., mediated by phonons) as in Eq. (3). Then, we introduce a functional integral representation of $H = H_0 + H_{int}$, resulting in a quantum action S (see, e.g., Ref. [71]). Decoupling the interaction

term H_{int} in the pairing channel yields (we consider finite temperature T for generality)

$$S[\psi_\alpha, \bar{\psi}_\alpha, \Delta, \bar{\Delta}] = S_0[\psi_\alpha, \bar{\psi}_\alpha] + S_{int}[\psi_\alpha, \bar{\psi}_\alpha, \Delta, \bar{\Delta}],$$

$$S_{int} = \sum_{\omega_n, \mathbf{q}} g|\Delta(\omega_n, \mathbf{q})|^2 - \frac{g}{\sqrt{\beta V}} \sum_{\substack{\mathbf{k},\mathbf{q} \\ \nu_n, \omega_n}} \bar{\Delta}(\omega_n, \mathbf{q})$$

$$\times \psi_\downarrow(\omega_n - \nu_n, \mathbf{q} - \mathbf{k})\psi_\uparrow(\nu_n, \mathbf{k}) + \text{H.c.}, \tag{A2}$$

where S_0 is the bare fermionic action derived from Eq. (A1), ω_n, ν_n are bosonic and fermionic Matsubara frequencies, respectively, and β is the inverse temperature. The subsequent mean-field analysis shows that the superconducting susceptibility is peaked at momenta $\pm\mathbf{Q}_{FFLO} = (k_{F,\uparrow} - k_{F,\downarrow})\mathbf{e}_x$, where $k_{F,\alpha}$ are the respective Fermi momenta of the two spin species. Consequently, electrons interact with superconducting fluctuations Δ predominantly at so-called hot spots on the Fermi surface which are connected by $\mathbf{Q}_{FFLO}$, found at $k_y = 0, k_x = \pm k_{F,\alpha}$. For this reason, within a low-energy theory sufficient for a universal RG analysis, we can expand the fermion fields as well as the fermion dispersions near these hot spots. In this manner, we introduce four low-energy fields $\psi_{\uparrow/\downarrow}^{L/R}$. Furthermore, expanding Δ near $\pm\mathbf{Q}_{FFLO}$ readily yields action (4) in the limit $V \to \infty, T \to 0$ apart from different boson kinetic and mass terms, which automatically arise in the RG flow as discussed in the main text.

A standard Landau-Ginzburg analysis of Eq. (A2), which indicates a continuous phase transition, can be performed by integrating out the fermions.[3] This yields an effective bosonic action

$$S_\Delta[\Delta, \bar{\Delta}] = \sum_{\omega_n, \mathbf{q}} g|\Delta(\omega_n, \mathbf{q})|^2 - \text{Tr} \ln G^{-1}, \tag{A3}$$

where Tr denotes the trace in spin and energy-momentum space, and G^{-1} is a matrix propagator:

$$G^{-1}(\nu_n, \nu_n', \mathbf{k}, \mathbf{k}') = \begin{pmatrix} \beta\delta_{\nu_n,\nu_n'}\delta_{\mathbf{k},\mathbf{k}'}(i\nu_n - \xi_\uparrow(\mathbf{k})) & \sqrt{\beta/V}\,g\Delta(\nu_n - \nu_n', \mathbf{k} - \mathbf{k}') \\ \sqrt{\beta/V}\,g\bar{\Delta}(\nu_n - \nu_n', \mathbf{k} - \mathbf{k}') & \beta\delta_{\nu_n,\nu_n'}\delta_{\mathbf{k},\mathbf{k}'}(i\nu_n + \xi_\downarrow(-\mathbf{k})) \end{pmatrix}. \tag{A4}$$

To generally treat Eq. (A3) on mean-field level, one would proceed by solving for the saddle point, $\delta/(\delta\Delta)S_\Delta \overset{!}{=} 0$, making an appropriate mean-field ansatz for the (static) boson. The Larkin-Ovchinnikov ansatz, around which our dynamical boson in the main part is expanded, reads

$$\Delta_{LO}(\omega_n, \mathbf{q}) = \Delta_0\delta_{\omega_n,0}(\delta_{\mathbf{q},\mathbf{Q}_{FFLO}} + \delta_{\mathbf{q},-\mathbf{Q}_{FFLO}}), \tag{A5}$$

where the amplitude Δ_0 can be chosen real. However, a derivation of a closed-form saddle point equation ($\hat{=}$ mean-field self-consistency equation) is difficult since it requires the inversion of Eq. (A4), which is hindered by the involved momentum dependence in Eq. (A5). To avoid this difficulty, one can plug

in the ansatz (A5) into S_Δ and expand in powers of Δ_0 up to fourth order. Since the odd terms trivially vanish by symmetry, one obtains an effective Landau-Ginzburg functional

$$S_{LG}[\Delta_0] = m[h]\Delta_0^2 + a_4[h]\Delta_0^4, \tag{A6}$$

where we have indicated the magnetic field dependence explicitly. A strong indication for a continuous transition at mean-field level is then given if (see, e.g., Ref. [73]) the boson mass m can be tuned to zero for appropriate h, while $a_4 > 0$. The second condition was shown to be true in Ref. [14] (see Appendix A within). Let us focus on the first one here. As easily shown, the coefficient m is given by

$$m = 2g^2(1/g - \Pi_0[h]), \tag{A7}$$

$$\Pi_0[h] = \sum_{\mathbf{k}} \frac{1 - n_F[\xi_\uparrow(\mathbf{k})] - n_F[\xi_\downarrow(\mathbf{Q}_{FFLO} - \mathbf{k})]}{\xi_\uparrow(\mathbf{k}) + \xi_\downarrow(\mathbf{Q}_{FFLO} - \mathbf{k}))}, \tag{A8}$$

---

[3]This is dangerous for 2D fermionic systems, see, e.g., Ch. 18 of Ref. [72]; a proper analysis requires an RG procedure as presented in this paper.
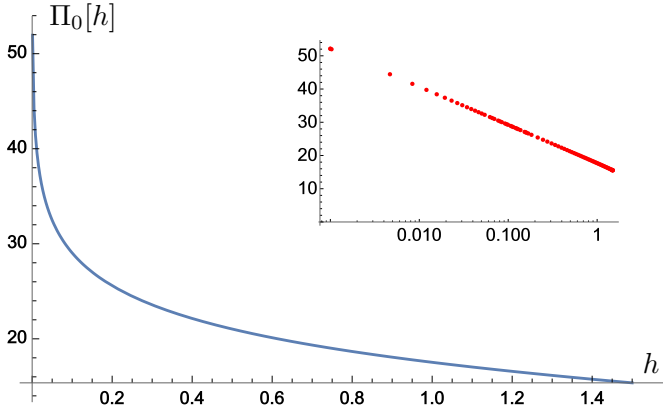
FIG. 8. $\Pi_0[h]$ numerically computed from Eq. (A8), with parameters $\mu = 3.3, t = 0.5$. The inset shows the same plot on a log-linear scale.

where n_F is the Fermi distribution, and Π_0 the static inverse pair propagator, respectively, the boson self-energy. Evaluating Eq. (A8) for general external boson momenta, one can easily check that it is indeed peaked at $\mathbf{Q}_{\text{FFLO}}$ as claimed before. We limit ourselves to a numerical evaluation of $\Pi_0[h]$ in the limit $T \to 0$; a plot for generic parameters is shown in Fig. 8.

As clearly seen in Fig. 8, $\Pi_0[h]$ diverges as $h \to 0$. In fact, this divergence is logarithmic, as pinpointed in the inset. This is in accordance with the analytical evaluation for the low-energy action in Appendix D (where $\delta v \propto h$), and also with Ref. [15]. Therefore, at any arbitrarily small value of the coupling g, there is a critical magnetic field $h_c \propto \exp(-1/g)$ where the mass term m in Eq. (A7) changes sign, and the mean-field phase transition between the normal metal and the FFLO superconductor occurs. Close to h_c, the field dependence of the mass term scales as $m \propto (h - h_c)/h_c$, as claimed in the main text.

The mean-field treatment presented above is fairly simplistic. First, it does not describe the phase transition between the FFLO and homogeneous superconductor—to this aim, one would have to make a homogeneous mean-field ansatz as well, which we avoid since we are only interested in the QCP shown in Fig. 1. One could also improve the mean-field ansatz, say, by allowing for more complicated periodic functions than the $\cos(\mathbf{Q} \cdot \mathbf{x})$ LO dependence, as done, e.g., in Ref. [52]. We do not pursue this further since the mean-field treatment is not the focus of this work, and the general outcome that a mean-field transition exists and is continuous in 2D is generally agreed upon in the literature.

APPENDIX B: COMPUTATION OF SELF-ENERGIES

1. Boson self-energy

Here, we present the evaluation of the boson self-energy, given by Eq. (23),

$$\Pi(k) = -g^2\mu^\epsilon \int_{p^{d+1}} \sum_{\alpha \neq \alpha'} \text{Tr}[G_\alpha(-p)M_1 G_{\alpha'}(k-p)M_2].$$

(B1)

To evaluate the trace, we use

$$\text{Tr}[\sigma_i M_1 \sigma_j M_2] = \begin{array}{c} \\ i=x \\ i=y \\ i=z \end{array} \begin{array}{ccc} j=x & j=y & j=z \\ \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}. \quad \text{(B2)}$$

In the limit $\delta v \to 0$ discussed in the main text, this leads to

$$\Pi(k) = 2g^2\mu^\epsilon \int_{p^{d+1}} \frac{(\delta_{-p} + ip_0)(\delta_{k-p} + i(k_0 - p_0))}{[\mathbf{P}^2 + (\delta_{-p})^2][(\mathbf{K} - \mathbf{P})^2 + (\delta_{k-p})^2]}.$$

(B3)

Changing to energy variables $x = \delta_{-p}, y = \delta_{k-p}$, with Jacobian $1/|2k_d|$, $\Pi(k)$ is rewritten as

$$\Pi(k) = \frac{g^2\mu^\epsilon}{|k_d|} \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{d-2}} \frac{dx}{2\pi} \frac{dy}{2\pi} \frac{x + ip_0}{\mathbf{P}^2 + x^2}$$
$$\times \frac{y + i(k_0 - p_0)}{(\mathbf{K} - \mathbf{P})^2 + y^2}, \quad \mathbf{P} = (p_0, \vec{p}). \quad \text{(B4)}$$

Note that the limit $\delta v \to 0$ is already required at this stage: for general velocity detuning, the Jacobian of the transformation to energy variables is more involved, and the integration range is nontrivial as well, obstructing further evaluation.

Taking the elementary x, y integrals (note that the log-divergent parts vanish by antisymmetry), results in

$$\Pi(k) = \frac{g^2\mu^\epsilon}{4|k_d|} \int \frac{dp_0}{2\pi} \int \frac{d\vec{p}}{(2\pi)^{d-2}} \frac{(k_0 - p_0)p_0}{|\mathbf{P}||\mathbf{K} - \mathbf{P}|}. \quad \text{(B5)}$$

To proceed (the remaining steps are similar to Sec. A1 of Ref. [11]), we introduce a Feynman parameter, using

$$\frac{1}{\sqrt{A_1}\sqrt{A_2}} = \frac{1}{\pi} \int_0^1 dt \frac{1}{\sqrt{t(1-t)}} \frac{1}{tA_1 + (1-t)A_2}. \quad \text{(B6)}$$

Shifting $\mathbf{P} \to \mathbf{P} + (1-t)\mathbf{K}$, this gives

$$\Pi(k) = \frac{g^2\mu^\epsilon}{4\pi|k_d|} \int \frac{dp_0}{2\pi} \int \frac{d\vec{p}}{(2\pi)^{d-2}} \int_0^1 dt \frac{1}{\sqrt{t(1-t)}}$$
$$\times \frac{(tk_0 - p_0)(p_0 + (1-t)k_0)}{\mathbf{P}^2 + t(1-t)\mathbf{K}^2}. \quad \text{(B7)}$$

We note that the terms of the numerator linear in p_0 give no contribution by antisymmetry. After rescaling $\mathbf{P} \to \sqrt{t(1-t)}\mathbf{P}$, we are left with a t integral of the form

$$\int_0^1 dt(t(1-t))^{d/2-1} = \frac{\Gamma\left(\frac{d}{2}\right)^2}{\Gamma(d)}. \quad \text{(B8)}$$

Going to polar coordinates, the remaining integrals read

$$\Pi(k)$$
$$= \frac{g^2\mu^\epsilon}{|k_d|} \frac{\Gamma\left(\frac{d}{2}\right)^2}{\Gamma(d)} \cdot \frac{2^{1-d}}{\pi^{d/2}\Gamma\left(\frac{d}{2}-1\right)} \int_{-\infty}^{\infty} \frac{dp_0}{2\pi}$$
$$\times \int_0^\infty d|\vec{p}| \, |\vec{p}|^{d-3} \left(\frac{p_0^2 - k_0^2}{|\vec{p}|^2 + p_0^2 + |\vec{k}|^2 + k_0^2} - \frac{p_0^2}{|\vec{p}|^2 + p_0^2}\right). \quad \text{(B9)}$$

Here, we have also subtracted $\Pi(0, k_d)$ for UV regularization. As discussed in the main text, the residual momentum dependence of this subtraction can be seen as an artefact of

the $\delta v \to 0$ limit, and is further discussed in Appendix D. Formally, this subtraction can also be justified by referring to Veltman's formula (see, e.g., Ref. [46]).

It is instructive to study the $|\vec{p}|$ integral as $d \to 2$. In this limit, the extra dimensions vanish and the $|\vec{p}|$ integral should be absent. Indeed, as $d \to 2$, the integral becomes IR logarithm-divergent, and so comes from $\vec{p} = 0$ only; the logarithmic divergence is asymptotically canceled by the prefactor $\sim 1/\Gamma(d/2 - 1)$. The remaining integrations are straightforward, resulting in Eq. (25) of the main text:

$$\Pi(k) = \chi_d \frac{g^2 \mu^\epsilon}{|k_d|} \left(d\, k_0^2 + |\vec{k}|^2\right) \left(k_0^2 + |\vec{k}|^2\right)^{\frac{d-3}{2}},$$

$$\chi_d = \frac{\Gamma((1-d)/2)}{2^{d+2}\pi^{(d+1)/2}} \frac{\Gamma(d/2)^2}{\Gamma(d)},$$

$$\chi_{5/2} \simeq -0.0178. \tag{B10}$$

2. Fermion self-energy

We continue with evaluation of the fermion self-energy with external spin index κ, starting from Eq. (27):

$$\Sigma_1^\kappa(k) = g^2 \mu^\epsilon \int_{p^{d+1}} D(p) M_1 G_\beta(k+p) M_2 \sigma_y^{\kappa\beta} \sigma_y^{\beta\kappa}. \tag{B11}$$

The sums in spinor space can be performed using

$$M_1 \sigma_i M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{cases} 1 & i = x \\ -i & i = y \\ 0 & i = z \end{cases}. \tag{B12}$$

In the spin-independent limit $\delta v \to 0$, this leads to

$$\Sigma_1 = \underbrace{\begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix} g^2}_{\equiv c_1} \mu^\epsilon \int_{p^{d+1}} \frac{(k_0 + p_0)\, i + \delta_{k+p}}{(\boldsymbol{K} + \boldsymbol{P})^2 + (\delta_{k+p})^2} \frac{1}{p_d^2 - \Pi(p)}. \tag{B13}$$

Inserting the boson self-energy, one can elementarily evaluate the p_{d-1}, p_d integrals, resulting in

$$\Sigma_1 = \frac{ic_1 \mu^\epsilon}{3\sqrt{3}} \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{d-2}} \frac{(k_0 + p_0)}{|\boldsymbol{K} + \boldsymbol{P}|}$$

$$\times \frac{1}{\mu^{\epsilon/3} \chi^{1/3} \left(d\mathbf{c}, p_0^2 + |\vec{p}|^2\right)^{1/3} \left(p_0^2 + |\vec{p}|^2\right)^{\frac{d-3}{6}}},$$

$$\chi = -\chi_d g^2. \tag{B14}$$

We apply a Feynman parametrization:

$$\frac{1}{A_1^{\alpha_1} A_2^{\alpha_2} A_3^{\alpha_3}} = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^1 dt_1 \int_0^{1-t_1} dt_2$$

$$\times \frac{t_1^{\alpha_1 - 1} t_2^{\alpha_2 - 1} (1 - t_1 - t_2)^{\alpha_3 - 1}}{(t_1 A_1 + t_2 A_2 + (1 - t_1 - t_2) A_3)^{\alpha_1 + \alpha_2 + \alpha_3}}. \tag{B15}$$

With $\alpha_1 = 1/2, \alpha_2 = 1/3, \alpha_3 = (d - 3)/6$, Eq. (B14) is rewritten as

$$\Sigma_1 = c_2 \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{d-2}} \int_0^1 dt_1 \int_0^{1-t_1} dt_2 \frac{(k_0 + p_0) t_1^{-\frac{1}{2}} t_2^{-\frac{2}{3}} (1 - t_1 - t_2)^{\frac{d-9}{6}}}{\left[t_1 (\boldsymbol{K} + \boldsymbol{P})^2 + t_2 \left(dp_0^2 + |\vec{p}|^2\right) + (1 - t_1 - t_2)\left(p_0^2 + |\vec{p}|^2\right)\right]^{\frac{d+2}{6}}},$$

$$c_2 = \frac{ic_1 \mu^{2/3\epsilon} \Gamma\left(\frac{2+d}{6}\right)}{3\sqrt{3}\, \chi^{1/3} \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{3}\right) \Gamma\left(\frac{d-3}{6}\right)}. \tag{B16}$$

Strictly speaking, the Feynman parametrization of Eq. (B16) is only well-defined for $d > 3$, as the t_2 integral is otherwise divergent. We will circumvent this problem by evaluating the t_2 integral for general $d > 3$ below (after the momentum integrals), and then analytically continue the result to $d < 3$; the divergence at $d = 3$ will cancel against the term $\Gamma((d-3)/6)$ contained in the factor c_2. As there certainly is a strip of convergence of the original integral (B14), and we also recover the $d = 2$ result of Ref. [14], this procedure should be legitimate. To proceed, in Eq. (B16), we shift

$$p_0 \to p_0 + \frac{-t_1}{1 + t_2(d-1)} k_0, \quad \vec{p} \to \vec{p} - t_1 \vec{k}. \tag{B17}$$

Disregarding the linear terms in p_0, which vanish by antisymmetry, we then obtain

$$\Sigma_1 = c_2 \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{d-2}} \int_0^1 dt_1 \int_0^{1-t_1} dt_2 \frac{\left(1 + \frac{-t_1}{1+t_2(d-1)}\right) k_0\, t_1^{-\frac{1}{2}} t_2^{-\frac{2}{3}} (1 - t_1 - t_2)^{\frac{d-9}{6}}}{\left(d_1 k_0^2 + d_2 p_0^2 + |\vec{p}|^2 + |\vec{k}|^2 t_1 (1 - t_1)\right)^{\frac{d+2}{6}}},$$

$$d_1 = t_1 - \frac{t_1^2}{1 + (d-1)t_2}, \quad d_2 = 1 + (d-1)t_2. \tag{B18}$$

For $2 \leqslant d < 5/2$, the momentum integrals can be straightforwardly evaluated by going to polar coordinates, yielding

$$\Sigma_1 = k_0 \frac{c_2\, 2^{1-d}\, \Gamma\left(\frac{4}{3} - \frac{d}{3}\right) \Gamma\left(\frac{\epsilon}{3}\right)}{\pi^{(d-1)/2} \Gamma\left(\frac{2+d}{6}\right) \Gamma\left(\frac{1}{2} + \frac{\epsilon}{3}\right)} \int_0^1 dt_1 \int_0^{1-t_1} dt_2 \frac{\left(1 - \frac{t_1}{1+(d-1)t_2}\right) t_1^{-\frac{1}{2}} t_2^{-\frac{2}{3}} (1 - t_1 - t_2)^{\frac{d-9}{6}}}{\sqrt{1 + (d-1)t_2}} \left(\frac{\mu^2}{d_1 k_0^2 + t_1(1 - t_1)|\vec{k}|^2}\right)^{\epsilon/3}. \tag{B19}$$

Following the procedure described below Eq. (B16), let us evaluate the t_2 integral for $d > 3$. With an eye for the final limit $\epsilon \to 0$, we still set the last dimensionless prefactor in Eq. (B19) equal to one, which should be fine as this is a perfectly regular function in t_1 and t_2. We have also checked this numerically on a simplified integral. Furthermore, note that in $d = 2$, we can extract a factor of $|k_0|^{-1/3}$ from the integral, and obtain a self-energy $\propto |k_0|^{2/3}$ as found in Ref. [14].

Evaluation of the t_2 integral yields, without the other prefactors, the fairly involved expression

$$\tilde{F}(d) = -\Gamma\left(\frac{d-3}{6}\right) \times (1-t_1)^{\frac{d-7}{6}} \Gamma\left(\frac{1}{3}\right) \left\{\left[\left(d\left(t_1^2 - 3t_1 + 3\right) - t_1(t_1 + 3)\right) {}_2F_1\left(\frac{1}{3}, \frac{1}{2}; \frac{d-1}{6}; (d-1)(t_1 - 1)\right)\right.\right.$$
$$\left.\left. + (4-d)t_1 \, {}_2F_1\left(-\frac{1}{2}, \frac{1}{3}; \frac{d-1}{6}; (d-1)(t_1 - 1)\right)\right]\right\} \Bigg/ \left[3\sqrt{t_1}(d(t_1 - 1) - t_1)\Gamma\left(\frac{d-1}{6}\right)\right], \quad \text{(B20)}$$

where ${}_2F_1$ is the hypergeometric function. $\tilde{F}(d)$ is divergent for $d \searrow 3$ due to the prefactor $\Gamma((d-3)/6)$, but this factor cancels against the same factor contained in the overall prefactor c_2 [cf. (B16)]. The remainder $F(d) \equiv \tilde{F}(d)/\Gamma((d-3)/6)$ is a well-behaved function. Its numerical integration leads to

$$\int_0^1 dt_1 F\left(\frac{5}{2}\right) \simeq 1.166. \quad \text{(B21)}$$

Collecting all prefactors, and expanding the Gamma functions from Eq. (B19) in ϵ, one obtains

$$\Sigma_1 = \frac{u_g g^{4/3}}{\epsilon}\begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}(-ik_0) + \text{finite terms},$$
$$u_g = -0.0813. \quad \text{(B22)}$$

Evaluation of Σ_2 given in Eq. (28) proceeds analogously. In total, one arrives at Eq. (29) of the main text:

$$\Sigma(k) = \Sigma_1(k) + \Sigma_2(k) \quad \text{(B23)}$$
$$= \frac{u_g g^{4/3}}{\epsilon}\sigma_y(-ik_0) + \text{finite terms}.$$

APPENDIX C: COMPUTATION OF VERTEX CORRECTIONS FOR COMPETING INSTABILITIES

In this Appendix, we compute the anomalous dimensions of possible competing orders, which are summarized in Tables II and III.

1. Type-1 orders

As in the main text, we start with type-1 orders, computing one-loop corrections V to the fermion bilinear of Eq. (45). Fixing the signs with Wick's theorem, in the limit $\delta v \to 0$, where the Green's functions become spin-independent, they have the general form

$$V = \lambda \int_{k^{d+1}} \bar{\Psi}_\gamma(k)\Omega(k)\Psi_\delta(k)\left(\sigma_y^{\beta\delta}\sigma_y^{\gamma\alpha}B_{\alpha\beta}\right), \quad \text{(C1)}$$

$$\Omega(k) = \Omega^1(k) + \Omega^2(k),$$

$$\Omega^1(k) = g^2\mu^\epsilon \int_{p^{d+1}} M_1 G(p)AG(p)M_2 D(k-p),$$

$$\Omega^2(k) = \Omega^1[M_1 \leftrightarrow M_2]. \quad \text{(C2)}$$

Let us fix $A = \mathbb{1}$ and compute Ω^1. The sums in spinor space are determined from

$$M_1\sigma_i\mathbb{1}\sigma_j M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{matrix} i=x \\ i=y \\ i=z \end{matrix}\overset{\begin{matrix} j=x & j=y & j=z \end{matrix}}{\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & i \\ 1 & -i & 0 \end{pmatrix}}. \quad \text{(C3)}$$

Since $G \propto -\mathbf{\Gamma}\cdot\mathbf{P} + \sigma_x\delta_p$ and we take the Gamma matrices $\vec{\Gamma}$ in the extra dimensions to be proportional to σ_z [cf. Eq. (15)], it immediately follows that Ω^1 is of the form

$$\Omega^1 \propto \int_{p_{d+1}} (\vec{\Gamma}\cdot\vec{p})\, f(|\vec{p}|), \quad \text{(C4)}$$

where f is some function. This expression vanishes as discussed in the main text below Eq. (49). The same conclusion holds for $A = \sigma_z$. For $A = \sigma_x$, using

$$M_1\sigma_i\sigma_x\sigma_j M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{matrix} i=x \\ i=y \\ i=z \end{matrix}\overset{\begin{matrix} j=x & j=y & j=z \end{matrix}}{\begin{pmatrix} 1 & -i & 0 \\ -i & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}, \quad \text{(C5)}$$

we obtain

$$\Omega^1 = -g^2\mu^\epsilon\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\int_{p^{d+1}}\frac{-\mathbf{P}^2 + \delta_p^2 + 2ip_0\delta_p}{\left(\delta_p^2 + \mathbf{P}^2\right)^2}D(k-p). \quad \text{(C6)}$$

This expression has the same form as the vertex correction in the Ising-nematic case [11]. Since the boson propagator D is independent of p_{d-1}, after shifting $p_{d-1} \to \delta_p$, Eq. (C6) vanishes due to the identity

$$\int dx\, \frac{x^2 - a^2}{(x^2 + a^2)^2} = 0. \quad \text{(C7)}$$

Last, we consider $A = \sigma_y$. Using

$$M_1\sigma_i\sigma_y\sigma_j M_2 = \begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}\begin{matrix} i=x \\ i=y \\ i=z \end{matrix}\overset{\begin{matrix} j=x & j=y & j=z \end{matrix}}{\begin{pmatrix} -1 & i & 0 \\ i & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}, \quad \text{(C8)}$$

we find

$$\Omega^1 = -\underbrace{g^2\begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}}_{\equiv c_1}\mu^\epsilon \int_{p^{d+1}} \frac{p_0^2 - p^2 - \delta_p^2 - 2i\delta_p p_0}{\left(\boldsymbol{P}^2 + \delta_p^2\right)^2} D(k-p). \tag{C9}$$

Performing the p_{d-1} integral (by shifting $p_{d-1} \to \delta_p$), we get

$$\Omega = c_1\mu^\epsilon \int \frac{dp_0}{2\pi}\frac{d\vec{p}}{(2\pi)^{d-2}}\frac{dp_d}{2\pi}\frac{|\vec{p}|^2}{2|\boldsymbol{P}|^3}D(k-p). \tag{C10}$$

Note that for $d=2$, Eq. (C9) vanishes by Cauchy's integral theorem, which can be seen by reducing the fraction; accordingly, the integrand in Eq. (C10) is proportional to the external momenta $|\vec{p}|^2$. To further evaluate Eq. (C10), we focus on $k=(k_0,0,\ldots)$, which is sufficient in leading order in ϵ. Shifting $p_0 \to p_0 + k_0$ for convenience and performing the p_d integral gives

$$\Omega^1 = \frac{c_1\mu^{2\epsilon/3}}{3\sqrt{3}\chi^{1/3}}\int\frac{dp_0}{2\pi}\frac{d\vec{p}}{(2\pi)^{d-2}}\frac{p^2}{((p_0+k_0)^2+p^2)^{3/2}\left(dp_0^2+p^2\right)^{1/3}\left(p_0^2+p^2\right)^{(d-3)/6}}, \tag{C11}$$

with $\chi \simeq 0.0178g^2$. Applying the Feynman parametrization (B15), with $\alpha_1 = 3/2, \alpha_2 = 1/3, \alpha_3 = (d-3)/6$, we obtain

$$\Omega^1_{\alpha\beta} = \underbrace{\frac{c_1}{3\sqrt{3}\chi^{1/3}}\frac{\Gamma\left(\frac{d+8}{6}\right)}{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{d-3}{6}\right)}}_{\equiv c_2}\mu^{2\epsilon/3}\int\frac{dp_0}{2\pi}\frac{d\vec{p}}{(2\pi)^{d-2}}\int_0^1 dt_1\int_0^{1-t_1}dt_2\,t_1^{1/2}t_2^{-2/3}(1-t_1-t_2)^{\frac{d-9}{6}}$$

$$\times\frac{p^2}{\left[(p_0+k_0)^2+p^2)t_1 + \left(dp_0^2+p^2\right)t_2 + \left(p_0^2+p^2\right)(1-t_1-t_2)\right]^{\frac{d+8}{6}}}, \tag{C12}$$

where we follow the same logic as in the evaluation of Eq. (B16). Shifting $p_0 \to p_0 - \frac{t_1}{(d-1)t_2+1}k_0$ and going to polar coordinates yields

$$\Omega^1 = \frac{c_2\mu^\epsilon 2^{3-d}}{\pi^{d/2-1}\Gamma\left(\frac{d}{2}-1\right)}\int_{-\infty}^\infty\frac{dp_0}{2\pi}\int_0^\infty d|\vec{p}|\int_0^1 dt_1\int_0^{1-t_1}dt_2 t_1^{1/2}t_2^{-2/3}(1-t_1-t_2)^{\frac{d-9}{6}}\frac{|\vec{p}|^{d-1}}{\left(d_1k_0^2+d_2p_0^2+|\vec{p}|^2\right)^{\frac{d+8}{6}}}, \tag{C13}$$

where $d_{1/2}$ were defined in Eq. (B18). Performing the $|\vec{p}|$ and p_0 integrals is then straightforward and results in

$$\Omega^1 = \frac{c_2\Gamma\left(\frac{d}{2}\right)\Gamma\left(\frac{\epsilon}{3}\right)}{4\sqrt{\pi}\Gamma\left(\frac{d+8}{6}\right)}\int_0^1 dt_1\int_0^{1-t_1}dt_2$$

$$\times\frac{t_1^{1/2}t_2^{2/3}(1-t_1-t_2)^{\frac{d-9}{6}}}{\sqrt{d_2}\Gamma\left(\frac{d-3}{6}\right)}\left(\frac{\mu^2}{d_1k_0^2}\right)^{\epsilon/3}. \tag{C14}$$

Approximating the last expression in parentheses in Eq. (C14) by 1, the t_2 integral can be evaluated analytically for $d > 3$; the divergence as $d \to 3$ cancels against the factor $\Gamma[(d-3)/6]$ contained in c_2, cf. Eq. (C12). Then, the t_1 integral can be computed numerically for $d = 5/2$, yielding

$$\Omega^1 \simeq \begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}0.081\frac{g^{4/3}}{\epsilon}. \tag{C15}$$

Ω^2 [cf. Eq. (C1)] is evaluated in the same vein, and in total, we obtain

$$\Omega = \sigma_y 0.081\frac{g^{4/3}}{\epsilon}. \tag{C16}$$

Now, we need to evaluate the factor involving the spin matrix B in Eq. (C1), which yields

$$\left(\sigma_y^{\beta\delta}\sigma_y^{\gamma\alpha}B_{\alpha\beta}\right) = \begin{cases} B_{\gamma\delta}, & B = \mathbb{1}, \sigma_y, \\ -B_{\gamma\delta}, & B = \sigma_x, \sigma_z. \end{cases} \tag{C17}$$

Altogether, the quantum correction V therefore reads

$$V = \lambda\frac{u_\lambda g^{4/3}}{\epsilon}\int_{k^{d+1}}\bar{\Psi}_\gamma(k)\sigma_y\Psi_\delta(k)\times\begin{cases} B_{\gamma\delta}, & B = \mathbb{1}, \sigma_y, \\ -B_{\gamma\delta}, & B = \sigma_x, \sigma_z, \end{cases}$$

$$u_\lambda = 0.081. \tag{C18}$$

Using Eq. (48), this readily yields Table II.

2. Type-2 orders

We proceed with type-2 orders, computing corrections V to the fermion bilinear of Eq. (46). Analogous to the previous case, they are of the form

$$V = \lambda\int_{k^{d+1}}\Psi_\gamma^T(k)\Omega(k)\Psi_\delta(-k)\left(\sigma_y^{\beta\delta}\sigma_y^{\alpha\gamma}B_{\alpha\beta}\right) + \text{H.c.}, \tag{C19}$$

$$\Omega(k) = \Omega^1(k) + \Omega^2(k),$$

$$\Omega^1(k) = g^2\mu^\epsilon\int_{p^{d+1}}M_1 G^T(p)AG(-p)M_2 D(k-p),$$

$$\Omega^2(k) = \Omega^1[M_1 \leftrightarrow M_2]. \tag{C20}$$

For $A = \mathbb{1}, \sigma_z$, Ω^1 vanishes as in the previous case. For $A = \sigma_x$, the required product in spinor space reads

$$M_1 \sigma_i^T \sigma_x \sigma_j M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{matrix} i=x \\ i=y \\ i=z \end{matrix} \begin{matrix} j=x & j=y & j=z \\ \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{matrix}, \tag{C21}$$

resulting in

$$\Omega^1 = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{\equiv c_1} g^2 \mu^\epsilon \int_{p^{d+1}} \frac{p_0^2 - p^2 + ip_0\delta_{-p} + ip_0\delta_p - \delta_p\delta_{-p}}{(\boldsymbol{P}^2 + \delta_p^2)(\boldsymbol{P}^2 + \delta_{-p}^2)} D(k-p). \tag{C22}$$

To evaluate this expression, we restrict ourselves to $k = (0, \ldots, k_d)$. Then, the linear terms in p_0 vanish by antisymmetry. Taking the p_{d-1} integral results in

$$\Omega^1 = \frac{c_1 \mu^\epsilon}{2} \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{(d-2)}} \frac{dp_d}{2\pi} \frac{p_0^2}{|\boldsymbol{P}|(p_d^4 + \boldsymbol{P}^2)} \frac{1}{(k_d + p_d)^2 + \frac{\chi\mu^\epsilon}{|k_d+p_d|}(dp_0^2 + \vec{p}^2)(p_0^2 + \vec{p}^2)^{\frac{d-3}{2}}}, \tag{C23}$$

with $\chi \simeq 0.0178g^2$. To evaluate Eq. (C23), we shift $p_d \to p_d - k_d$. Then, following Ref. [11], we may approximately disregard the p_d dependence of the fermion part in leading order in g (and hence in leading order in ϵ). We can then perform the p_d integral, yielding

$$\Omega^1 = \left(\frac{\mu}{k_d^2}\right)^{2/3\epsilon} \frac{c_1}{\chi^{1/3} 3\sqrt{3}} \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{(d-2)}} \frac{p_0^2}{(\boldsymbol{P}^2)^{d/6}(\boldsymbol{P}^2 + 1)(dp_0^2 + \vec{p}^2)^{1/3}}, \tag{C24}$$

where we have also rescaled $\boldsymbol{P} \to \boldsymbol{P}/k_d^2$. In leading order in ϵ, the first factor can be approximated by 1. The Feynman parametrization (B15) with $\alpha_1 = d/6$, $\alpha_2 = 1$, and $\alpha_3 = 1/3$ then leads to

$$\Omega^1 = \frac{c_1}{6\sqrt{3}} \frac{\Gamma\left(\frac{4}{3} + \frac{d}{6}\right)}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{d}{6}\right)} \int \frac{dp_0}{2\pi} \frac{d\vec{p}}{(2\pi)^{(d-2)}} \int_0^1 dt_1 \int_0^{1-t_1} dt_2 \, t_1^{d/6-1}(1 - t_1 - t_2)^{-2/3}$$

$$\times \frac{p_0^2}{\left(t_1(p_0^2 + |\vec{p}|^2) + t_2(p_0^2 + |\vec{p}|^2 + 1) + (1 - t_1 - t_2)(dp_0^2 + |\vec{p}|^2)\right)^{4/3+d/6}}. \tag{C25}$$

Changing to polar coordinates, the integrals over $p_0, \vec{p}$, and t_2 are straightforwardly computed. The remaining t_1 integral can be evaluated numerically for $d = 5/2$. Performing the same steps for Ω^2 [cf. Eq. (C20)], in total one obtains, in leading order in ϵ:

$$\Omega = \Omega^1 + \Omega^2 \simeq \sigma_x 0.138 \frac{g^{4/3}}{\epsilon}. \tag{C26}$$

Let us now consider $A = \sigma_y$. Using

$$M_1 \sigma_i^T \sigma_y \sigma_j M_2 = \begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}\begin{matrix} i=x \\ i=y \\ i=z \end{matrix} \begin{matrix} j=x & j=y & j=z \\ \begin{pmatrix} -1 & i & 0 \\ -i & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{matrix}, \tag{C27}$$

we obtain

$$\Omega^1 = \begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix}(-g^2\mu^\epsilon)\int_{p^{d+1}} \frac{\boldsymbol{P}^2 + ip_0\delta_{-p} + ip_0\delta_p - \delta_p\delta_{-p}}{(\boldsymbol{P}^2 + \delta_p^2)(\boldsymbol{P}^2 + \delta_{-p}^2)}$$

$$\times D(k-p). \tag{C28}$$

The computations proceed largely analogous to the previous case of $A = \sigma_x$; in total, we obtain

$$\Omega \simeq -\sigma_y 0.219 \frac{g^{4/3}}{\epsilon}. \tag{C29}$$

To proceed, we need to evaluate the factor involving the spin matrix B in Eq. (C19), which yields

$$\left(\sigma_y^{\beta\delta}\sigma_y^{\alpha\gamma}B_{\alpha\beta}\right) = \begin{cases} -B_{\gamma\delta}, & B = \mathbb{1}, \sigma_y \\ B_{\gamma\delta}, & B = \sigma_x, \sigma_z \end{cases}. \tag{C30}$$

Before denoting which contributions are enhanced and which are suppressed, we notice that some products under consideration vanish trivially:

$$\Psi_\alpha^T \sigma_x \Psi_\beta B_{\alpha\beta} = \left(\bar{\psi}_\alpha^L \psi_\beta^R - \bar{\psi}_\beta^L \psi_\alpha^R\right)B_{\alpha\beta} = 0 \tag{C31}$$

$$\text{for } B = \mathbb{1}, \sigma_x, \sigma_z,$$

$$\Psi_\alpha^T \sigma_y \Psi_\beta \sigma_y^{\alpha\beta} = i\left(\bar{\psi}_\beta^L \psi_\alpha^R + \bar{\psi}_\alpha^L \psi_\beta^R\right)\sigma_y^{\alpha\beta} = 0. \tag{C32}$$

Altogether, the nonvanishing quantum corrections are, for $A = \sigma_x$,

$$V = \lambda \frac{u_\lambda g^{4/3}}{\epsilon} \int_{k^{d+1}} \Psi_\gamma^T(k)\sigma_x \Psi_\delta(-k)\sigma_y^{\gamma\delta}, \quad u_\lambda = -0.138. \tag{C33}$$

For $A = \sigma_y$,

$$V = \lambda \frac{u_\lambda g^{4/3}}{\epsilon} \int_{k^{d+1}} \Psi_\gamma^T(k) \sigma_y \Psi_\delta(-k) \times \begin{cases} -B_{\gamma\delta}, & B = \mathbb{1} \\ B_{\gamma\delta}, & B = \sigma_x, \sigma_z \end{cases},$$
$$u_\lambda = -0.219. \tag{C34}$$

Using Eq. (48), Eqs. (C33) and (C34) readily yield Table III.

APPENDIX D: SUPERCONDUCTING LOGARITHM

To clarify the role of the limit $\delta v \to 0$ applied in this paper, it is instructive to reevaluate the boson self-energy of Eq. (23) for $\delta v \neq 0$ and $d = 2$. Equation (23) then reads, up to constant prefactors,

$$\Pi(k) \propto \int_{p^{2+1}} \sum_{\alpha \neq \alpha'} \frac{1}{i p_0 - \delta_{-p}^\alpha} \frac{1}{i(k_0 - p_0) - \delta_{k-p}^{\alpha'}},$$
$$\delta_p^\alpha = v_\alpha p_x + p_y^2. \tag{D1}$$

Performing the integral over p_0 with help of Cauchy's theorem gives

$$\Pi(k) \propto \int dp_x dp_y \frac{1}{i k_0 - \delta_{-p}^\alpha - \delta_{k-p}^{\alpha'}}$$
$$\times \left(\theta(\delta_{-p}^\alpha) \theta(\delta_{k-p}^{\alpha'}) - \theta(-\delta_{-p}^\alpha) \theta(-\delta_{k-p}^{\alpha'}) \right), \tag{D2}$$

where $\alpha' \neq \alpha$. We introduce momentum cutoffs in the two directions p_x and p_y, Λ_x and Λ_y, with $\Lambda_x \simeq \Lambda_y^2$. Then, the p_x integral in (D2) gives

$$\Pi(k) \frac{1}{v_+} \int dp_y \left\{ \ln \left(i k_0 + v_+ p_x - p_y^2 - v_{\alpha'} k_x - (k_y - p_y)^2 \right) \Big|_{-\Lambda_x}^{\text{Mi}} - \ln \left(i k_0 + v_+ p_x - p_y^2 - v_{\alpha'} k_x - (k_y - p_y)^2 \right) \Big|_{\text{Ma}}^{\Lambda_x} \right\}, \tag{D3}$$

$$v_+ = v_\alpha + v_\alpha',$$

$$\text{Mi/Ma} = \min/\max \left(\frac{p_y^2}{v_\alpha}, k_x + \frac{(k_y - p_y)^2}{v_\alpha'} \right). \tag{D4}$$

Inserting the boundaries $\pm \Lambda_x$ yields terms of the form $2 \ln(v_+ \Lambda_x) + i\pi \text{sign}(k_0) + \mathcal{O}[(k,p)/\Lambda_x]$. These constant terms vanish once we subtract $\lim_{k_0 \to 0} \Pi(k_0, 0)$, which is legitimate when working at the critical point. By noticing that, if $\text{Ma} = p_y^2/v_\alpha$ in some integration region R_1, then $\text{Mi} = p_y^2/v_\alpha$ in $\mathbb{R} \setminus R_1$, we can recast the remainder in the following form:

$$\Pi(k) \sim \int dp_y \left\{ \ln \left(i k_0 + v p_y^2 - v_{\alpha'} k_x - (k_y - p_y)^2 \right) \right.$$
$$\left. + \ln \left(i k_0 + v^{-1}(k_y - p_y)^2 + v_\alpha k_x - p_y^2 \right) \right\}, \tag{D5}$$

where $v = v_{\alpha'}/v_\alpha$, and w.l.o.g. we assume $v > 1$. The remaining integral can be straightforwardly evaluated; inserting the boundaries $\pm \Lambda_y$ yields a long expression, which is of the schematic form

$$\Pi(k) \propto \frac{1}{v-1} \sqrt{(i k_0 - v_{\alpha'} k_x)(v-1) - k_y^2 v}$$
$$+ \frac{1}{|v^{-1} - 1|} \sqrt{(i k_0 + k_x v_\alpha)(v^{-1} - 1) - k_y^2 v^{-1}} + \Pi_{\text{div}},$$
$$\Pi_{\text{div}} \simeq \Lambda_y + \ln \left((\delta v)^2 + \frac{|k_y|}{\Lambda_y} \right). \tag{D6}$$

The first two terms of Eq. (D6) reproduce the result of Ref. [14]. For these terms, the limit $\delta v \to 0$, which is equivalent to $v \to 1$, can be taken, and results in a standard damping term; see also Appendix E. Let us now consider Π_{div}, the divergent part of Eq. (D6). For the first summand, Λ_y, the limit $\Lambda_y \to \infty$ corresponds to a pure UV divergence, which effectively arises from expansion of the fermion dispersion in the low-energy action (4). If higher order terms in the dispersion are taken into account, this UV singularity is absent, as numerically demonstrated in Ref. [14]; we can therefore disregard this

term. The second term is finite for $\delta v \neq 0$. In a fully realistic model of the FFLO transition, this condition is always fulfilled; increasing the magnetic field leads to increasing δv, and the phase transition takes place when $g - \Pi(\delta v, 0)$ vanishes (on mean-field level); here, g is the strength of the original four-fermion interaction. This happens at a small but nonzero value $\delta v = \delta v_c$. Thus, for $\delta v \simeq \delta v_c$, and $k_y \ll \Lambda_y$, i.e., when taking the limit $\Lambda_y \to \infty$ first, Π_{div} is just a finite mass term, which can be dropped when performing computations at the critical point. The remainder is regular in δv, and one can take the limit $\delta v \to 0$ to simplify the computation.

On the other hand, in the DIMREG computation we have to take the limit $\delta v \to 0$ first, [cf. Eq. (B3)], and are therefore left with the IR divergent quantity $\ln(|k_y|/\Lambda_y)$, a standard "BSC logarithm." To correct for this unphysical way of taking the limits, one must subtract $\Pi_{\text{div}}(k_y)$, as effectively done in Eq. (B9).

APPENDIX E: DIMENSIONAL REGULARIZATION WITH FIXED CO-DIMENSION

In this work, we have performed a DIMREG procedure by increasing the codimension of the Fermi surface. An alternative approach, shortly discussed in this Appendix, is to keep the codimension fixed, following Refs. [44,45,74]. That is, in the higher-dimensional action, the kinetic term for the fermions is modified to

$$\int_{k^{d+1}} \bar{\Psi}_\alpha(k)(-i k_0 \sigma_y + i(v_\alpha k_1 + \boldsymbol{K}^2)\sigma_x)\Psi_\alpha(k),$$
$$\boldsymbol{K} = (k_2, \ldots, k_d), \tag{E1}$$

with all other terms in the action unchanged. The leading terms in the action are then scale-invariant

under

$$k_0 = \frac{k_0'}{b}, \quad k_1 = \frac{k_1'}{b}, \quad \boldsymbol{K} = \frac{\boldsymbol{K}'}{\sqrt{b}}, \tag{E2}$$

$$\Psi = b^{\frac{d}{4}+\frac{5}{4}} \Psi'(k'), \quad \Delta(k) = b^{\frac{d}{4}+\frac{5}{4}} \Delta'(k'). \tag{E3}$$

With this scaling, the interaction term becomes marginal in $d = 3$, such that one can expand in $\epsilon = 3 - d$. In this scheme, evaluation of the Bose self-energy is very similar to the 2D-case sketched in Appendix D. It can be performed in the general case $\delta v \neq 0$ by employing the trivial reshuffling described above Eq. (D5). Taking all momentum cutoffs to infinity, and subtracting $\Pi(0)$ for regularization (which works for $\delta v \neq 0$, see Appendix D), one arrives at

$$\Pi(k) = \sum_{\alpha \neq \alpha'} \frac{\beta_d}{\cos(d\pi/2)} \left[\left(\frac{i k_0 - k_1 v_{\alpha'} - \boldsymbol{K}^2 \frac{v}{v-1}}{v - 1} \right)^{\frac{d-1}{2}} \right.$$
$$\left. + \left(\frac{i k_0 + k_1 v_\alpha - \boldsymbol{K}^2 \frac{v^{-1}}{v^{-1}-1}}{v^{-1} - 1} \right)^{\frac{d-1}{2}} \right], \tag{E4}$$

where $\beta_d > 0$ is a d-dependent factor of order 1. For $d \to 3$, Π is ϵ divergent due to the term $\cos(d\pi/2)$. To gain analytical control, one can again expand in δv, which leads to

$$\lim_{\delta v \to 0} \Pi(k) = \frac{\beta_d}{\cos(d\pi/2)|\delta v|^{d-1}} \left[4 \cos \left(\frac{(d-1)\pi}{2} \right) |\boldsymbol{K}|^{d-1} \right.$$
$$\left. - 4 \frac{(d-1)}{2} \cos(d\pi/2)|k_0 \delta v||\boldsymbol{K}|^{d-3} \right]. \tag{E5}$$

In $d = 2$, the prefactor of the term $\propto |\boldsymbol{K}|^{d-1}$ vanishes, and the remainder is the damping term of Ref. [14], and regular as $\delta v \to 0$. However, for $2 < d < 3$, the first term does not vanish, and $\Pi(k)$ is divergent as $\delta v \to 0$. This can be seen as an instance of UV/IR mixing [44]. As discussed in the main text (see Fig. 5), for $\delta v \to 0$, spin-up and spin-down Fermi sheets have the same curvature. As a result, any spin-up electron with momentum $\mathbf{k}_1$ on the Fermi surface can scatter against a spin-down electron with momentum $-\mathbf{k}_1$. However, if the Fermi surface is one-dimensional, the final states of this scattering event must have momenta $\pm \mathbf{k}_2 \simeq \mathbf{k}_1$; otherwise, the tangent vectors to the Fermi surface differ strongly, and the phase space for the scattering is negligible. By contrast, for

a Fermi surface with dimension greater than one, all points of the Fermi surface share a mutual tangent vector. Therefore low-energy scattering events entangle the full Fermi surface, and the hot spot theory breaks down, as signaled by the $\delta v \to 0$ divergence of Eq. (E5).

APPENDIX F: MAGNETIC SUSCEPTIBILITY

In this Appendix, we shortly present the evaluation of the magnetic susceptibility close to criticality. We limit ourselves to evaluation of the functional behavior (up to a constant prefactor).

If the contribution of the fermions is neglected (or, phrased differently, they have been integrated out on one-loop level), the free energy on the normal metal side reads, for $d = 2$:

$$F_\Delta = -\ln[Z_\Delta]$$
$$= -\ln \left[\int \mathcal{D}(\Delta, \bar{\Delta}) \exp \left(-\int d^3k \, D^{-1}(k)|\Delta(k)|^2 \right) \right]$$
$$\propto \ln[\det(D^{-1})] = \int d^3k [\ln(D^{-1}(k))]. \tag{F1}$$

Therefore the fluctuation contribution to the magnetic susceptibility is given by [15,37]

$$\chi_\Delta \propto -\frac{\partial^2 F_\Delta}{\partial h^2} \propto -\frac{\partial^2 F_\Delta}{\partial m^2} \propto \int d^3k \frac{-1}{\left(m + k_y^2 + \frac{\alpha|k_0|}{|k_y|} \right)^2}, \tag{F2}$$

where we reintroduced the mass term ($m > 0$) into the 2D boson propagator [see Eqs. (2) and (43)], and used that m is proportional to the reduced magnetic field, $m = m_0 \frac{h-h_c}{h_c}$; m_0 and $\alpha \propto g^2$ are constants. Easy integration yields

$$\chi_\Delta \propto \Lambda_x \ln \left(\frac{\Lambda_y^2}{\frac{h-h_c}{h_c} m_0} \right), \tag{F3}$$

where Λ_x and Λ_y are UV cutoffs in the x and y directions (of order of Fermi energies). Normalizing χ_Δ with the Pauli spin susceptibility in the normal state χ_P as in Ref. [15], and fixing the prefactors, on can conclude

$$\frac{\chi_\Delta}{\chi_P} \simeq \frac{\Delta_0}{E_F} \ln \left(\frac{h_c}{h - h_c} \right), \tag{F4}$$

where Δ_0 is the BCS gap and E_F the Fermi energy.

[1] H. v. Löhneysen, A. Rosch, M. Vojta, and P. Wölfle, Rev. Mod. Phys. **79**, 1015 (2007).

[2] G. R. Stewart, Rev. Mod. Phys. **73**, 797 (2001).

[3] B. L. Altshuler, L. B. Ioffe, and A. J. Millis, Phys. Rev. B **50**, 14048 (1994).

[4] Y. B. Kim, A. Furusaki, X.-G. Wen, and P. A. Lee, Phys. Rev. B **50**, 17917 (1994).

[5] W. Metzner, D. Rohe, and S. Andergassen, Phys. Rev. Lett. **91**, 066402 (2003).

[6] S.-S. Lee, Phys. Rev. B **80**, 165102 (2009).

[7] M. A. Metlitski and S. Sachdev, Phys. Rev. B **82**, 075127 (2010).

[8] M. A. Metlitski and S. Sachdev, Phys. Rev. B **82**, 075128 (2010).

[9] E. Berg, M. A. Metlitski, and S. Sachdev, Science **338**, 1606 (2012).

[10] T. Senthil and R. Shankar, Phys. Rev. Lett. **102**, 046406 (2009).

[11] D. Dalidovich and S.-S. Lee, Phys. Rev. B **88**, 245106 (2013).

[12] P. Fulde and R. A. Ferrell, Phys. Rev. **135**, A550 (1964).

[13] A. I. Larkin and Y. N. Ovchinikov, Sov. Phys. JETP **20**, 762 (1965).

[14] F. Piazza, W. Zwerger, and P. Strack, Phys. Rev. B **93**, 085112 (2016).

[15] K. V. Samokhin and M. S. Mar'enko, Phys. Rev. B **73**, 144502 (2006).

[16] M. D. Croitoru and A. I. Buzdin, Condens. Matter **2**, 30 (2017).

[17] J. Shinagawa, Y. Kurosaki, F. Zhang, C. Parker, S. E. Brown, D. Jérome, K. Bechgaard, and J. B. Christensen, Phys. Rev. Lett. **98**, 147002 (2007).

[18] S. Yonezawa, S. Kusaba, Y. Maeno, P. Auban-Senzier, C. Pasquier, K. Bechgaard, and D. Jérome, Phys. Rev. Lett. **100**, 117002 (2008).

[19] H. Mayaffre, S. Krämer, M. Horvatić, C. Berthier, K. Miyagawa, K. Kanoda, and V. Mitrović, Nat. Phys. **10**, 928 (2014).

[20] J. Wosnitza, Ann. Phys. (Leipzig) **530**, 1700282 (2017).

[21] Y. Matsuda and H. Shimahara, J. Phys. Soc. Jpn. **76**, 051005 (2007).

[22] A. Ptok, J. Phys.: Condens. Matter **29**, 475901 (2017).

[23] K. Cho, H. Kim, M. A. Tanatar, Y. J. Song, Y. S. Kwon, W. A. Coniglio, C. C. Agosta, A. Gurevich, and R. Prozorov, Phys. Rev. B **83**, 060502 (2011).

[24] D. A. Zocco, K. Grube, F. Eilers, T. Wolf, and H. v. Löhneysen, Phys. Rev. Lett. **111**, 057007 (2013).

[25] P. W. Adams, H. Nam, C. K. Shih, and G. Catelani, Phys. Rev. B **95**, 094520 (2017).

[26] V. V. Ryazanov, V. A. Oboznov, A. Y. Rusanov, A. V. Veretennikov, A. A. Golubov, and J. Aarts, Phys. Rev. Lett. **86**, 2427 (2001).

[27] V. A. Oboznov, V. V. Bol'ginov, A. K. Feofanov, V. V. Ryazanov, and A. I. Buzdin, Phys. Rev. Lett. **96**, 197003 (2006).

[28] K. Miyake, S. Schmitt-Rink, and C. M. Varma, Phys. Rev. B **34**, 6554 (1986).

[29] D. J. Scalapino, E. Loh, and J. E. Hirsch, Phys. Rev. B **34**, 8190 (1986).

[30] S. Sachdev, M. A. Metlitski, and M. Punk, J. Phys.: Cond. Mat. **24**, 294205 (2012).

[31] S. Lederer, Y. Schattner, E. Berg, and S. A. Kivelson, Phys. Rev. Lett. **114**, 097001 (2015).

[32] M. A. Metlitski, D. F. Mross, S. Sachdev, and T. Senthil, Phys. Rev. B **91**, 115111 (2015).

[33] I. Mandal, Phys. Rev. B **94**, 115138 (2016).

[34] Y. Schattner, M. H. Gerlach, S. Trebst, and E. Berg, Phys. Rev. Lett. **117**, 097002 (2016).

[35] S. Lederer, Y. Schattner, E. Berg, and S. A. Kivelson, Proc. Natl. Acad. Sci. USA **114**, 4905 (2017).

[36] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, Cambridge, 2011).

[37] A. Larkin and A. Varlamov, in *Superconductivity* (Springer, Berlin, Heidelberg, 2008), pp. 369–458.

[38] J. Wosnitza (private communication).

[39] R. M. Lutchyn, M. Dzero, and V. M. Yakovenko, Phys. Rev. A **84**, 033609 (2011).

[40] D. E. Sheehy and L. Radzihovsky, Ann. Phys. (NY) **322**, 1790 (2007).

[41] J. J. Kinnunen, J. E. Baarsma, J.-P. Martikainen, and P. Törmä, Rep. Prog. Phys. **81**, 046401 (2018).

[42] A. I. Buzdin and H. Kachkachi, Phys. Lett. A **225**, 341 (1997).

[43] L. Radzihovsky and A. Vishwanath, Phys. Rev. Lett. **103**, 010404 (2009).

[44] I. Mandal and S.-S. Lee, Phys. Rev. B **92**, 035141 (2015); I. Mandal, Eur. Phys. J. B **89**, 278 (2016).

[45] S.-S. Lee, Ann. Rev. Condens. Matter Phys. **9**, 227 (2018).

[46] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4-Theories* (World Scientific, Singapore, 2001).

[47] A. N. Vasiliev, *The Field Theoretic Renormalization Group in Critical Behavior Theory and Stochastic Dynamics* (CRC press, Boca Raton, 2004).

[48] P. Jakubczyk, Phys. Rev. A **95**, 063626 (2017).

[49] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Westview Press, Boulder, USA, 1995).

[50] G. D. Mahan, *Many Particle Physics*, 3rd ed. (Kluwer Academic/Plenum, New York and London, 2000).

[51] H. Burkhardt and D. Rainer, Ann. Phys. (Leipzig) **506**, 181 (1994).

[52] C. Mora and R. Combescot, Europhys. Lett. **66**, 833 (2004).

[53] R. A. Ferrell, J. Low Temp. Phys. **1**, 423 (1969).

[54] D. J. Scalapino, Phys. Rev. Lett. **24**, 1052 (1970).

[55] J. T. Anderson and A. M. Goldman, Phys. Rev. Lett. **25**, 743 (1970).

[56] J.-H. She, B. J. Overbosch, Y.-W. Sun, Y. Liu, K. E. Schalm, J. A. Mydosh, and J. Zaanen, Phys. Rev. B **84**, 144527 (2011).

[57] A. Abanov, A. V. Chubukov, and J. Schmalian, Adv. Phys. **52**, 119 (2003).

[58] L. Dell'Anna and W. Metzner, Phys. Rev. B **73**, 045127 (2006).

[59] M. Punk, Phys. Rev. B **94**, 195113 (2016).

[60] A. Eberlein, I. Mandal, and S. Sachdev, Phys. Rev. B **94**, 045133 (2016).

[61] A. A. Patel, P. Strack, and S. Sachdev, Phys. Rev. B **92**, 165105 (2015).

[62] I. Mandal, Ann. Phys. (NY) **376**, 89 (2017).

[63] R. Lortz, Y. Wang, A. Demuer, P. H. M. Böttger, B. Bergk, G. Zwicknagl, Y. Nakazawa, and J. Wosnitza, Phys. Rev. Lett. **99**, 187002 (2007).

[64] C. C. Agosta, N. A. Fortune, S. T. Hannahs, S. Gu, L. Liang, J.-H. Park, and J. A. Schleuter, Phys. Rev. Lett. **118**, 267001 (2017).

[65] M. A. Metlitski and S. Sachdev, New J. Phys. **12**, 105007 (2010).

[66] J. Gukelberger, S. Lienert, E. Kozik, L. Pollet, and M. Troyer, Phys. Rev. B **94**, 075157 (2016).

[67] S. Sur and S.-S. Lee, Phys. Rev. B **91**, 125136 (2015).

[68] R. M. Nandkishore and S. A. Parameswaran, Phys. Rev. B **95**, 205106 (2017); I. Mandal and R. M. Nandkishore, *ibid.* **97**, 125121 (2018).

[69] L. G. Aslamazov, Sov. Phys. JETP **28**, 773 (1969).

[70] A. G. Lebed and S. Wu, Phys. Rev. B **82**, 172504 (2010).

[71] A. Altland and B. Simons, *Condensed Matter Field Theory*, 2nd ed. (Cambridge Univerisity Press, Cambridge, UK, 2010).

[72] S. Sachdev, *Quantum Phase Transitions* (Cambridge Univerisity Press, Cambridge, UK, 2011).

[73] P. Zdybel and P. Jakubczyk, J. Phys.: Condens. Matter **30**, 305604 (2018).

[74] S. Chakravarty, R. E. Norton, and O. F. Syljuåsen, Phys. Rev. Lett. **74**, 1423 (1995).

3.3 Incommensurate $2k_F$ density wave quantum criticality

3.3.1 Overview

Arguably the simplest manifestation of a charge density wave (CDW) transition is the Peierls instability (see, e.g., [Hal19]): one considers a monoatomic chain with lattice constant a, and one valence electron per atom, corresponding to a commensurate Fermi wave vector $k_F = \pi/2a$. If the electrons are weakly coupled to phonons, a gap at k_F opens in the fermion dispersion, and occupied electronic states are shifted downwards in energy. As can be shown via mean field analysis, this results in a periodic modulation of the electron density $\rho(x) = \langle \bar{\psi}(x)\psi(x) \rangle = \rho_0 + \rho_1 \cos(2k_F x)$.

The Peierls instability only requires a weak coupling to the phonons, since the ultimate driving force behind it are electron-hole excitations of the Fermi sea. This is most clearly seen in $d = 1$, where the Lindhard function (imaginary part of the particle-hole bubble) is logarithmically peaked at $Q = 2k_F$ [Mih11], which can be incommensurate as well. In 2D, CDWs mostly determined by electronic correlations can still arise, but usually require some variety of Fermi surface nesting, akin to an effective one-dimensionality of the problem.

The prime example for two-dimensional materials with a CDW driven by Fermi surface nesting are the rare earth tellurites [GDC$^+$98, YRKK06], where the Fermi surface nesting can e.g. be deduced from angle-resolved photoemission spectroscopy. Another interesting material class, where CDWs can be directly observed in X-ray diffraction [FWJ$^+$12, FvWW$^+$15, WMH$^+$08] or scanning tunneling microscopy [SYH$^+$13, CPC$^+$18] are transition-metal dichalcogenides (TMDs) such as NbSe$_2$, VaSe$_2$ or TaS$_2$. The precise origin of CDW formation in the TMDs is still under debate: while Fermi surface nesting appears to be important in some cases, e.g. in VaSe$_2$ [CPC$^+$18], strong electron-phonon coupling could also play a role [FvW15] in these materials.

At $T = 0$, an CDW quantum critical point can be reached in the TMDs by applying pressure [FWJ$^+$12, FvWW$^+$15], or inserting a different transition metal [WMH$^+$08]. As usual, this QCP is hidden beneath a superconducting dome, but its precise experimental characterization is not yet available and is a worthwhile goal for future studies.

From the theory side, CDW quantum criticality was first studied by Altshuler et al. [AIM95] in a hot spot type setup as already spelled out in Eq. (3.5): a CDW order parameter field ϕ peaked at $\mathbf{Q} = 2k_F$ couples to fermions at two hot spots in a pseudo-nested configuration. Altshuler and coworkers considered both the commensurate and incommensurate case; for the latter, it was found that two-loop contributions lead to strong infrared divergences in the bosonic self-energy, suggesting a first order transition.

This pioneering work was later reconsidered by Metzner and collaborators [HM14, SHM18] in a RPA-type approach. While they reproduced (and extendend) the results of Altshuler et al., in the second paper [SHM18] they also obtained logarithmic corrections to the fermion dispersion, leading to an increased dynamical nesting. They emphasized that this effect could potentially render the phase transition continuous again.

Applying the controlled ϵ-expansion, we have rigorously confirmed this notion in the article below. In particular, we found that the linear and quadratic terms in the hot spot fermion dispersion $\delta_k = k_x + k_y^2$ renormalize independently. We obtained several fixed points, but at the one which is physically meaningful the Fermi surface is strongly flattened as shown in Fig. 3.7. Quite unexpectedly, Landau damping turned out to be unimportant as $\epsilon \to 0$. Furthermore, our fermion self-energy is frequency independent to leading order, and thus the fermionic quasiparticles retain their integrity at the fixed point; still, we obtain interesting power law behaviour for the correlation functions. For example, the bosonic two-point function $D(\omega)$, which corresponds to the density susceptibility and could be directly

resolved e.g. via inelastic X-ray scattering, scales as $D(\omega) \propto |\omega|^{-0.616}$. An experimental observation of this behaviour in the scaling region of the fixed point could help to clarify the origin of CDW formation in the TMDs in favour of electronic nesting.
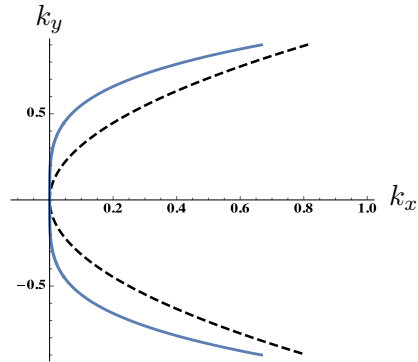

Figure 3.7 Dynamical flattening of the Fermi surface. The black dashed line shows the bare (UV) Fermi surface expanded around a hot spot, the blue line shows the renormalized Fermi surface at the CDW fixed point. Adapted from our publication [HPP19].

Incommensurate 2kF density wave quantum criticality in two dimensional metals

by

J. Halbinger, **D. Pimenov**,[1] and M. Punk[1]

[1] Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for

NanoScience, and Munich Center for Quantum Science and Technology (MCQST),

Ludwig-Maximilians-Universität München, 80333 München, Germany

Incommensurate $2k_F$ density wave quantum criticality in two-dimensional metals

Johannes Halbinger, Dimitri Pimenov, and Matthias Punk

Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for NanoScience,
and Munich Center for Quantum Science and Technology (MCQST), Ludwig-Maximilians University, Munich, Germany

We revisit the problem of two-dimensional metals in the vicinity of a quantum phase transition to incommensurate $\mathbf{Q} = 2k_F$ charge-density-wave order, where the order-parameter wave vector $\mathbf{Q}$ connects two hot spots on the Fermi surface with parallel tangents. Earlier theoretical works argued that such critical points are potentially unstable, if the Fermi surface at the hot spots is not sufficiently flat. Here we perform a controlled, perturbative renormalization-group analysis and find a stable fixed point corresponding to a continuous quantum phase transition, which exhibits a strong dynamical nesting of the Fermi surface at the hot spots. We derive scaling forms of correlation functions at the critical point and discuss potential implications for experiments with transition-metal dichalcogenides and rare-earth tellurides.

I. INTRODUCTION

While Landau's Fermi liquid theory has been tremendously successful in describing the properties of ordinary metals, a variety of strongly correlated electron materials show an unusual strange metal or non-Fermi-liquid behavior, which is not captured within the Fermi-liquid framework. It is usually characterized by a linear temperature dependence of the resistivity, as well as an absence of resistivity saturation at the Mott-Ioffe-Regel limit [1], which is taken as evidence for the absence of well-defined electronic quasiparticle excitations [2]. Examples include the strange metal phases observed in cuprate and iron pnictide superconductors [3,4], heavy-fermion materials [5], and also in twisted bilayer graphene [6].

Quantum critical points in metals are a promising theoretical scenario giving rise to non-Fermi-liquid phenomenology [7]. Indeed, in two spatial dimensions the strong coupling between gapless order-parameter fluctuations and particle-hole excitations at the Fermi surface leads to a loss of electronic quasiparticle coherence and to a strong damping of order-parameter fluctuations. Even though the computation of transport coefficients in these models remains a big challenge, two notable theoretical developments have considerably advanced our understanding of metallic quantum critical points: First, some lattice models of electrons coupled to a bosonic order parameter can be numerically studied using determinant quantum Monte Carlo methods avoiding the infamous fermion sign problem [8–11]. Second, epsilon expansions have been developed for hot-spot models of quantum critical points in metals, allowing us to characterize non-Fermi-liquid fixed points within a controlled renormalization-group approach, where the bosonic and fermionic degrees of freedom are treated on an equal footing [12–15].

Due to the fact that strange metals are often found in the regime between a magnetically ordered phase and an ordinary Fermi liquid, a lot of theoretical work has focused on the study of commensurate spin-density-wave quantum criticality

in metals [10,14,16–22]. In this work, we consider incommensurate charge-density-wave (CDW) quantum critical points in quasi-two-dimensional metals instead, where the electron density spontaneously breaks translational symmetries and develops a density modulation with a wave vector $\mathbf{Q}$ that is incommensurate with the underlying crystalline lattice. In particular, we are interested in systems in which the CDW ordering wave vector $\mathbf{Q} = 2k_F$ is determined by a partial nesting condition of the Fermi surface and connects two points on the Fermi surface with parallel tangents. This is to be distinguished from perfect nesting, where entire sections of the Fermi surface are connected by the same $2k_F$ wave vector.

The properties of $2k_F$ density wave quantum critical points in two-dimensional metals have been analyzed in previous theoretical works [23–29]. While an early study by Altshuler *et al.* [23] concluded that the incommensurate transition is of first order due to strong fluctuations, a more recent article by Sykora *et al.* [28] pointed out that the transition is potentially continuous if the Fermi surface is sufficiently flat at the hot spots. In this work, building upon the epsilon expansion by Dalidovich and Lee [13], we resolve this open problem by performing a controlled renormalization-group (RG) analysis of such an incommensurate $\mathbf{Q} = 2k_F$ CDW transition. We show that there is a strong dynamical nesting of the Fermi surface at the two hot spots connected by the $2k_F$ wave vector, and we identify a stable RG fixed point corresponding to a continuous quantum phase transition. Furthermore, we compute critical exponents and the scaling form of correlation functions at the non-Fermi-liquid fixed point to leading order in epsilon, and we point out experimental signatures.

CDW order plays an interesting role in underdoped cuprates and has been observed in a variety of quasi-two-dimensional materials such as transition-metal dichalcogenides and rare-earth tellurides. Several of these materials exhibit a CDW ordering wave vector that is commensurate with the crystalline lattice, implying that the transition is likely driven by the coupling to phonons. A few notable

exceptions with incommensurate CDW order exist, such as the 2H forms of NbSe$_2$ and TaS$_2$ [30,31], VSe$_2$ [32], as well as SmTe$_3$ and TbTe$_3$ [33,34]. In some of these compounds, the CDW transition temperature can be tuned to zero across a potential quantum critical point by intercalation or applying pressure, indicating that electronic correlations could be the main driving force behind the CDW transition [35]. Moreover, some of these materials become superconducting at low temperatures in the vicinity of the putative CDW quantum critical point [36–38]. One of our aims is to provide clear experimental signatures of an incommensurate $2k_F$ CDW quantum phase transition, which would allow us to settle the question of whether the incommensurate CDW transition in materials like NbSe$_2$ is driven by electronic correlations, or by a different mechanism, such as the coupling to phonons.

The rest of this work is outlined as follows: In Sec. II we introduce the model of electrons coupled to charge-density-wave fluctuations in two dimensions, as well as a generalization to higher dimensions, which is amenable to dimensional regularization. Section III presents our results for the one-loop boson and fermion self-energies in arbitrary dimensions. The RG flow equations, their fixed-point structure, and the scaling form of the boson and fermion two-point correlators are presented in Sec. IV. A discussion of experimental signatures follows in Sec. V. Finally, results on superconducting instabilities in the vicinity of the QCP are presented in Sec. VI. We close with discussions and conclusions in Sec. VII.

II. MODEL

We start from a theory of electrons coupled to charge-density-wave fluctuations in two spatial dimensions described by the Euclidean action

$$S = \int_k \psi^\dagger(k)(-ik_0 + \xi_{\mathbf{k}})\psi(k) + \frac{1}{2}\int_q \phi(q)\,\chi_q^{-1}\,\phi(-q)$$
$$+ \lambda \int_{k,q} \phi(q)\psi^\dagger(k+q)\psi(k), \tag{1}$$

where the fermionic field $\psi(k)$ (spin index suppressed) describes electrons with frequency/momentum $k = (k_0, \mathbf{k})$, the electron dispersion measured from the Fermi energy is denoted by $\xi_{\mathbf{k}}$, the real field $\phi(q)$ describes CDW fluctuations, and $\chi_q = \chi_{-q}$ is the bare CDW susceptibility, which we assume to be peaked at the incommensurate $2k_F$ wave vectors $\pm\mathbf{Q}$. Consequently, electrons scatter predominantly in the vicinity of two hot spots connected by the vector $\mathbf{Q}$ (see Fig. 1). A finite order-parameter expectation value $\langle\phi\rangle \neq 0$ gives rise to a ground state with a spatially modulated electron density.

We expand the action around the hot spots by writing $\mathbf{q} = \pm\mathbf{Q} + \mathbf{p}$ and denoting $\phi(q_0, \pm\mathbf{Q} + \mathbf{p}) \equiv \phi^\pm(q_0, \mathbf{p}) \equiv \phi^\pm(p)$. Analogously, we denote the fermion fields in the vicinity of the two hot spots by $\psi(k_0, \pm\mathbf{Q}/2 + \mathbf{k}) \equiv \psi_\pm(k)$. Expanding the electron dispersion as well as the CDW susceptibility to second order around the hot-spot momenta, we thus obtain
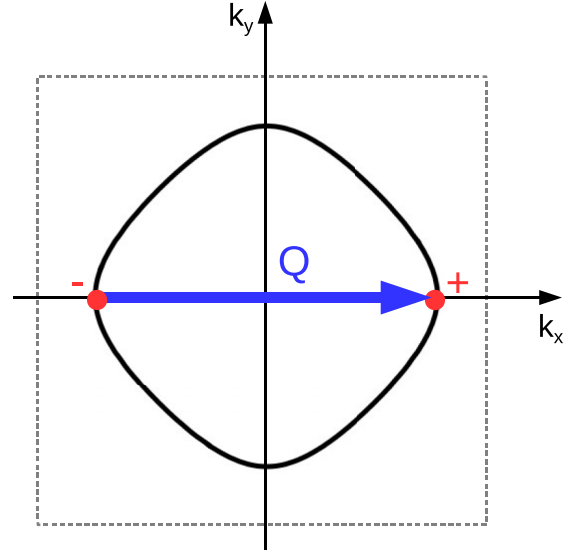

FIG. 1. Scattering geometry for electrons coupled to incommensurate $\mathbf{Q} = 2k_F$ order-parameter fluctuations. The order-parameter wave vector $\mathbf{Q}$ (blue arrow) couples electrons in the vicinity of two hot spots (denoted by $-$ and $+$), where the Fermi surface (thick black line) has parallel tangents.

the low-energy effective action

$$S = \sum_{s=\pm}\sum_{j=1}^{N}\int_k \psi_{s,j}^\dagger(k)\big(-ik_0 + sk_x + k_y^2\big)\psi_{s,j}(k)$$
$$+ \int_k \phi^+(k)\big(k_0^2 + k_x^2 + k_y^2\big)\phi^-(-k)$$
$$+ \frac{\lambda}{\sqrt{N}}\sum_{j=1}^{N}\int_{k,p}[\phi^+(p)\psi_{+,j}^\dagger(k+p)\psi_{-,j}(k)$$
$$+ \phi^-(-p)\psi_{-,j}^\dagger(k-p)\psi_{+,j}(k)], \tag{2}$$

where we have generalized the model to allow for N distinct fermionic species (with $N = 2$ for spin-1/2 fermions), and we tuned the model to the quantum critical point, where the mass term for the order-parameter field vanishes. Note that momenta in the fermion kinetic term are rescaled such that all proportionality constants are equal to unity. As discussed in detail below, only the $\sim k_y^2$ term in the kinetic part of the boson is relevant in the RG sense, while all other terms are irrelevant and will be discarded in the following. For this reason, we have set their proportionality constants to unity as well.

To perform an epsilon expansion, we generalize this action to arbitrary dimensions by increasing the codimension of the Fermi surface, following earlier work by Dalidovich and Lee [13]. For this reason, it is convenient to define the spinor

$$\Psi_j(k) = \begin{pmatrix} \psi_{+,j}(k) \\ \psi_{-,j}^\dagger(-k) \end{pmatrix}, \tag{3}$$

and to rewrite the action (2) as

$$S = \sum_j \int_k \overline{\Psi}_j(k)[-i\sigma_y k_0 + i\sigma_x \delta_k]\Psi_j(k)$$
$$+ \int_k \phi^+(k)\big(k_0^2 + k_x^2 + k_y^2\big)\phi^-(-k)$$

$$- \frac{i\lambda}{2\sqrt{N}} \sum_j \int_{k,p} \left[\phi^+(p)\overline{\Psi}_j(k+p)\sigma_y\overline{\Psi}_j^T(-k) \right.$$

$$\left. + \phi^-(-p)\Psi^T(p-k)\sigma_y\Psi(k) \right] \tag{4}$$

with

$$\overline{\Psi} = \Psi^\dagger\sigma_y, \quad \delta_k = k_x + k_y^2. \tag{5}$$

A canonical way to generalize this action to d spatial dimensions while keeping the action local is to write the fermion kinetic term as

$$\sum_j \int_{k^{d+1}} \overline{\Psi}_j(k)[-i\mathbf{\Gamma} \cdot \mathbf{K} + i\sigma_x\delta_k]\Psi_j(k), \tag{6}$$

where we defined

$$\mathbf{\Gamma} = (\sigma_y, \sigma_z, \ldots, \sigma_z), \quad \mathbf{K} = (k_0, k_1, \ldots, k_{d-2}) = (k_0, \mathbf{k}),$$

$$\delta_k = k_{d-1} + k_d^2, \quad \int_{k^{d+1}} = \int \frac{d^{d-1}\mathbf{K}dk_{d-1}dk_d}{(2\pi)^{d+1}}. \tag{7}$$

The additional momenta $k_1, \ldots, k_{d-2}$ correspond to the new directions perpendicular to the Fermi surface, and due to the Dirac structure of the action the fermions have a linear dispersion in these directions. The kinetic term of the boson is generalized accordingly to

$$\int_{k^{d+1}} \phi^+(k)\left(\mathbf{K}^2 + k_{d-1}^2 + k_d^2\right)\phi^-(-k) \tag{8}$$

in arbitrary dimensions.

The quadratic terms in the action are invariant under the scaling transformation

$$\mathbf{K} = \frac{\mathbf{K}'}{b}, \quad k_{d-1} = \frac{k'_{d-1}}{b}, \quad k_d = \frac{k'_d}{b^{\frac{1}{2}}},$$

$$\Psi(k) = \Psi'(k')b^{\frac{d}{2}+\frac{3}{4}}, \quad \phi^\pm(k) = \phi'^\pm(k')b^{\frac{d}{2}+\frac{3}{4}}. \tag{9}$$

A consequence of this scaling transformation is that all terms in the boson propagator apart from the $\sim k_d^2$ term are irrelevant at tree level and can be neglected in the following computations. However, as we will discuss in detail below, the renormalization-group flow generates a *linear* term $\sim k_{d-1}$ in the boson propagator, which is relevant, allowed by symmetry, and crucial to remove infrared divergences. This term arises from the fact that the susceptibility is enhanced along the line of $2k_F$ wave vectors $\mathbf{k}$ defined via $\xi_{(\mathbf{k}+\mathbf{G})/2} = 0$, with $\mathbf{G}$ an arbitrary reciprocal-lattice vector, connecting points on the Fermi surface with parallel tangents [25]. For this reason, we add the term ak_{d-1} to the boson propagator from the start, where a is a dimensionless parameter that will flow under the RG. The coupling constant λ transforms as

$$\lambda' = \lambda b^{\frac{1}{2}(\frac{5}{2}-d)}, \tag{10}$$

consequently interactions are irrelevant in $d > 5/2$ and we can perform a controlled expansion in small $\epsilon = \frac{5}{2} - d$. As usual, we define a dimensionless coupling constant by introducing an arbitrary mass scale μ via the replacement $\lambda \rightarrow \lambda\mu^{\epsilon/2}$. Our final form of the action in general dimensions thus

reads

$$S = \sum_j \int_{k^{d+1}} \overline{\Psi}_j(k)[-i\mathbf{\Gamma} \cdot \mathbf{K} + i\sigma_x\delta_k]\Psi_j(k)$$

$$+ \int_{k^{d+1}} \phi^+(k)\left(k_d^2 + ak_{d-1}\right)\phi^-(-k)$$

$$- \frac{i\lambda\mu^{\epsilon/2}}{2\sqrt{N}} \sum_j \int_{k^{d+1},p^{d+1}} \left[\phi^+(p)\overline{\Psi}_j(k+p)\sigma_y\overline{\Psi}_j^T(-k) \right.$$

$$\left. + \phi^-(-p)\Psi^T(p-k)\sigma_y\Psi(k) \right]. \tag{11}$$

In the following, we study this action within a field-theoretic renormalization-group approach using dimensional regularization and the minimal subtraction scheme. For this reason, we compute one-loop diagrams and extract the $1/\epsilon$ counterterms in the next section.

III. ONE-LOOP DIAGRAMS

The bare fermion and boson propagators for the theory in Eq. (11) take the form

$$G(k) = \langle\Psi(k)\overline{\Psi}(k)\rangle_0 = -i\frac{-\mathbf{\Gamma} \cdot \mathbf{K} + \sigma_x\delta_k}{\mathbf{K}^2 + \delta_k^2},$$

$$D_+(k) = \langle\phi^+(k)\phi^-(-k)\rangle_0 = \frac{1}{k_d^2 + ak_{d-1}}. \tag{12}$$

Analogously we define $D_-(k) = \langle\phi^-(k)\phi^+(-k)\rangle_0 \equiv D_+(-k)$. Even though this seems like a redundant definition, it is important to distinguish bosonic degrees of freedom in the vicinity of the $2k_F$ wave vector $\mathbf{Q}$ and $-\mathbf{Q}$ and the linear term $\sim ak_{d-1}$ in the boson propagator is allowed by this symmetry. The one-loop boson self-energy $\Pi_+(k) \equiv \Pi_-(-k)$ is given by the integral

$$\Pi_+(k) = -\frac{\lambda^2\mu^\epsilon}{2N} \int_{p^{d+1}} \text{Tr}[\sigma_y G(p)\sigma_y G^T(k-p)] \tag{13}$$

and evaluates to (details can be found in Appendix A)

$$\Pi_+(k) = -u_1\lambda^2\frac{e_k}{\epsilon} - u_2\lambda^2\frac{|\mathbf{K}|^{\frac{3}{2}}}{\sqrt{|e_k|}}\Theta(-e_k) + \cdots, \tag{14}$$

where $\Theta(x)$ is the unit step function, and we expanded the self-energy around $\epsilon \approx 0$ as well as around $|\mathbf{K}| \approx 0$ and defined

$$u_1 = \frac{\Gamma\left(\frac{5}{4}\right)}{8\sqrt{2}\pi^{\frac{7}{4}}} \approx 0.0108, \quad u_2 = \frac{\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{5}{4}\right)}{16\sqrt{2}\pi^{\frac{5}{4}}\Gamma\left(\frac{7}{4}\right)} \approx 0.0378,$$

$$e_k = k_{d-1} + \frac{1}{2}k_d^2. \tag{15}$$

Note that $e_k = 0$ defines the line of $2k_F$ momenta. The frequency-dependent term $\sim|\mathbf{K}|^{3/2}$ is the $d = 5/2$-dimensional analog of the Landau damping term $\sim|k_0|$ in two dimensions. Note that this term does not have a $1/\epsilon$ pole and thus does not renormalize. Consequently, no frequency-dependent terms are generated in the boson propagator during the one-loop RG flow. On the other hand, the $\sim 1/\epsilon$ term is proportional to k_d^2 and to k_{d-1} and thus generates an RG flow of the corresponding terms in the boson propagator.

The fermion self-energy is given by the integral

$$\Sigma(k) = \frac{\lambda^2 \mu^\epsilon}{N} \int_{p^{d+1}} \sigma_y G^T(p-k) \sigma_y D_+(p). \qquad (16)$$

As shown in Appendix B, this integral evaluates to

$$\Sigma(k) - \Sigma(0) = \frac{i\sigma_x}{\epsilon} \frac{2u_1 \lambda^2}{(1-a)\sqrt{|\tilde{a}|N}} \left(\frac{k_d^2}{1-a} - k_{d-1} \right) \\ + \text{finite terms}, \qquad (17)$$

where $\tilde{a} = \frac{a}{1-a}$. As discussed in detail below, these terms renormalize the fermion dispersion. The fact that all frequency-dependent terms in the boson propagator are irrelevant implies that the fermionic self-energy has no frequency dependence either. This seems strange in light of naive $1/N$ expansions, where the Landau damping term plays a prominent role and leads to a non-Fermi-liquid form of the fermion self-energy. In any case, simple $1/N$ expansions are known to break down for models of metallic quantum critical points [18,39,40]. Since no frequency-dependent terms renormalize the boson propagator at the one-loop level, Landau damping effects are less important within this controlled epsilon expansion scheme and only appear at two-loop order. Note that this is a crucial difference from the Ising-nematic QCP studied in Ref. [13], where the Landau damping term had to be included in the boson propagator to cure an IR divergence in the fermion self-energy, despite the fact that it does not renormalize at one-loop order. By contrast, in the problem studied here an analogous IR divergence is cured by the ak_{d-1} term in the boson propagator, as can be seen directly from the a dependence in Eq. (17).

Finally, we note that there is no one-loop vertex correction in our theory, because one simply cannot draw a one-loop vertex diagram given the structure of the interaction term in Eq. (11).

IV. RENORMALIZATION

A. Fixed points

We now use the minimal subtraction scheme to derive RG flow equations for all dimensionless parameters in Eq. (11). To make our theory UV-finite, we need to include counterterms in the action, which subtract the divergent terms in the limit $\epsilon \to 0$. These correspond to the $\sim 1/\epsilon$ terms in the one-loop diagrams evaluated above. Since we used the convention $\overline{\Psi}(G_0^{-1} - \Sigma)\Psi$ for the definition of the self-energy Σ, we need to add the divergent part of self-energy to cancel the $1/\epsilon$ poles (the same holds for the bosonic self-energy Π). Therefore, the renormalized action reads

$$S_{\text{ren}} = S + S_{ct}$$
$$= \sum_j \int_{k^{d+1}} \overline{\Psi}_j(k) \left[-i\mathbf{\Gamma} \cdot \mathbf{K} + i\sigma_x k_{d-1} Z_2 + i\sigma_x k_d^2 Z_3 \right] \Psi_j(k)$$
$$+ \int_{k^{d+1}} \phi^+(k) \left[k_d^2 Z_4 + ak_{d-1} Z_5 \right] \phi^-(-k) + S_{\text{int}}, \qquad (18)$$

where we defined $Z_i = 1 + \frac{Z_{i,1}}{\epsilon}$ and

$$Z_{2,1} = -\frac{2u_1 \lambda^2}{(1-a)\sqrt{|\tilde{a}|N}}, \quad Z_{3,1} = \frac{2u_1 \lambda^2}{(1-a)^2 \sqrt{|\tilde{a}|N}},$$

$$Z_{4,1} = -\frac{u_1 \lambda^2}{2}, \quad Z_{5,1} = -\frac{u_1 \lambda^2}{a}. \qquad (19)$$

Introducing the rescaled bare fields

$$\mathbf{K} = \mathbf{K}_B, \quad k_{d-1} = Z_2^{-1} k_{B,d-1}, \quad k_d = Z_3^{-\frac{1}{2}} k_{B,d},$$

$$\Psi(k) = Z_2^{\frac{1}{2}} Z_3^{\frac{1}{4}} \Psi_B(k_B), \quad \phi^\pm(k) = Z_2^{\frac{1}{2}} Z_3^{\frac{3}{4}} Z_4^{-1} \phi_B^\pm(k_B),$$

$$\lambda_B = \lambda \mu^{\frac{\epsilon}{2}} Z_2^{-\frac{1}{2}} Z_3^{\frac{1}{4}} Z_4^{-\frac{1}{2}}, \quad a = Z_2 Z_3^{-1} Z_4 Z_5^{-1} a_B \qquad (20)$$

brings the renormalized action back to its initial (bare) form in Eq. (11). The one-loop β-functions for the couplings λ and a follow straightforwardly from Eqs. (20) and take the form

$$\beta_\lambda = \mu \frac{d\lambda}{d\mu} = \frac{u_1 \lambda^3}{2} \left(\frac{3-2a}{(1-a)^2 \sqrt{|\tilde{a}|N}} + \frac{1}{2} \right) - \frac{\epsilon}{2}\lambda,$$

$$\beta_a = \mu \frac{da}{d\mu} = u_1 \lambda^2 \left(\frac{2a(2-a)}{(1-a)^2 \sqrt{|\tilde{a}|N}} + \frac{a}{2} - 1 \right). \qquad (21)$$

For the physical case $N = 2$ these β-functions describe three scale-invariant fixed points at

$$(\lambda_1^*, a_1^*) = (4.335\sqrt{\epsilon}, 0.152),$$
$$(\lambda_2^*, a_2^*) = (20.43\sqrt{\epsilon}, 3.383), \qquad (22)$$
$$(\lambda_3^*, a_3^*) = (25.137\sqrt{\epsilon}, 2.0),$$

where the first and second fixed points are stable and the third one is unstable. The line $a = 2$ separates the two domains of attraction of the two stable fixed points. A corresponding flow diagram is shown in Fig. 2. Note that the β-functions are singular at $a = 0, 1$, but the differential equation for the RG flow trajectory $\frac{da}{d\lambda}$ is regular at these points, giving rise to continuous solutions of the flow equations.

For the problem of interest here, namely a generic Fermi surface with two hot spots connected by a $2k_F$ wave vector, a physically sensible UV initial condition for the RG flow corresponds to a positive coupling λ as well as an infinitesimally small value of a, such that the density susceptibility is initially peaked at the $\mathbf{Q} = \pm 2k_F$ wave vector. Consequently, the RG flow is directed toward the first fixed point (λ_1^*, a_1^*), which we identify with the continuous quantum phase transition between an ordinary Fermi liquid metal and the incommensurate $2k_F$ charge-density-wave phase. Note that an initial condition with $a = 2$ would correspond to a perfectly circular Fermi surface, where the density susceptibility has degenerate maxima along the entire $2k_F$ line defined by $\pm 2k_x + k_y^2 = 0$. In this highly fine-tuned case, fermions along the entire Fermi surface can scatter resonantly, not just at the two hot spots. It is important to realize that this does not invalidate the hot-spot theory, however, because the scattering is local in momentum space and one obtains a theory with an infinite set of decoupled hot-spot pairs. This situation is similar to the Ising-nematic problem and to the quantum phase transition between a normal metal and a FFLO superconductor at vanishing velocity detuning studied in Ref. [41]. Interestingly, the Fermi surface retains its shape, and no dynamical nesting
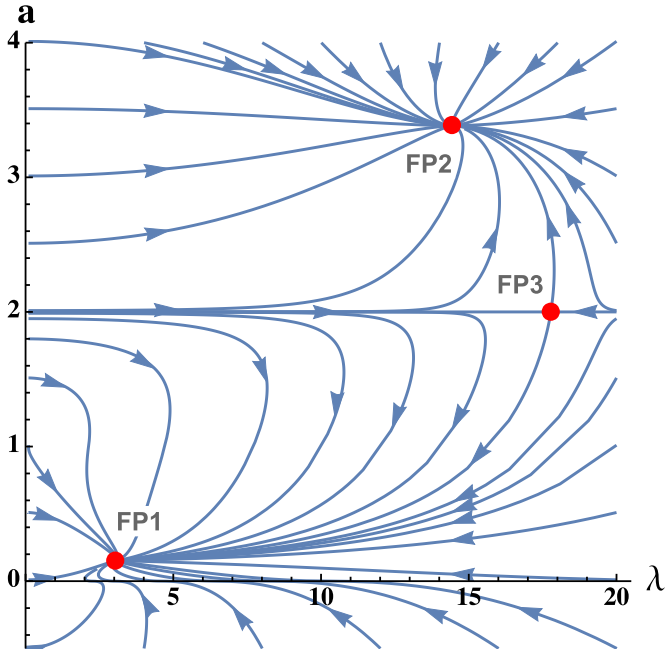
a

FIG. 2. RG flow in the λ-a plane for $\epsilon = 1/2$ and $N = 2$. The three fixed points are marked by red dots, and their coordinates are given in Eq. (22). We identify the stable fixed point FP1 with the $2k_F$ CDW quantum critical point of a metal with a generic Fermi surface, where two hot spots are connected by an incommensurate $2k_F$ wave vector.

occurs during the flow along the $a = 2$ line to the third fixed point (λ_3^*, a_3^*). This is in stark contrast to the flow toward the first stable fixed point, where a strong dynamical nesting of the Fermi surface at the two hot spots occurs during the RG flow, as discussed in the next section. Finally, we do not identify the second stable fixed point (λ_2^*, a_2^*) with a physically meaningful situation. A UV initial condition with an arbitrary value of a different from 0 or 2 would correspond to a density susceptibility with degenerate maxima that do not correspond to a $2k_F$ line.

B. Correlators

In the following, we discuss the general scaling form of the correlation functions,

$$
\begin{aligned}
&\langle \Psi(k_1) \cdots \Psi(k_m) \overline{\Psi}(k_{m+1}) \cdots \overline{\Psi}(k_{2m}) \\
&\phi^+(k_{2m+1}) \cdots \phi^+(k_{2m+n}) \phi^-(k_{2m+n+1}) \cdots \phi^-(k_{2m+2n}) \rangle \\
&= G^{(m,m,n,n)}(\{k_i\}, \mu, \lambda, a) \delta^{(d+1)}(\{k_i\}),
\end{aligned}
\tag{23}
$$

where $\delta^{(d+1)}(\{k_i\})$ ensures energy and momentum conservation. We define the dynamical critical exponents,

$$
z_{d-1}^{-1} = 1 + \frac{d \ln Z_2}{d \ln \mu} = 1 + \frac{2u_1 \lambda^2}{(1-a)\sqrt{|\tilde{a}|}N},
$$

$$
z_d^{-1} = 1 + \frac{d \ln Z_3}{d \ln \mu} = 1 - \frac{2u_1 \lambda^2}{(1-a)^2 \sqrt{|\tilde{a}|}N}, \tag{24}
$$

as well as the anomalous dimensions of the fermion and boson fields,

$$
\eta_\Psi = \frac{1}{2} \frac{d \ln Z_\Psi}{d \ln \mu} = \frac{u_1 \lambda^2}{2} \frac{2a-1}{(1-a)^2 \sqrt{|\tilde{a}|}N},
$$

$$
\eta_\phi = \frac{1}{2} \frac{d \ln Z_\phi}{d \ln \mu} = \frac{u_1 \lambda^2}{2} \left(\frac{2a+1}{(1-a)^2 \sqrt{|\tilde{a}|}N} + \frac{1}{2} \right), \tag{25}
$$

where $Z_\Psi = Z_2^{-1} Z_3^{-1/2}$ and $Z_\phi = Z_2^{-1} Z_3^{-3/2} Z_4$. Using these quantities, the renormalization-group equation for the correlation functions takes the form

$$
\begin{aligned}
&\left[\sum_{i=1}^{2m+2n} \left(\mathbf{K}_i \nabla_{\mathbf{K}_i} + \frac{k_{d-1,i}}{z_{d-1}} \frac{\partial}{\partial k_{d-1,i}} + \frac{k_{d,i}}{2z_d} \frac{\partial}{\partial k_{d,i}} \right) - \beta_\lambda \frac{\partial}{\partial \lambda} \right. \\
&- \beta_a \frac{\partial}{\partial a} - 2m \left(\eta_\Psi - \frac{4-\epsilon}{2} \right) - 2n \left(\eta_\phi - \frac{4-\epsilon}{2} \right) \\
&\left. + \left(\epsilon - \frac{3}{2} - \frac{1}{z_{d-1}} - \frac{1}{2z_d} \right) \right] G^{(m,m,n,n)}(\{k_i\}, \mu, \lambda, a) = 0.
\end{aligned}
\tag{26}
$$

At the fixed points, where the β-functions are zero, the solution of the RG equation for the fermion and boson two-point functions gives rise to the scaling forms

$$
G(k) = \frac{1}{|k_d|^{2z_d}} f_\Psi \left(\frac{|\mathbf{K}|}{|k_d|^{2z_d}}, \frac{\mathrm{sgn}(k_{d-1})|k_{d-1}|^{z_{d-1}}}{|k_d|^{2z_d}} \right), \tag{27}
$$

$$
D_+(k) = \frac{1}{|k_d|^{2pz_d}} f_\phi \left(\frac{|\mathbf{K}|}{|k_d|^{2z_d}}, \frac{\mathrm{sgn}(k_{d-1})|k_{d-1}|^{z_{d-1}}}{|k_d|^{2z_d}} \right), \tag{28}
$$

with

$$
p = \frac{1}{z_d} - \frac{u_1(\lambda^*)^2}{2}, \tag{29}
$$

and f_Ψ and f_ϕ are universal scaling functions. From the scaling form of the fermion propagator, we can infer the shape of the renormalized Fermi surface at the hot spots. In the noninteracting case, the fermion propagator has poles at the Fermi surface defined by $f_\Psi^{-1}(0, \pm 1) = 0$. Analogously, the renormalized shape of the Fermi surface is then determined by the equation

$$
\mathrm{sgn}(k_{d-1})|k_{d-1}|^{z_{d-1}} = \pm|k_d|^{2z_d}. \tag{30}
$$

For the physical case $N = 2$, the first fixed point (λ_1^*, a_1^*) is characterized by the dynamical critical exponents

$$
\left(z_{d-1}^{-1}, z_d^{-1} \right)_1^* = \left(1 + 0.566\epsilon, 1 - \frac{2}{3}\epsilon \right) \tag{31}
$$

and anomalous dimensions of the fermion and boson fields given by

$$
(\eta_\Psi, \eta_\phi)_1^* = (-0.116\epsilon, 0.268\epsilon). \tag{32}
$$

At this fixed point, the Fermi surface in the vicinity of the two hot spots takes the form $k_x = \pm|k_y|^{3.85}$ in $d = 2$ dimensions, which indicates a strong dynamical nesting with a vanishing Fermi surface curvature at the hot spots, as shown in Fig. 3. By contrast, for $a = 2$ the two critical exponents z_d and z_{d-1} in Eq. (24) are equal and thus the Fermi surface retains its shape for the RG flow along the $a = 2$ line. Note that all fixed
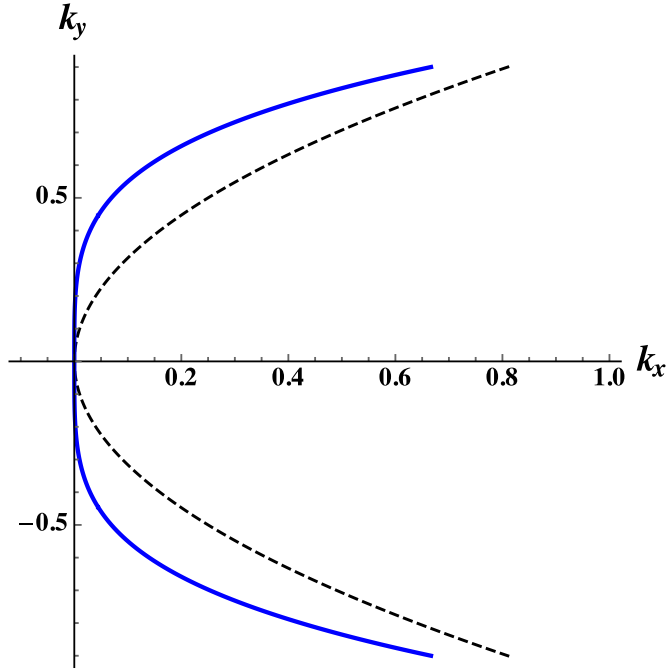
FIG. 3. Dynamical nesting of the Fermi surface. The blue solid line shows the form of the renormalized, flattened Fermi surface at one hot spot, as determined from Eq. (30) at the first RG fixed point (λ_1^*, a_1^*) for $N = 2$ and $\epsilon = 1/2$, i.e., in $d = 2$ dimensions. The black dashed line indicates the initial, parabolic form of the Fermi surface as a reference.

points describe non-Fermi-liquids, with a nonlinear, power-law fermion dispersion perpendicular (tangential) to the Fermi surface determined by the dynamical critical exponent z_{d-1} $(2z_d)$.

V. EXPERIMENTAL SIGNATURES

While there is strong evidence that the incommensurate CDW transition in rare-earth tellurides is driven by Fermi surface nesting [33,42], the situation in the transition-metal dichalcogenides NbSe$_2$ and TaS$_2$ is much less clear. Fermi surface nesting [35,43], saddle bands [44], as well as electron-phonon coupling [45,46] have been suggested as possible CDW mechanisms. Here we argue that it would be beneficial to study these materials in the vicinity of the quantum phase transition to the density wave ordered phase. Experimentally, the quantum phase transition can be driven by applying pressure [35,36] or intercalating different transition metals [38]. As we have laid out in this work, a quantum phase transition driven by Fermi surface nesting has definite signatures, which can be detected in the quantum critical regime at finite temperature above the quantum critical point. The most striking experimental signature of CDW quantum criticality would be the power-law behavior of the density susceptibility in the vicinity of the critical point. At the $2k_F$ wave vector **Q**, its characteristic power-law frequency dependence follows from Eq. (28) and has the form $D(\omega) \sim |\omega|^{-p}$ with the exponent $p = 1 - 0.769\epsilon \simeq 0.616$ for the physical case $N = 2$ and in $d = 2$ dimensions, which should be observable with a variety of experimental probes, such as Raman- or inelastic x-ray and

neutron scattering. The same signatures should be observable in rare-earth tellurides such as SmTe$_3$ and TbTe$_3$.

Moreover, in the quantum critical regime we expect ω/T scaling, i.e., the temperature T and frequency ω dependent density susceptibility at the $2k_F$ wave vector should obey the scaling relation

$$D(\omega, T) = |\omega|^{-p} f_T(T/\omega) \qquad (33)$$

with a universal scaling function f_T. Accordingly, the static density susceptibility at the $2k_F$ wave vector has a power-law temperature dependence $D(0) \sim T^{-p}$ with the same exponent p. Again, this characteristic temperature dependence should be observable using Raman- or elastic x-ray and neutron scattering.

Thermodynamic signatures are a bit more difficult to discern, unfortunately. The specific heat has singular contributions from "hot" fermions in the vicinity of the hot spots, as well as an ordinary linear in temperature contribution from "cold" fermions far away from the hot spots. The latter contribution dominates, however, which can be seen as follows: in analogy to the spin-density-wave critical point, we expect hyperscaling to be obeyed in the CDW hot-spot theory discussed here [47], consequently the scaling dimension of the free-energy density is given by $[f] = (d - 1) + z_{d-1}^{-1} + (2z_d)^{-1}$, where the first term comes from the scaling dimension of the timelike and the extra-dimensional directions, whereas the last two terms arise from the scaling dimension of the two spatial directions. The temperature dependence of the specific heat thus takes the form

$$C_v \sim T^{d-2+z_{d-1}^{-1}+(2z_d)^{-1}}. \qquad (34)$$

In two spatial dimensions, the specific-heat exponent is larger than 1 at the first fixed point, and thus the singular contribution is subleading compared to the $C_v \sim T$ contribution from cold fermions.

An alternative way to compute this contribution is via the scaling form of the electron propagator in Eq. (27), from which the electronic density of states $\nu(\omega)$ of the hot electrons can be computed. In $d = 2$ spatial dimensions, it takes the form

$$\nu(\omega) - \nu(0) \sim (1 - z_{d-1}) |\omega|^{z_{d-1}^{-1}+(2z_d)^{-1}-1}. \qquad (35)$$

The singular part of $\nu(\omega)$ gives rise to the same temperature dependence of C_v as determined above. Note, however, that the nonzero constant density of states $\nu(0)$ again leads to a dominant linear contribution $C_v \sim T$.

VI. SUPERCONDUCTING INSTABILITIES

Superconductivity has been observed in the vicinity of various metallic quantum critical points. For the $2k_F$ CDW critical point discussed here, a natural superconducting instability corresponds to the formation of spin-singlet Cooper pairs between electrons at the two antipodal hot spots. To investigate whether such a superconducting instability is enhanced or suppressed in the vicinity of the QCP, we compute the scaling dimension of the singlet Cooper-pair creation operator at the fixed point by including a corresponding source term in the

action:

$$S_{cp} = g \int_k [\psi_{+,\uparrow}(k)\psi_{-,\downarrow}(-k) - \psi_{+,\downarrow}(k)\psi_{-,\uparrow}(-k) + \text{c.c.}]. \tag{36}$$

In spinor representation and in general dimensions, this term can be written as

$$S_{cp} = g \int_{k^{d+1}} \tau_y^{\alpha\beta} \overline{\Psi}_\alpha(k)\Psi_\beta(k), \tag{37}$$

where the Greek indices are spin indices and $\tau_y^{\alpha\beta}$ is a Pauli matrix in spin space. The vertex factor and the one-loop correction read $-g\tau_y^{\alpha\beta}\mathbb{1} + g\tau_y^{\alpha\beta}V$, where V is a matrix in spinor space given by the integral

$$V = \frac{\lambda^2 \mu^\epsilon}{N} \int_p \sigma_y G^2(p) \sigma_y D_+(p+q). \tag{38}$$

Calculating V (see Appendix C) leads to the $1/\epsilon$ pole

$$V = -\mathbb{1}\frac{4u_1\lambda^2}{(1-a)\sqrt{|\tilde{a}|}N\epsilon} + \text{finite terms}. \tag{39}$$

The renormalized action, which we get by adding V, is given by

$$S_{cp}^r = g\mu Z_g \int_{k^{d+1}} \tau_y^{\alpha\beta} \overline{\Psi}_\alpha(k)\Psi_\beta(k), \tag{40}$$

where we introduced an arbitrary mass scale μ to make the coupling dimensionless and $Z_g = 1 + \frac{Z_{g,1}}{\epsilon}$ with $Z_{g,1} = -\frac{4u_1\lambda^2}{(1-a)\sqrt{|\tilde{a}|}N}$. The β-function for the source field g reads

$$\beta_g = g(-1 - \eta_g) \tag{41}$$

with the anomalous dimension

$$\eta_g = \frac{d \ln Z_g}{d \ln \mu} = \frac{4u_1\lambda^2}{(1-a)\sqrt{|\tilde{a}|}N}. \tag{42}$$

The first fixed point is the only one with $a < 1$ and therefore $\eta_g = 1.131\epsilon > 0$ for $N = 2$. Consequently, superconducting instabilities are enhanced at the first fixed point. At the other two fixed points, $\eta_g < 0$ and superconductivity is suppressed.

VII. DISCUSSION AND CONCLUSIONS

We have presented a controlled, perturbative renormalization-group study of incommensurate $2k_F$ CDW quantum critical points in two-dimensional metals, which treats electronic as well as bosonic order-parameter fluctuations on an equal footing, and we found a stable fixed point corresponding to a continuous quantum phase transition with a strongly renormalized, flattened Fermi surface at the hot spots. This result is in contrast to the early theoretical observation of a first-order incommensurate transition due to strong fluctuations in Ref. [23], which did not take the crucial dynamical nesting of the Fermi surface into account, however. Indeed, Sykora *et al.* pointed out in Ref. [28] that the Fermi surface is strongly renormalized and flattened at the two hot spots connected by the $2k_F$ wave vector, and our results strengthen their observation. Moreover, we have presented experimentally testable predictions for the density susceptibility in the vicinity of the quantum critical point,

which could help to clarify the controversially debated CDW mechanism in the transition-metal dichalcogenides $NbSe_2$ and TaS_2.

One important point that we have not discussed so far is how to tune our model away from criticality by adding a boson mass term to the action. Naively, adding a boson mass in Eq. (11) would turn the boson massless at a momentum different from $k = 0$ rather than leading to a gap, due to the linear term $\sim ak_{d-1}$ in the boson propagator. This seemingly invalidates our assumption that the density susceptibility is peaked at the $2k_F$ momentum $\mathbf{Q}$, i.e., at $k = 0$. It is crucial to realize, however, that the boson propagator strongly renormalizes during the RG flow, and the dynamical nesting of the Fermi surface ensures that the boson self-energy remains peaked at $k = 0$. Perturbing the fixed-point action with a mass term thus gaps out the boson at all momenta. A simple way to see this is as follows: the detailed form of the one-loop boson self-energy in Eqs. (A7) and (A8) computed with the bare fermion propagators shows that it is peaked at $k = 0$ in dimensions $d < 2$, but it does not have a peak at $k = 0$ for $d \geqslant 2$. This behavior is reminiscent of the Lindhard density susceptibility of a free Fermi gas. By contrast, computing the boson self-energy using a renormalized, nested fermion dispersion with vanishing Fermi surface curvature at the two hot spots of the form $\pm k_{d-1} + |k_d|^\alpha$ with $\alpha > 2$ leads to a well-defined peak in the $d = 2$ density susceptibility at the $2k_F$ wave vector $\mathbf{Q}$, i.e., at $k = 0$.

Another technical detail worth discussing is that our computation differs crucially from an analogous approach to Ising-nematic quantum criticality in one aspect: while Ref. [13] had to reorganize the perturbation expansion by including the Landau damping term in the boson propagator from the start to remove IR divergences, this is not necessary for the CDW problem considered here, where similar IR divergences are cured by the linear $\sim k_{d-1}$ term in the boson propagator. As a consequence, Landau damping effects are subleading and only appear at two-loop order in the problem considered here. We leave the challenging computation of two-loop effects for future investigation.

APPENDIX A: BOSON SELF-ENERGY

The integral for the one-loop boson self-energy in Eq. (13) can be evaluated as follows: Using the properties of the Pauli matrices, the trace can be simplified using

$$\text{Tr}(\sigma_y \sigma_i \sigma_y \sigma_j) = \begin{cases} 2, & i = j = y, \\ -2, & i = j \neq y, \\ 0, & i \neq j, \end{cases} \tag{A1}$$

which yields

$$\Pi(k) = -\lambda^2 \mu^\epsilon \int_p \frac{\delta_p \delta_{k-p} - \mathbf{P} \cdot (\mathbf{P} - \mathbf{K})}{(\mathbf{P}^2 + \delta_p^2)[(\mathbf{P} - \mathbf{K})^2 + \delta_{k-p}^2]}. \tag{A2}$$

Shifting $p_{d-1} \to p_{d-1} - p_d^2$ and introducing the new integration variable $y = \frac{1}{\sqrt{2}}(2p_d - k_d)$ as well as $e_k = k_{d-1} + \frac{k_d^2}{2}$ leads to

$$\Pi(k) = \frac{\lambda^2 \mu^\epsilon}{\sqrt{2}} \int_{\mathbf{P}, p_{d-1}, y} \frac{p_{d-1}(p_{d-1} - e_k - y^2) + \mathbf{P} \cdot (\mathbf{P} - \mathbf{K})}{(\mathbf{P}^2 + p_{d-1}^2)[(\mathbf{P} - \mathbf{K})^2 + (p_{d-1} - e_k - y^2)^2]}. \tag{A3}$$

Using the Feynman parametrization, the above integral can be written as

$$\Pi(k) = \frac{\lambda^2 \mu^\epsilon}{\sqrt{2}} \int_0^1 dx \int \frac{d^{d-1}\mathbf{P} dp_{d-1} dy}{(2\pi)^{d+1}} \frac{p_{d-1}^2 - (e_k + y^2)p_{d-1} + \mathbf{P}^2 - x(1-x)\mathbf{K}^2}{[p_{d-1}^2 - 2x(e_k + y^2)p_{d-1} + x(e_k + y^2)^2 + \mathbf{P}^2 + x(1-x)\mathbf{K}^2]^2}. \tag{A4}$$

Integrating over p_{d-1} and then rescaling $\mathbf{P} \to \sqrt{x(1-x)}\mathbf{P}$ yields

$$\Pi(q) = \frac{\lambda^2 \mu^\epsilon}{\sqrt{8}} \int_0^1 dx \, [x(1-x)]^{\frac{d}{2}-1} \int \frac{d^{d-1}\mathbf{P} dy}{(2\pi)^d} \frac{\mathbf{P}^2}{[\mathbf{P}^2 + \mathbf{K}^2 + (e_k + y^2)^2]^{\frac{3}{2}}}. \tag{A5}$$

The x-integral is elementary, and after switching to hyperspherical coordinates, subtracting $\Pi(0)$ for UV regularization, and performing the integral over the radial coordinate, we obtain

$$\Pi(q) - \Pi(0) = \lambda^2 \mu^\epsilon \frac{\Gamma^2(\frac{d}{2})\Gamma(\frac{d+1}{2})\Gamma(1 - \frac{d}{2})}{2^{d-\frac{1}{2}}\pi^{\frac{d}{2}+1}\Gamma(d)\Gamma(\frac{d-1}{2})} \int_0^\infty dy \big[[\mathbf{K}^2 + (e_k + y^2)^2]^{\frac{d}{2}-1} - (y^4)^{\frac{d}{2}-1}\big]. \tag{A6}$$

Since the integral has different solutions for $e_k > 0$ and $e_k < 0$, we need to distinguish the two cases. After setting $d = \frac{5}{2} - \epsilon$, the self-energy for $e_k > 0$ reads

$$\Pi(q) - \Pi(0) = \lambda^2 \mu^\epsilon \frac{\Gamma^2(\frac{5}{4} - \frac{\epsilon}{2})\Gamma(-\frac{1}{4} + \frac{\epsilon}{2})\Gamma(-1 + \epsilon)}{2^{4-\epsilon}\pi^{\frac{7}{4}-\frac{\epsilon}{2}}\Gamma(\frac{5}{2} - \epsilon)\Gamma(-\frac{1}{2} + \epsilon)} \left(\frac{3}{2} - \epsilon\right) e_k^{1-\epsilon} \, {}_2F_1\left[\frac{\epsilon-1}{2}, \frac{\epsilon}{2}, \frac{2\epsilon+1}{4}, -\frac{\mathbf{K}^2}{e_k^2}\right], \tag{A7}$$

which gives the $1/\epsilon$ pole of Eq. (14) when expanding the hypergeometric function ${}_2F_1$ around $\epsilon = 0$. In the case of $e_k < 0$, the solution to the y integration is

$$\Pi(q) - \Pi(0) = \lambda^2 \mu^\epsilon \frac{\Gamma^2(\frac{5}{4} - \frac{\epsilon}{2})(\frac{3}{2} - \epsilon)}{2^{4-\epsilon}\pi^{\frac{9}{4}-\frac{\epsilon}{2}}\Gamma(\frac{5}{2} - \epsilon)} |\mathbf{K}|^{-\epsilon} \left(2|\mathbf{K}|\Gamma\left(\frac{5}{4}\right)\Gamma\left(\frac{\epsilon-1}{2}\right){}_2F_1\left[\frac{1}{4}, \frac{\epsilon-1}{2}, \frac{1}{2}, -\frac{e_k^2}{\mathbf{K}^2}\right]\right.$$

$$\left. - e_k \Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{\epsilon}{2}\right){}_2F_1\left[\frac{3}{4}, \frac{\epsilon}{2}, \frac{3}{2}, -\frac{e_k^2}{\mathbf{K}^2}\right]\right), \tag{A8}$$

which leads to the same $1/\epsilon$ pole as the solution for $e_k > 0$ above. Expanding this expression in $|\mathbf{K}|$ and afterward around $\epsilon = 0$ yields the expression in (14). Note that doing the same with (A7), the lowest-order term in $|\mathbf{K}|$ is $\propto \mathbf{K}^2$.

APPENDIX B: FERMION SELF-ENERGY

The product of Pauli matrices in Eq. (16) can be simplified to

$$\sigma_y G^T(k)\sigma_y = -i\sigma_y \frac{k_0 \sigma_y - \mathbf{k} \cdot \sigma_z + \delta_k \sigma_x}{\mathbf{K}^2 + \delta_k^2} \sigma_y = -i \frac{k_0 \sigma_y + \mathbf{k} \cdot \sigma_z - \delta_k \sigma_x}{\mathbf{K}^2 + \delta_k^2} = -G(k) \tag{B1}$$

and therefore

$$\Sigma(k) = -\frac{\lambda^2 \mu^\epsilon}{N} \int_p G(p-k) D_+(p) = \frac{i\lambda^2 \mu^\epsilon}{N} \int_p \frac{-\boldsymbol{\Gamma} \cdot (\mathbf{P} - \mathbf{K}) + \sigma_x \delta_{p-k}}{(\mathbf{P} - \mathbf{K})^2 + \delta_{p-k}^2} \frac{1}{p_d^2 + a p_{d-1}}. \tag{B2}$$

Shifting $\mathbf{P} \to \mathbf{P} + \mathbf{K}$, $p_{d-1} \to p_{d-1} - p_d^2 + 2k_d p_d$, and defining $\tilde{a} = \frac{a}{1-a}$ leads to

$$\Sigma(k) = \frac{i\lambda^2 \mu^\epsilon \sigma_x}{(1-a)N} \int_p \frac{p_{d-1} + \delta_{-k}}{\mathbf{P}^2 + (p_{d-1} + \delta_{-k})^2} \frac{1}{p_d^2 + 2\tilde{a}k_d p_d + \tilde{a}p_{d-1}}, \tag{B3}$$

where $\delta_{-k} = -k_{d-1} + k_d^2$. The p_d integral can be evaluated using the principal value, which leads to

$$\Sigma(k) = \frac{i\lambda^2 \mu^\epsilon \sigma_x}{2(1-a)\sqrt{|\tilde{a}|}N} \int \frac{d^{d-1}\mathbf{P} dp_{d-1}}{(2\pi)^d} \frac{p_{d-1} + \delta_{-k}}{\mathbf{P}^2 + (p_{d-1} + \delta_{-k})^2} \frac{\Theta[\text{sgn}(\tilde{a})p_{d-1} - |\tilde{a}|k_d^2]}{\sqrt{\text{sgn}(\tilde{a})p_{d-1} - |\tilde{a}|k_d^2}}. \tag{B4}$$

After the substitution $y = \text{sgn}(\tilde{a})p_{d-1} - |\tilde{a}|k_d^2$, the fermion self-energy reads

$$\Sigma(q) = \frac{i\lambda^2\mu^\epsilon\sigma_x\text{sgn}(\tilde{a})}{2(1-a)\sqrt{|\tilde{a}|}N}\int\frac{d^{d-1}\mathbf{P}}{(2\pi)^{d-1}}\int_0^\infty\frac{dy}{2\pi}\frac{y+c(k)}{\mathbf{P}^2+[y+c(k)]^2}\frac{1}{\sqrt{y}}, \tag{B5}$$

where we defined $c(k) = |\tilde{a}|k_d^2 + \text{sgn}(\tilde{a})\delta_{-k}$. The y integration is elementary, and using hyperspherical coordinates we get

$$\Sigma(k) - \Sigma(0) = \frac{i\lambda^2\mu^\epsilon\sigma_x\text{sgn}(\tilde{a})}{(1-a)\sqrt{|\tilde{a}|}N}\frac{1}{2^d\pi^{\frac{d-1}{2}}\Gamma\left(\frac{d-1}{2}\right)}\text{Re}\left\{\int_0^\infty dr\, r^{d-2}\left(\frac{1}{\sqrt{c(k)+ir}} - \frac{1}{\sqrt{ir}}\right)\right\}, \tag{B6}$$

where we subtracted $\Sigma(0)$ for UV regularization since the above integral initially converges only for $d < \frac{3}{2}$. Carrying out the r integral gives the result

$$\Sigma(k) - \Sigma(0) = \frac{i\lambda^2\mu^\epsilon\sigma_x\text{sgn}(\tilde{a})}{(1-a)\sqrt{|\tilde{a}|}N}\frac{\Gamma\left(\frac{3}{2}-d\right)\Gamma(d-1)}{2^d\pi^{\frac{d}{2}}\Gamma\left(\frac{d-1}{2}\right)}\text{Re}\left\{i^{1-d}[c(k)]^{d-\frac{3}{2}}\right\}. \tag{B7}$$

Setting $d = \frac{5}{2} - \epsilon$ and expanding around $\epsilon = 0$ gives the pole contribution in Eq. (17).

APPENDIX C: SUPERCONDUCTING VERTEX

The matrix product in the integral (38) can be evaluated to

$$\sigma_y G^2(p)\sigma_y = -\mathbb{1}\frac{p_0^2 + \sum_{i,j=1}^{d-2}p_ip_j + \delta_p^2}{\left(\mathbf{P}^2+\delta_p^2\right)^2} = -\mathbb{1}\frac{p_0^2 + \mathbf{p}^2 + \sum_{i\neq j}p_ip_j + \delta_p^2}{\left(\mathbf{P}^2+\delta_p^2\right)^2}. \tag{C1}$$

The term $\sum_{i\neq j}p_ip_j$ vanishes by antisymmetry under $p_i \to -p_i$ since the boson propagator is independent of p_i, and so we get after shifting $p_{d-1} \to p_{d-1} - p_d^2$

$$V = -\mathbb{1}\frac{\lambda^2\mu^\epsilon}{(1-a)N}\int_p\frac{1}{\mathbf{P}^2+p_{d-1}^2}\frac{1}{p_d^2+\frac{2q_d}{1-a}p_d+\frac{1}{1-a}q_d^2+\tilde{a}p_{d-1}+\tilde{a}q_{d-1}}. \tag{C2}$$

Here the p_d integral is straightforward using the principal value, and with the substitution $y = \text{sgn}(\tilde{a})p_{d-1} - f(q)$, where $f(q) = -\text{sgn}(\tilde{a})p_{d-1} + \text{sgn}(\tilde{a})\frac{1}{1-a}q_d^2$, the vertex correction reads

$$V = -\mathbb{1}\frac{\lambda^2\mu^\epsilon}{2(1-a)\sqrt{|\tilde{a}|}N}\int\frac{d^{d-1}\mathbf{P}}{(2\pi)^{d-1}}\int_0^\infty\frac{dy}{2\pi}\frac{1}{\mathbf{P}^2+[y+f(q)]^2}\frac{1}{\sqrt{y}}. \tag{C3}$$

Integrating over y and changing to hyperspherical coordinates yields

$$V = -\mathbb{1}\frac{\lambda^2\mu^\epsilon}{(1-a)\sqrt{|\tilde{a}|}N}\frac{i}{2^{d+1}\pi^{\frac{d-1}{2}}\Gamma\left(\frac{d-1}{2}\right)}\int_0^\infty dr\, r^{d-3}\left(\frac{1}{\sqrt{f(q)+ir}} - \frac{1}{\sqrt{f(q)-ir}}\right), \tag{C4}$$

which can be evaluated to

$$V = \mathbb{1}\frac{\lambda^2\mu^\epsilon}{(1-a)\sqrt{|\tilde{a}|}N}\frac{i\,\Gamma\left(\frac{5}{2}-d\right)\Gamma(d-2)[f(q)]^{\frac{5}{2}-d}}{2^{d+1}\pi^{\frac{d}{2}}\Gamma\left(\frac{d-1}{2}\right)}[i^{-d}-(-i)^{-d}]. \tag{C5}$$

Setting $d = \frac{5}{2} - \epsilon$ and expanding around $\epsilon = 0$ leads to the expression in (39).

---

[1] O. Gunnarsson, M. Calandra, and J. E. Han, Rev. Mod. Phys. **75**, 1085 (2003).

[2] V. J. Emery and S. A. Kivelson, Phys. Rev. Lett. **74**, 3253 (1995).

[3] N. E. Hussey, J. Phys: Condens. Matter **20**, 123201 (2008).

[4] S. Kasahara, T. Shibauchi, K. Hashimoto, K. Ikada, S. Tonegawa, R. Okazaki, H. Shishido, H. Ikeda, H. Takeya, K. Hirata, T. Terashima, and Y. Matsuda, Phys. Rev. B **81**, 184519 (2010).

[5] G. R. Stewart, Rev. Mod. Phys. **73**, 797 (2001).

[6] Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigorda, K. Watanabe, T. Taniguchi, T. Senthil, and P. Jarillo-Herrero, arXiv:1901.03710.

[7] H. V. Löhneysen, A. Rosch, M. Vojta, and P. Wölfle, Rev. Mod. Phys. **79**, 1015 (2007).

[8] E. Berg, M. Metlitski, and S. Sachdev, Science **338**, 1606 (2012).

[9] S. Lederer, Y. Schattner, E. Berg, and S. A. Kivelson, Phys. Rev. Lett. **114**, 097001 (2015).

[10] M. H. Gerlach, Y. Schattner, E. Berg, and S. Trebst, Phys. Rev. B **95**, 035124 (2017).

[11] E. Berg, S. Lederer, Y. Schattner, and S. Trebst, Annu. Rev. Condens. Matter Phys. **10**, 63 (2018).

[12] T. Senthil and R. Shankar, Phys. Rev. Lett. **102**, 046406 (2009).

[13] D. Dalidovich and S.-S. Lee, Phys. Rev. B **88**, 245106 (2013).

[14] S. Sur and S.-S. Lee, Phys. Rev. B **91**, 125136 (2015).

[15] S.-S. Lee, Annu. Rev. Condens. Matter Phys. **9**, 227 (2018).

[16] A. Abanov and A. V. Chubukov, Phys. Rev. Lett. **84**, 5608 (2000).

[17] A. Abanov, A. V. Chubukov, and J. Schmalian, Adv. Phys. **52**, 119 (2003).

[18] M. A. Metlitski and S. Sachdev, Phys. Rev. B **82**, 075128 (2010).

[19] H. Meier, C. Pépin, M. Einenkel, and K. B. Efetov, Phys. Rev. B **89**, 195115 (2014).

[20] A. A. Patel and S. Sachdev, Phys. Rev. B **90**, 165146 (2014).

[21] S. A. Maier and P. Strack, Phys. Rev. B **93**, 165114 (2016).

[22] A. Schlief, P. Lunts, and S.-S. Lee, Phys. Rev. X **7**, 021010 (2017).

[23] B. L. Altshuler, L. B. Ioffe, and A. J. Millis, Phys. Rev. B **52**, 5563 (1995).

[24] D. Bergeron, D. Chowdhury, M. Punk, S. Sachdev, and A.-M. S. Tremblay, Phys. Rev. B **86**, 155123 (2012).

[25] T. Holder and W. Metzner, Phys. Rev. B **90**, 161106(R) (2014).

[26] Y. Wang and A. Chubukov, Phys. Rev. B **88**, 024516 (2013).

[27] M. Punk, Phys. Rev. B **91**, 115131 (2015).

[28] J. Sykora, T. Holder, and W. Metzner, Phys. Rev. B **97**, 155159 (2018).

[29] I. Jang, G. Duvjir, B. K. Choi, J. Kim, Y. J. Chang, and K.-S. Kim, Phys. Rev. B **99**, 014106 (2019).

[30] J. A. Wilson, F. J. Di Salvo, and S. Mahajan, Adv. Phys. **24**, 117 (1975).

[31] G. A. Scholz, O. Singh, R. F. Frindt, and A. E. Curzon, Solid State Commun. **44**, 1455 (1982).

[32] P. Chen, W. W. Pai, Y.-H. Chan, V. Madhavan, M. Y. Chou, S.-K. Mo, A.-V. Fedorov, and T.-C. Chiang, Phys. Rev. Lett. **121**, 196402 (2018).

[33] G.-H. Gweon, J. D. Denlinger, J. A. Clack, J. W. Allen, C. G. Olson, E. D. DiMasi, M. C. Aronson, B. Foran, and S. Lee, Phys. Rev. Lett. **81**, 886 (1998).

[34] A. Fang, N. Ru, I. R. Fisher, and A. Kapitulnik, Phys. Rev. Lett. **99**, 046401 (2007).

[35] Y. Feng, J. V. Wezel, J. Wang, F. Flicker, D. M. Silevitch, P. B. Littlewood, and T. F. Rosenbaum, Nat. Phys. **11**, 865 (2015).

[36] Y. Feng, J. Wang, R. Jaramillo, J. v. Wezel, S. Haravifard, G. Srajer, Y. Liu, Z.-A. Xu, P. B. Littlewood, and T. F. Rosenbaum, Proc. Natl. Acad. Sci. USA **109**, 7224 (2012).

[37] J. B. He, P. P. Wang, H. X. Yang, Y. J. Long, L. X. Zhao, C. Ma, M. Yang, D. M. Wang, X. C. Shangguan, M. Q. Xue, P. Zhang, Z. A. Ren, J. Q. Li, W. M. Liu, and G. F. Chen, Supercond. Sci. Technol. **29**, 065018 (2016).

[38] K. E. Wagner, E. Morosan, Y. S. Hor, J. Tao, Y. Zhu, T. Sanders, T. M. McQueen, H. W. Zandbergen, A. J. Williams, D. V. West, and R. J. Cava, Phys. Rev. B **78**, 104520 (2008).

[39] S.-S. Lee, Phys. Rev. B **80**, 165102 (2009).

[40] M. A. Metlitski and S. Sachdev, Phys. Rev. B **82**, 075127 (2010).

[41] D. Pimenov, I. Mandal, F. Piazza, and M. Punk, Phys. Rev. B **98**, 024510 (2018).

[42] H. Yao, J. A. Robertson, E.-A. Kim, and S. A. Kivelson, Phys Rev. B **74**, 245126 (2006).

[43] Th. Straub, Th. Finteis, R. Claessen, P. Steiner, S. Hüfner, P. Blaha, C. S. Oglesby, and E. Bucher, Phys. Rev. Lett. **82**, 4504 (1999).

[44] T. M. Rice and G. K. Scott, Phys. Rev. Lett. **35**, 120 (1975).

[45] F. Weber, S. Rosenkranz, J.-P. Castellan, R. Osborn, R. Hott, R. Heid, K.-P. Bohnen, T. Egami, A. H. Said, and D. Reznik, Phys. Rev. Lett. **107**, 107403 (2011).

[46] F. Flicker and J. van Wezel, Nat. Commun. **6**, 7034 (2015).

[47] A. A. Patel, P. Strack, and S. Sachdev, Phys. Rev. B **92**, 165105 (2015).

4 Deconfined criticality in $SU(3)$ antiferromagnets on the triangular lattice

4.1 Introduction

4.1.1 Two-sided symmetry breaking

In our discussion of metallic criticality in the preceding section, we have seen that a description of the phase transitions solely in terms of an order parameter ϕ can fail. Still, the basic distinction between a symmetric phase with $\langle \phi \rangle = 0$ and a symmetry-broken one with $\langle \phi \rangle \neq 0$ has remained valid. But even this notion is not set in stone: as we will discuss in this chapter, sometimes a distinct symmetry can be broken in both phases – i.e., there is no relation $G' \subsetneq G$ between the symmetry groups of the two phases as in Sec. 2.

In the classical realm, such a situation is e.g. encountered in crystal structure transitions at high temperatures for a variety of materials (see Sec. IV.5 of [TT87]). As regards quantum phase transitions, the "two-sided symmetry breaking" frequently occurs in quantum magnets defined on a lattice. This comes about as follows: naively, one might assume that a generic quantum phase transition in a quantum magnet separates a state with magnetic order (spontaneously broken spin-rotation symmetry) from a featureless, fully symmetric paramagnet. But it turns out that the simple paramagnet is not an option for systems with a half-integer spin per unit cell. This is the core result of a theorem mainly due to Lieb, Schultz and Mattis [LSM61] (in 1D) and Oshikawa [Osh00] and Hastings [Has04] (generalization to higher dimensions). The precise statement is that, in an $SU(2)$-symmetric system of linear size L, the gap between the ground and the first excited state is bounded by $\log(L)/L$.[1] In the thermodynamic limit, this leaves two options: Either, the system is gapped and the ground state is degenerate (topological order or spontaneous breaking of a discrete symmetry). Alternatively, the system is gapless – a gapless spin-liquid or continuous symmetry breaking, along with a manifold of Goldstone modes. In any case, a featureless gapped paramagnet with a unique ground state is excluded. Here, we will focus on quantum phase transitions where a discrete symmetry is broken in the paramagnetic phase, i.e. we have two-sided symmetry breaking.

In principle, this scenario can still be captured within a LGW description, by introducing two order parameters ϕ_1, ϕ_2 coupled to each other. For example, the most simple Landau functional could read [TT87]

$$F_L(\phi_1, \phi_2) = r_1\phi_1^2 + u_1\phi_1^4 + r_2\phi_2^2 + u_2\phi_2^4 + \gamma\phi_1^2\phi_2^2 \,, \tag{4.1}$$

and r_1, r_2 change their signs at two distinct values of some parameter r driving the transition. Here, we are interested in the transition between two phases with $\langle \phi_1 \rangle \neq 0, \langle \phi_2 \rangle = 0$ (phase A) and $\langle \phi_2 \rangle \neq 0, \langle \phi_1 \rangle = 0$ (phase B). As seen from minimization of the functional F_L, this transition can either be first order (with a discontinuous jump of one order parameter, see Fig. 4.1(a)) or continuous, but via a phase of coexistence (Phase C, see Fig. 4.1(b)). By

---

1 While for integer-spin systems a finite gap can remain, as evidenced by the famous Haldane conjecture [Hal83] for a Heisenberg spin chain.

contrast, a true continuous phase transition where both order parameters vanish at the transition point (Fig. 4.1(c)) requires unnatural finetuning, e.g. $r_1 = r_2$ in the above.
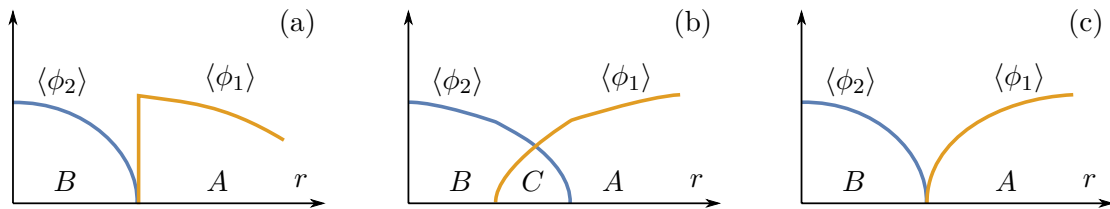

Figure 4.1 Two-order parameter description of a two-sided symmetry breaking transition. (a) Generic first order transition. (b) Generic continuous transition across coexistence phase (c) Fine-tuned continuous transition. For a different Landau functional, a disordered intermediate phase where both $\langle \phi_1 \rangle = \langle \phi_2 \rangle = 0$ would be possible as well.

Thus, a direct continuous transition between the two ordered phases seems unlikely from the LGW viewpoint. But yet again, the phenomenological order parameter description can fail, and a continuous transition is possible within the framework of "deconfined criticality", first introduced in 2004 by Senthil and coworkers [SVB$^+$04, SBS$^+$04].

4.1.2 Deconfined Criticality

For our short review of deconfined criticality, we will focus on its paradigmatic application, which is the phase transition between a magnetically ordered Néel state and a paramagnetic valence bond solid (VBS) state on the square lattice. To this end, we consider the antiferromagnetic spin-1/2 Heisenberg model on the square lattice,

$$H = J \sum_{\langle i,j \rangle} \hat{\boldsymbol{S}}_i \hat{\boldsymbol{S}}_j - \tilde{J} H_{\text{pert}}, \quad J, \tilde{J} > 0 \ , \tag{4.2}$$

with $\hat{\boldsymbol{S}}$ a vector of Pauli matrices times 1/2, and $\tilde{J} H_{\text{pert}}$ some appropriate $SU(2)$-invariant perturbation. For example, the Hamiltonian studied numerically in Ref. [San07] has $H_{\text{pert}} = \sum_{\langle ijkl \rangle} (\hat{\boldsymbol{S}}_i \hat{\boldsymbol{S}}_j - 1/4)(\hat{\boldsymbol{S}}_k \hat{\boldsymbol{S}}_l - 1/4)$, with $\langle ijkl \rangle$ referring to the corners of a plaquette. This term favours a specific singlet pattern formation (since the terms in parentheses are projectors on the singlet state). However, one should just view (4.2) as a single microscopic example, and all $SU(2)$-invariant Hamiltonians which realize the phases of interest should fall into the same universality class.

Different ground states of H can be accessed by tuning $r \equiv \tilde{J}/J$. For small r, the ground state has Néel order, characterized by a non-vanishing staggered magnetization $\boldsymbol{N} = 1/L^2 \sum_i (-1)^{\mathbf{r}_i} \langle \vec{S}_i \rangle$.[2] I.e., the $SU(2)$ symmetry of the Hamiltonian is spontaneously broken down to a $U(1)$ (and the translational invariance is broken as well). Increasing r, eventually a different ground state with VBS order will be realized – a trivial paramagnetic ground state is excluded by the Lieb-Schultz-Mattis theorem. There are, in principle, two VBS ordering patterns, a columnar and plaquette VBS order, but these entail the same critical properties at the phase transition [SBS$^+$04], and thus we focus on the columnar order here. A cartoon of the relevant ground states is shown in Fig. 4.2. As seen in the figure, in the VBS state lattice symmetries (translation and rotation) are broken. A suitable order

---

2 Of course, the simple Néel product state is obtained only in the classical limit ($S \to \infty$), but the true quantum ground state should show a staggered magnetization as well.

parameter is, e.g. the two-component dimer correlation [SBS$^+$04][3]

$$(\hat{\psi}_x, \hat{\psi}_y) = \left(\frac{1}{L^2}\sum_i(-1)^{x_i}\hat{\boldsymbol{S}}_i\hat{\boldsymbol{S}}_{i+\mathbf{e}_x}, \frac{1}{L^2}\sum_i(-1)^{y_i}\hat{\boldsymbol{S}}_i\hat{\boldsymbol{S}}_{i+\mathbf{e}_y}\right) . \tag{4.3}$$

In the VBS states of Fig. 4.2 it takes the values $(\pm 1/8, 0), (0, \pm 1/8)$, corresponding to a Z_4 "clock" order. This is visualized by the four orientations of an arrow in Fig. 4.2.
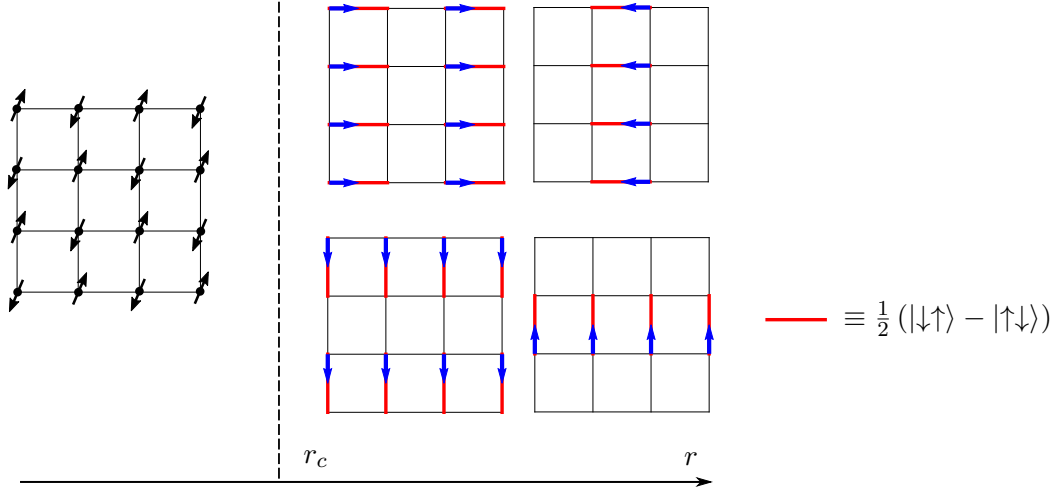

Figure 4.2 Néel to VBS transition, occuring at a critical value of the tuning parameter r. Fig. adapted from Ref. [SBS$^+$04].

We want to understand the phase transition between the two phases. Let us jump ahead, and state the key finding of the theory of deconfined criticality: in the continuum limit, the phase transition is not described by a coupled order parameter theory of the type (4.1), but rather by the so-called euclidean NCCP^1-model[4] with (2+1)-dimensional action

$$S_{\text{NCCP}^1} = \int d^3x\left[\sum_{\alpha,\mu}|(\partial_\mu - i\mathcal{A}_\mu)z_\alpha|^2 + rz_\alpha\bar{z}_\alpha + \rho(z_\alpha\bar{z}_\alpha)^2 + \frac{1}{4e^2}\mathcal{F}_{\mu\nu}^2\right] . \tag{4.4}$$

Here, $\mathbf{z}(\tau, \mathbf{x}) = (z_1(\tau, \mathbf{x}), z_2(\tau, \mathbf{x}))$ is a two-component complex spinor (i.e., it transforms under the fundamental representation of $SU(2)$), and represents a **spinon**, i.e. a spin-1/2 excitation of the VBS order (see below). One can furthermore identify the (fluctuating) Néel order parameter with the spinons $\mathbf{z}$ as $\boldsymbol{N} = \bar{z}_\alpha\boldsymbol{\sigma}_{\alpha\beta}z_\beta$, with $\boldsymbol{\sigma}$ the vector of Pauli matrices. The spinons are therefore "fractionalized" degrees of freedom. They are minimally coupled to a non-compact real $U(1)$ gauge field $\mathcal{A}_\mu(\tau, \mathbf{x})$, with $\mu \in (\tau, x, y)$, and kinematics determined by a standard Maxwell term $\mathcal{F}_{\mu\nu}$, with a gauge charge e.[5] The Néel phase is reached for $r < 0$ where $\langle\mathbf{z}\rangle \neq 0$ and hence $\boldsymbol{N} \neq 0$ by the above identification. By contrast, the VBS phase is associated with $\langle\mathbf{z}\rangle = 0$. The Z_4 order of the VBS state is not evident from Eq. (4.4): it arises as a result of a "dangerously irrelevant" perturbation, which is irrelevant at the fixed point associated with the phase transition, but relevant in the paramagnetic phase.[6]

---

3 To be precise, in the Hamiltonian formulation by order parameter we mean an operator which takes non-zero values in the symmetry-broken phase only.

4 An NCCP^N model is a "non-compact complex projective" theory, consisting of $N + 1$ dimensional complex fields defined modulo a global phase.

5 This term will not arise in our microscopic derivation presented in the next section, but it will be dynamically generated once we run the RG.

6 Recall that in the RG (ir)relevance is always associated with a specific fixed point.

Despite its innocuous appearance, Eq. (4.4) is still a complicated strongly coupled theory. Strictly speaking, it is not yet known whether it really admits a critical fixed point and hence a true second order phase transition. A critical fixed point is suggested by large-N-type arguments, and is further supported by a direct FRG analysis [Bar13] and various numerical studies of lattice models similar to (4.2) (see Sec. IX of Ref. [WNM+17] for a review of the current status). However, some finite-size numerical studies show unexpected system-size scaling, which could be due to a weakly first order transition, albeit with a very large (but finite) correlation length ξ. Such large ξ could result from the proximity of a true critical point not captured by the model (4.4), but "close-by" in parameter space [WNM+17] (e.g. at a slighly different dimension). Anyhow, the model (4.4) definitely admits (quasi-)universal behaviour on scales smaller than ξ, and is emphatically not a simple order parameter theory. In fact, since the critical modes are the spinons and not the order parameter fluctuations $\boldsymbol{N}$ themselves, one can conclude that the latter will pick up a large anomalous dimension at the transition, which will show in the experimentally relevant spin response function [SBS+04] (although no unambigious experimental signatures of deconfined criticality have been seen to date).

How does the NCCP^1 model come about? Amazingly, it can be derived from both sides of the phase transition, which we will quickly describe in turn.

Derivation from the Néel side

In starting from the Néel side (see, e.g., [Sen07, Aue94] for further details), we first compute a thermal partition function from the Hamiltonian of Eq. (4.2) (without the perturbation) by using spin coherent states $|\boldsymbol{n}\rangle$, with $\langle \boldsymbol{n}|\hat{\boldsymbol{S}}|\boldsymbol{n}\rangle = 1/2\boldsymbol{n}$. This gives us

$$Z = \int \prod_i \mathcal{D}\boldsymbol{n}_i(\tau) \exp\left(-\mathcal{S}_{\mathrm{B}} - \int_0^\beta d\tau \frac{J}{4} \sum_{\langle i,j \rangle} \boldsymbol{n}_i(\tau)\boldsymbol{n}_j(\tau)\right), \quad \mathcal{S}_{\mathrm{B}} = \int_0^\beta d\tau \langle \boldsymbol{n}(\tau)|\frac{d}{d\tau}|\boldsymbol{n}(\tau)\rangle ,$$

$$(4.5)$$

with unit length spin fields $\boldsymbol{n}_i(\tau)$, which fulfill the boundary condition $\boldsymbol{n}_i(0) = \boldsymbol{n}_i(\beta)$. The Berry phase term, which arises in the standard fashion in the path integral construction, is all-important: it knows about topological defects, and is ultimately responsible for the failure of the LGW reasoning as we will see below.

To describe the Néel state, we rewrite the unit length spin field $\boldsymbol{n}$ in terms of a staggered component $\boldsymbol{N}$ of unit length and a small uniform component $\boldsymbol{L}$. Of course, $\boldsymbol{N}$ is nothing but a fluctuating Néel order parameter field. Then, we go to a continuum limit, expand to quadratic order in gradients of $\boldsymbol{N}$ and in $\boldsymbol{L}$, and integrate out the $\boldsymbol{L}$ fields. After some manipulations, the same can be done for the Berry phase term. Ultimately, this results in the non-linear sigma model with Berry phase term

$$\mathcal{S} = \int d^2x d\tau \frac{1}{4} \left[(\nabla_{\mathbf{x}} \boldsymbol{N}(\tau, \mathbf{x}))^2 + (\partial_\tau \boldsymbol{N}(\tau, \mathbf{x}))^2 \right] + c \cdot \mathcal{S}_{\mathrm{B}}(\boldsymbol{N}(\tau, \mathbf{x})) ,$$

$$(4.6)$$

where we also took the limit $T \to 0$, and lumped together all constants in the (unimportant) prefactor c of the Berry phase term. Without Berry phases, this theory is simply a three-dimensional version of our usual Landau-Ginzburg theory from chapter 2, apart from the unit length constraint; however, it is known that the critical properties of the theory with length constraint are retained if the constrained is "softened" [Sac11] (i.e., only imposed on mean field level) with an appropriate potential of the form $r\boldsymbol{N}^2 + u\boldsymbol{N}^4$, and so we are truly back at the LGW description.

But the Berry phase term is indeed crucial. Still on the lattice, it can be given the compact form

$$\mathcal{S}_{\mathrm{B}} = \frac{i}{2} \sum_i (-1)^{\mathbf{r}_i} A_i \ , \tag{4.7}$$

where A_i is the area on the unit sphere contained within the closed loop swept out by the time evolution of $\boldsymbol{N}_i(\tau)$, which inherits the boundary condition of $\boldsymbol{n}_i(\tau)$. To understand the effect of $\mathcal{S}_{\mathrm{B}}$, one must consider the topological configurations $\boldsymbol{N}(\mathbf{x}, \tau)$ (in the continuum limit). For fixed time τ, and appropriate boundary conditions,[7] we can view $\boldsymbol{N}(\tau, \mathbf{x})$ as a map from a sphere to a sphere, equivalence classes of which can be characterized by the homotopy group $\pi_2(S^2) = \mathbb{Z}$, i.e. an integer topological winding number Q. Topologically non-trivial configurations are known as "skyrmions". Now, it can be shown that the Berry phase term vanishes for all smooth configurations $\boldsymbol{N}(\tau, \mathbf{x})$, including the skyrmions. However, the Berry phase term does contribute for spin configurations where the skyrmion winding number Q jumps in time, known as "hedgehogs". Detailed analysis [SBS$^+$04] shows that the Berry phases associated with these hedgehogs rapidly oscillate on the dual lattice formed by the centers of the square lattice plaquettes, and only quadrupled hedgehog events contribute to the partition function. So, effectively the role of the Berry phase is to disallow all non-quadrupled hedgehog events.

To incorporate this intricate effect, it is convenient to change to the so-called CP (complex projective) parametrization, and write

$$\boldsymbol{N}(\tau, \mathbf{x}) = \bar{z}_\alpha(\tau, \mathbf{x}) \boldsymbol{\sigma}_{\alpha\beta} z_\beta(\tau, \mathbf{x}) \ , \tag{4.8}$$

with a normalized complex field $\mathbf{z} = (z_1, z_2)$. Crucially, this parametrization entails a $U(1)$ gauge invariance $\mathbf{z} \to \mathbf{z} e^{i\chi(\tau, \mathbf{x})}$, with some function χ. This gauge invariance can be made explicit by rewriting [Aue94]:

$$\frac{1}{4}(\partial_\mu \boldsymbol{N})^2 = |\partial_\mu \mathbf{z}|^2 - \left(\tilde{\mathcal{A}}_\mu(\bar{\mathbf{z}}, \mathbf{z})\right)^2, \quad \tilde{\mathcal{A}}_\mu(\bar{\mathbf{z}}, \mathbf{z}) \equiv -i\left[\bar{\mathbf{z}}\partial_\mu \mathbf{z} - (\partial_\mu \bar{\mathbf{z}})\mathbf{z}\right] \ , \tag{4.9}$$

where we employed the normalization of $\mathbf{z}$. It is easily seen that the bilinear $\tilde{\mathcal{A}}_\mu$ transforms as expected under $U(1)$ gauge transformations, $\tilde{\mathcal{A}}_\mu \to \tilde{\mathcal{A}}_\mu + \partial_\mu \chi$. Via a Hubbard-Stratonovich transformation, it can be traded for a dynamical real gauge field $\mathcal{A}_\mu$ (i.e, an independent degree of freedom). In this way, the non-linear sigma model becomes

$$\mathcal{S} = \int d^3 x \sum_{\alpha, \mu} |(\partial_\mu - i\mathcal{A}_\mu) z_\alpha|^2 + \mathcal{S}_{\mathrm{B}} \ . \tag{4.10}$$

As discussed above, the main effect of the Berry phase term should be to allow for quadrupled hedgehogs only. Recall that the hedgehogs correspond to events where the winding number changes in time. Now, the winding number can be related to the flux of the gauge field $\mathcal{A}_\mu$ in units of 2π, and thus in the gauge theory language a hedgehog event corresponds to the sudden appearance of a magnetic monopole – an "instanton" event. Of course, an ordinary $U(1)$ gauge theory does not admit magnetic monopoles; their appearance means that we deal with a compact gauge theory, where $\mathcal{A}$ is defined modulo 2π. So, the Berry phase term effectively forbids non-quadrupled instantons.

Ref. [SBS$^+$04] suggested the following ultimate fate of these instantons: at the critical point of the model (4.10) (provided it exists), the instantons are irrelevant due to the quadrupling, and thus the only effect of the Berry phase term is to render the field $\mathcal{A}_\mu$ non-compact.

---

7 For spin configurations finite in energy, we demand $\lim_{|\mathbf{x}| \to \infty} \boldsymbol{N}(\tau, \mathbf{x}) = \mathrm{const.}$

Softening the unit length constraint, the theory (4.10) then assumes the form of the NCCP[1] model of Eq. (4.4), which therefore is the correct critical theory. The irrelevance of the instantons implies that the gauge flux is conserved at the critical point; in fact, such an emergent topological conservation law can be taken as the definition of a deconfined quantum critical point.

However, in the paramagnetic phase where $\langle \mathbf{z} \rangle = 0$, the instantons are relevant. Furthermore, by comparing symmetry transformation properties, the skyrmion creation/instanton operator is shown to be proportional to the VBS order parameter of Eq. (4.3). Thus, when the instantons proliferate, VBS order results, and the quadrupling of the instantons induces the Z_4 structure of the latter. This is, of course, very different from the LGW first-order picture of Sec. 4.1.1: the topological defects that destroy the Néel phase carry quantum numbers of the VBS phase, and thus the phase transition is expected to be continuous.

While the scenario described above cannot be proven rigorously for the model with two spinon flavors ($N = 2$), it can be substantiated by comparison to the $N = \infty$ and $N = 1$ models; the first case can be solved by saddle point analysis, and the latter by duality mapping on the XY model with quartic anisotropy. A similar duality argument is also made in the more intuitive derivation of the critical theory from the VBS side, which we will discuss next.

Derivation from the VBS side

Let us now approach the problem from the VBS side. If we were to write down a phenomenological Landau-Ginzburg theory that captures the Z_4 order plus fluctuations, a natural candidate would be a $(2 + 1)$ dimensional XY model with quartic anisotropy, given by

$$\mathcal{S}_{XY} = \int d^2 x d\tau \; K \sum_\mu (\partial_\mu \theta(\tau, \mathbf{x}))^2 - g \cos[4\theta(\tau, \mathbf{x})] \; , \quad K, g > 0 \; , \tag{4.11}$$

with an angle variable $\theta(\tau, \mathbf{x})$. The cosine term yields four minima, and induces Z_4 order for large g, while the gradient term parametrizes the energy cost of wave-like fluctuations.

But obviously this cannot be the full story: to begin with, it does not include the elementary excitations formed by breaking up a singlet, i.e. creating a spin-1 excitation, as shown in Fig. 4.3(a). We can think of them as being composed from two spin-1/2 "spinons". Deep in the VBS phase, these spinons are "confined", as sketched in Fig. 4.3(b): when they are pulled apart, a line of defects (with wrong clock order orientation) is created, and thus spinons are confined by a potential proportional to the length of the defect line.
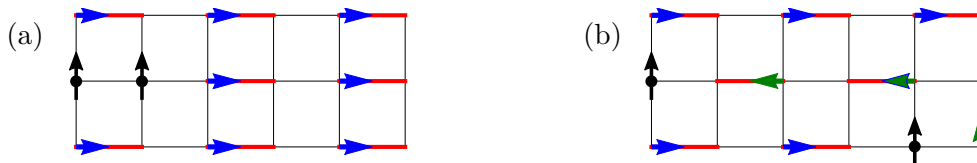

Figure 4.3 (a) Elementary spin-1 excitation, composed from two spin-1/2 spinons. (b) Separation of the spinons deep in the VBS phase creates a line of defects. Fig. adapted from [SBS[+]05].

Furthermore, we know that vortex-like topological defects are crucial in the XY model. As laid out in Ref. [LS04], in the VBS case these defects are Z_4 vortices, formed when four Z_4 domain walls terminate at a single point – see Fig. 4.4. As indicated in the Figure, an interesting pictorial property of these vortices is that they carry an unpaired spin – a spinon – in their core.

Without the spin structure, we would therefore describe the transition by a 3D XY theory with quartic anisotropy and vortices. This already points towards our CP gauge theory: as
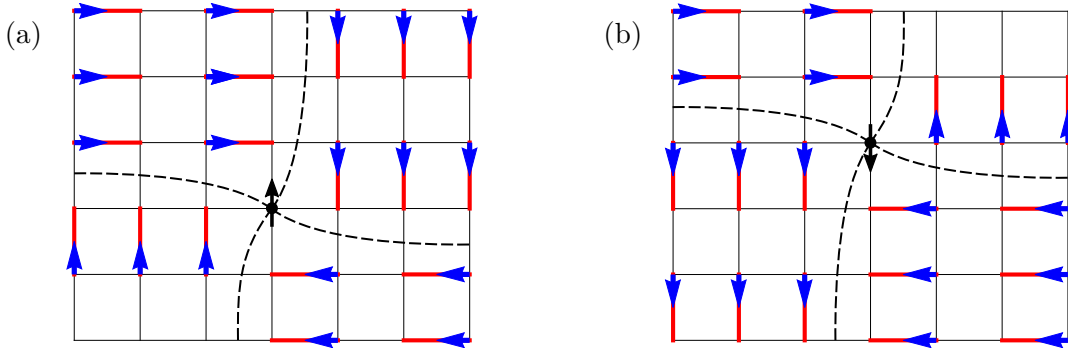
Figure 4.4 Topological excitations, with a spinon in the core. (a) Vortex. (b) Antivortex. The dashed lines indicate the domain walls. The free spin orientations in the core are chosen in a suggestive manner pointing towards the Néel state, as discussed further below. Fig. adapted from [LS04].

can be shown by a combination of explicit mappings of the form $K\partial_\mu\theta = \frac{1}{2\pi}\epsilon_{\mu\nu\lambda}\partial_\nu\mathcal{A}_\lambda$ and symmetry arguments (see, e.g., Sec. 6.3 of [Wen04]), the XY model with vortices is dual[8] to a $U(1)$ gauge theory with charged particles, which is known as charge-vortex duality [DH81]. Furthermore, the quartic anisotropy of the XY model corresponds to instantons in the gauge theory. But this quartic anisotropy is in fact known to be dangerously irrelevant at the critical point of the XY model [LSB07] – i.e., it is relevant in the ordered phase of the XY model only, but irrelevant at the critical point. Thus, instantons can be dropped at the critical point, and a non-compact gauge theory emerges. The analogies to the previous Néel description are obvious, only with a swapped labeling which phase is the "ordered one".

So, all we need to do is to include the spin structure. As we have seen in Fig. 4.4, the vortices carry a free spinon in their core, which we want to transform as a $SU(2)$ spinor. Provided that the RG picture carries over from the spinless case, we therefore need to write down a non-compact $U(1)$ gauge theory where the charges are spinors – and the NCCP[1] model of Eq. (4.4) is the minimal theory that does the job.

To complete the description, we need to understand how the identification of the Néel order parameter $\boldsymbol{N} = \bar{z}_\alpha\boldsymbol{\sigma}_{\alpha\beta}z_\beta$ emerges from the VBS side. The argument goes like this: we demand that both vortex and antivortex transform as spinors under $SU(2)$, and choose to represent the vortex by z_α. Since the antivortex has to carry opposite charge in the gauge picture, it has to be proportional to $\bar{z}$. Thus, for it to transform as a spinor, we can conclude that antivortices are represented by $\epsilon_{\alpha\beta}\bar{z}_\beta$ (which induces the correct transformation due to the unit determinant of the $SU(2)$ matrices). Furthermore, in the lattice picture of Fig. 4.4, we see that vortices and antivortices are associated with different sublattices of the square lattice. As a result, a translation by a lattice constant corresponds to $z_\alpha \to \epsilon_{\alpha\beta}\bar{z}_\beta$, and the associated three-dimensional vector $\bar{z}_\alpha\boldsymbol{\sigma}_{\alpha\beta}z_\beta$ changes sign under lattice translation, which is the defining property of the Néel order parameter. Thus, again we see that the topological defects in one phase carry the seeds of the other.

In the above line of argument, we have identified the spinor fields z of the NCCP[1] model with the spinon excitations of the VBS phase, which are fragments of the spin-1 excitations deep in the VBS phase, where they are confined due to the Z_4 domain structure. At the critical point, the Z_4 anisotropy is irrelevant, and the spinons are no longer bound by the linear potential – they are "deconfined", which finally explains the name of the theory. Still, they are not free particles, but interact via a Coulomb interaction, which is logarithmic in two dimensions.

---

8 Here, we will call two theories dual when they show the same universal behaviour, i.e. flow to the same IR fixed points, and we will not delve into the duality analysis further.

To conclude this section, let us summarize the failure of the LGW paradigm at the deconfined quantum critical points:

> In systems with "two-sided symmetry breaking", the LGW paradigm predicts a first order transition. However, a pure order parameter theory does not properly include topological defects, whose proliferation can drive the phase transition. An accurate treatment of these defects requires careful consideration of Berry phases neglected in the LGW approach. If the topological defects carry the quantum number of the other symmetry broken phase, a direct continuous transition between the two symmetry broken phases is possible, as a condensation of these defects destroys one order and simultaneously creates the other one.

4.1.3 Functional renormalization group

If we accept the NCCP[1] model as valid theory describing the phase transition, an RG analysis is the best way to extract the critical properties, and to see whether in this model there is a second order phase transition in the first place. A particular convenient method to derive the RG equations, which we will also apply in our original work discussed in Sec. 4.2 and in the polaron problem of Sec. 5.4.3, is the functional renormalization group (FRG). Let us give a quick recap of the method and its gauge-theory version; a thorough introduction to FRG can e.g. be found in Refs. [KBS10, MSH+12].

While in the Wilsonian-type RG employed in the previous chapter we studied the flow of isolated couplings, the basic idea of FRG is to compute the flow an "effective action" (a functional of the fields) as a whole. This is, for instance, required in the application of FRG to fermionic problems [MSH+12] where an infinite set of couplings can be relevant as seen in Sec. 3.1.2. In our case of interest, only a handful of couplings will suffice, but their flow can be obtained from the projection of the full functional flow in a computationally effective manner.

To derive this functional flow, for starters let us focus on a simple theory with an action $\mathcal{S}(\chi)$ depending on a single real scalar field χ (e.g. the ϕ^4 theory discussed before). Then, we can define a scale-dependent generating functional as

$$Z_\Lambda[J] = \int \mathcal{D}(\chi) \exp\left(-\mathcal{S}[\chi] - \Delta\mathcal{S}_\Lambda[\chi] + \int_k \chi(k)J(-k)\right), \quad \int_k = \int \frac{d^d k}{(2\pi)^d}, \qquad (4.12)$$

$$\Delta\mathcal{S}_\Lambda[\chi] = \frac{1}{2}\int_k \chi(k)R_\Lambda(k^2)\chi(-k).$$

Without the term $\Delta\mathcal{S}_\Lambda$ this would just be the ordinary generating functional, from which time-ordered correlation functions of χ can be derived by computing functional derivatives w.r.t. J evaluated at $J = 0$. The role of $\Delta\mathcal{S}_\Lambda$ is that of an IR cutoff, which ensures that only UV modes are integrated out: $\Delta\mathcal{S}_\Lambda$ involves a regulator function R_Λ, with the defining properties $R_\Lambda(k^2) > 0$ for $|k| \lesssim \Lambda$, $R_\Lambda(k^2) = 0$ for $|k| \gtrsim \Lambda$. Thus, for $\Lambda \to 0$, we recover the full generating functional.

One can now derive a differential equation for Z_Λ in a two-liner, by pulling out the cutoff term $\Delta\mathcal{S}_\Lambda[\chi]$ as $\exp(-\Delta\mathcal{S}_\Lambda[\delta/\delta J])$, and simply taking the Λ-derivative of the resulting form. This gives

$$\frac{dZ_\Lambda[J]}{d\Lambda} = -\frac{1}{2}\int_k \frac{\delta^2 Z_\Lambda[J]}{\delta J(k)\delta J(-k)}\frac{dR_\Lambda(k^2)}{d\Lambda}. \qquad (4.13)$$

This equation is formally exact, but in practice it is more convenient to change to a different formulation. To this end, we first introduce

$$W_\Lambda[J] = \ln(Z_\Lambda[J]), \quad \phi = \frac{\delta}{\delta J} W_\Lambda[J] \; . \tag{4.14}$$

As $\Lambda \to 0$, W_Λ becomes the generating functional for connected correlation functions, and ϕ simply the expectation value of the field χ. Now, we change from $W_\Lambda[J]$ to the **effective action** $\Gamma_\Lambda[\phi]$ via a modified Legendre transform as

$$\Gamma_\Lambda[\phi] = \int_k \phi[k] J[-k] - W_\Lambda[J] - \Delta\mathcal{S}_\Lambda[\phi] \; , \quad J = J[\phi] \; , \tag{4.15}$$

where $J[\phi]$ can be determined from inverting the relation (4.14). In the limit $\Lambda \to 0$ where $\Delta\mathcal{S}_\Lambda$ vanishes, this simplifies to the ordinary Legendre transformation.

As one can show, $\Gamma_\Lambda[\phi]$ has the following convenient properties: for $\Lambda \to \infty$, it reduces to the bare action $\mathcal{S}[\phi]$. In the IR limit $\Lambda \to 0$, it becomes the generating functional for the one-particle irreducible vertex functions. Thus, we can extract the flow of the vertex functions from taking functional derivatives of Γ_Λ.[9] In using simple properties of the Legendre transformation and our previous flow equation (4.13), one can now easily derive an exact flow equation for Γ_Λ, the so-called "Wetterich equation", which will be the starting point for our FRG computations. It reads

$$\frac{d\Gamma_\Lambda[\phi]}{d\Lambda} = \frac{1}{2} \int_p \left[\left(\frac{d}{d\Lambda} R_\Lambda(p^2) \right) \left(\frac{\delta^2 \Gamma_\Lambda[\phi]}{\delta\phi(p)\delta\phi(-p)} + R_\Lambda(p^2) \right)^{-1} \right] \tag{4.16}$$

$$\equiv \frac{1}{2} \mathrm{Tr} \left[\left(\frac{d}{d\Lambda} R_\Lambda \right) \left(\Gamma_\Lambda^{(2)} + R_\Lambda \right)^{-1} \right] \; .$$

The second line indicates the general formulation valid for theories with several field variables such as our NCCP[1] model; in this case the objects R_Λ and $\Gamma_\Lambda^{(2)}$ are generalized matrices in field space.

Of course, in a realistic model the complicated self-consistent equation (4.16) can never be solved unless significant approximations are made. Usually one proceeds by projecting the flow of (4.16) on a finite number of couplings, setting the rest to zero – this corresponds to a "truncation"-ansatz for the allowed terms in Γ_Λ. Due to the self-consistent structure of the Wetterich equation, this still is a non-perturbative approximation, but the accuracy of a given truncation is often hard to estimate.

In the FRG application of Sec. 4.2, we will choose a very simple truncation: our main motivation will be to show that deconfined criticality can arise in principle, and we will not claim numerical accuracy of critical exponents. Thus, for simplicity we essentially replace the effective action Γ_Λ on the right hand side of Eq. (4.16) by the bare action $\mathcal{S}$ with scale-dependent couplings. Since the Wetterich equation involves a trace, the flow equations obtained this way are formally correct to one-loop order (the bare action is zeroth loop order in this language). Without proving it, we expect that this loop expansion correctly reproduces the qualitative results of the studied phase transition; furthermore, we also consider a large-N extension of our model where the one-loop result should be accurate.

---

9 One-particle reducible contributions will not play any role within our level of approximations.

Gauge theory version

The NCCP[1] model, and similarly the model we will study in Sec. 4.2, contain complex multicomponent fields z_α coupled to a non-compact $U(1)$ gauge potential $\mathcal{A}_\mu$. While the multicomponent fields do not present difficulties, the incorporation of gauge fields into the FRG formalism is far from trivial (see [Gie12] for a detailed review in the context of $SU(N)$ gauge theories).

Naively, we could try to redo the derivation above, but with a field configuration $(z, \bar{z}, \mathcal{A}_\mu)$ (we'll suppress the components of the scalar fields and the dependence on the conjugated fields in the following). However, already the definition of the generating functional Z_Λ in step one, Eq. (4.12), does not go through: if we simply integrate over all possible fields $\mathcal{A}_\mu$, this is a gross overcounting, since the fields related by gauge transformations are physically equivalent. Thus, we must restrict ourselves to one "gauge-orbit", i.e. fix the gauge. For the $U(1)$ gauge theory this can simply be done by multiplication with a gauge-fixing term of the form

$$\exp\left(-S_{\mathrm{gf}}[\mathcal{A}]\right) \ , \quad S_{\mathrm{gf}}[\mathcal{A}] = \frac{1}{2\alpha} \int_x (\partial_\mu \mathcal{A}_\mu)^2 \ . \tag{4.17}$$

For $\alpha \to 0$, this can be seen as an exponential representation of a functional delta-function, which enforces the transversal gauge $\partial_\mu \mathcal{A}_\mu = 0$.[10] So, in principle we could rederive a Wetterich equation in a gauge fixed version. However, this comes at a big price: when the constraints due to gauge invariance are abandoned, it is not clear how to choose a meaningful truncation for the effective action Γ_Λ.

One way to circumvent this problem is the so-called **background field** formalism, which was introduced in the FRG context by Reuter and Wetterich in Ref. [RW94], and also used in the FRG study of the NCCP[1] model of Ref. [Bar13]; in our FRG analysis in Sec. 4.2, we follow these two references closely. Let us spell out the main idea, paraphrasing Ref. [RW94]: First, one defines the generating functional as

$$Z_\Lambda[J, K; A^0] = \int \mathcal{D}(z)\mathcal{D}(\mathcal{A}) \exp\left(-\mathcal{S}[z, A^0 + \mathcal{A}] - \mathcal{S}_{\mathrm{gf}}[\mathcal{A}] - \Delta\mathcal{S}_\Lambda^{(\mathrm{s})}[z; A^0] \right. \tag{4.18}$$
$$\left. - \Delta\mathcal{S}_\Lambda^{(\mathrm{g})}[\mathcal{A}] + \int_k \left(\bar{J}(k)z(k) + J(k)\bar{z}(k) + K_\mu \mathcal{A}_\mu\right)\right) \ .$$

Here, we have expanded the gauge fields around an arbitrarily chosen background field A^0, fixed the gauge for the fluctuating fields $\mathcal{A}$, and introduced infrared cutoffs for both the gauge field $\mathcal{A}$ and the scalar field z. The latter cutoff depends on the background field A^0 via the covariant derivative $\partial_\mu - iA^0_\mu$ (in real space).

Now, we can introduce $W_\Lambda = \ln(Z_\Lambda)$ as before. Defining

$$\phi = \frac{\delta W_\Lambda}{\delta \bar{J}}, \quad \tilde{\mathcal{A}}_\mu = \frac{\delta W_\Lambda}{\delta K_\mu} \ , \tag{4.19}$$

we can then change to the effective action:

$$\tilde{\Gamma}_\Lambda[\phi, \tilde{\mathcal{A}}, A^0] = \tag{4.20}$$
$$\int_k \left(\bar{J}(k)\phi(k) + J(k)\bar{\phi}(k) + K_\mu \tilde{\mathcal{A}}_\mu\right) - W_\Lambda[J, K; A^0] - \Delta\mathcal{S}_\Lambda^{(\mathrm{s})}[\phi; A^0] - \Delta\mathcal{S}_\Lambda^{(\mathrm{g})}[\tilde{\mathcal{A}}] \ .$$

---

10 For $SU(N)$ gauge theories [Gie12], one additionally picks up a so-called Faddeev-Popov determinant. To treat it properly, one must introduce fermionic "ghost" fields. Fortunately, for the $U(1)$ gauge theory, the determinant is simply a constant, or, in other words, the ghost fields do not couple to the physical fields.

Recombining $A \equiv \tilde{\mathcal{A}} + A^0$, one can define an effective action that depends on the two fields A, A^0 as

$$\Gamma_\Lambda[\phi, A, A^0] = \tilde{\Gamma}_\Lambda[\phi, A - A^0; A^0] . \tag{4.21}$$

Importantly, this new effective action is gauge-invariant under simultaneous gauge transformations of ϕ, A, A^0. It fulfills an exact Wetterich equation [RW94]:

$$\frac{d\Gamma_\Lambda[\phi, A, A^0]}{d\Lambda} = \frac{1}{2}\mathrm{Tr}\left[\left(\frac{d}{d\Lambda}R_\Lambda[A^0]\right)\left(\Gamma_\Lambda^{(2)}[\phi, A, A^0] + R_\Lambda[A^0]\right)^{-1}\right] . \tag{4.22}$$

Here, the scalar part of the regulator depends on A^0 as discussed below Eq. (4.18). Thus, we obtained a gauge-invariant Wetterich equation, which however has an unpleasant dependence on the arbitrary background field A^0. Furthermore, this version of gauge invariance is not very restrictive, since any function of $A - A^0$ is gauge invariant.

As a way out, Reuter and Wetterich propose to identify A^0 with A. This however leads to problems with the second functional derivative $\Gamma_\Lambda^{(2)}$ on the right hand side of (4.22), since

$$\left.\frac{\delta^2}{\delta A \delta A}\Gamma_\Lambda[\phi, A, A^0]\right|_{A^0 = A} \neq \frac{\delta^2}{\delta A \delta A}\Gamma_\Lambda[\phi, A, A] . \tag{4.23}$$

To plug this new hole, one first defines

$$\Gamma_\Lambda[\phi, A] \equiv \Gamma_\Lambda[\phi, A, A^0 = A] + C_\Lambda[A] , \tag{4.24}$$

where $C_\Lambda[A]$ is a freely chosen counterterm discussed below. Now, one can write

$$\Gamma_\Lambda^{(2)}[\phi, A, A] \simeq \Gamma_\Lambda^{(2)}[\phi, A] + \Gamma_{\mathrm{gf}}^{(2)}, \tag{4.25}$$

where $\Gamma_{\mathrm{gf}}^{(2)}$ is a constant contribution stemming from the original gauge-fixing term in (4.18), with

$$\Gamma_{\mathrm{gf}} = \frac{1}{2\alpha}\int_x \left(\partial_\mu(A_\mu - A_\mu^0)\right)^2 . \tag{4.26}$$

In the approximate equation (4.25), we dropped some terms arising due to the relation (4.23): the counterterm C_Λ is chosen precisely such that these extra terms are small. The discussion of the detailed form of C_Λ would take us too far afield – details can be found in Ref. [RW94]), and in our publication of Sec. 4.2.2. Even for the optimal choice of C_Λ, Eq. (4.25) is not exact, but can be formally assessed to hold to leading loop order, which will be sufficient for us.

Including the counterterm, an approximate gauge-invariant evolution equation for the functional $\Gamma_\Lambda[\phi, A]$ defined by Eq. (4.24) finally reads

$$\frac{d\Gamma_\Lambda[\phi, A]}{d\Lambda} \simeq \frac{dC_\Lambda[A]}{d\Lambda} + \frac{1}{2}\mathrm{Tr}\left[\left(\frac{d}{d\Lambda}R_\Lambda[A]\right)\left(\Gamma_\Lambda^{(2)}[\phi, A] + \Gamma_{\mathrm{gf}}^{(2)} + R_\Lambda[A]\right)^{-1}\right]. \tag{4.27}$$

Using this version of the Wetterich equation, and a rather involved truncation, the critical properties of the NCCP[1] model were analyzed in Ref. [Bar13]. Indeed, a critical fixed point (with one relevant RG direction) was found, indicating a second order phase transition as predicted by the deconfined criticality theory. Furthermore, an irrelevant RG eigenvalue of small magnitude was located. In part, this may explain the unusual scaling properties of the numerical studies, see the discussion below Eq. (4.4).

4.2 Deconfined criticality in $SU(3)$ magnets

4.2.1 Overview

The continuous nature of the Néel-VBS transition is contingent on the two-sublattice structure of the square lattice: for instance, this enables the nice vortex-antivortex pinning seen in Fig. 4.4. Is there a way for obtain deconfined criticality for non-bipartite lattices, say the triangular lattice? Clearly, for ordinary spin-1/2 Heisenberg models, the answer is no – due to frustration there is no Néel state in the first place.

In the publication below we show a way around this obstacle: instead of $SU(2)$ spins, we consider an $SU(3)$ model, where the Hilbert space on each lattice side is spanned by three-component complex spinors z_α, $\alpha \in \{1, 2, 3\}$, which transform under the fundamental representation of $SU(3)$. This is very different from previous $SU(N)$ approaches to deconfined criticality for instance studied in [Kau11, KS12], where bipartite lattices with a fundamental $SU(3)$ on one sublattice, and an antifundamental $SU(3)$ on the other sublattice were considered.

Similar to the spin-1/2-case, as parent Hamiltonian of our model we may think of the antiferromagnetic $SU(3)$ Heisenberg model

$$H_{SU(3)} = J \sum_{\langle i,j \rangle} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j - \tilde{J} H_{\text{pert}} , \quad J, \tilde{J} > 0, \tag{4.28}$$

where λ is an 8-dimensional vector of Gell-Mann matrices, the generators of $SU(3)$. The Heisenberg part is $SU(3)$ invariant, and H_{pert} indicates further $SU(3)$-symmetric perturbations, for example ring-exchange terms as studied in [Lai13]. Such an $SU(3)$-model is not just a theorists' fantasy, but for instance arises as a fine-tuned point in a spin-1 Hamiltonian with "quadrupolar" interactions (the so-called BBQ model, see Ref. [SS13]); furthermore, $SU(N)$ models can be realized in ultra-cold gas setups, see the paper below for references.

What are the phases of this model? In the Heisenberg limit, $\tilde{J} = 0$, DMRG studies [BCL+12] indicate a "color-ordered" ground state which is the three-state analog of the Néel state. As $r = \tilde{J}/J$ is increased, a paramagnetic phase should arise. While there is no $SU(3)$-version of the Lieb-Schultz-Mattis theorem in two dimensions yet which prohibits a featureless gapped paramagnet (related results for chains can be found in Ref. [LWMA17]), here we focus on a ground state that is an analog of the $SU(2)$ VBS state. In this sixthfold-degenerate state, the lattice is covered by $SU(3)$ singlets – as can be obtained explicitly within slave-fermion constructions [Lai13]. A cartoon of the two phases is shown in Fig. 4.5.
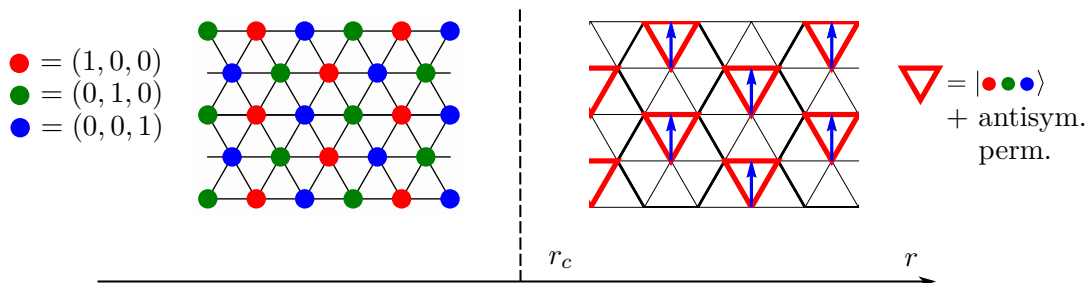

Figure 4.5 Sketch of the transition between a three-color-ordered state (with the colors representing the single site basis vectors), and the VBS state. The red triangle represents an $SU(3)$ singlet, obtained from antisymmetrized permutations of a three-site product state. Five additional VBS configurations can be obtained by rotations around the center of a hexagon, and the blue arrow indicates the sixfold clock order.

Borrowing intuition from the $SU(2)$ case, we have studied the phase transition occurring at $r = r_c$. We find that a deconfined quantum critical point is possible, albeit with several interesting new features:

○ While in the classical $SU(2)$ Néel configuration the choice of a spin orientation on one sublattice uniquely fixes the (antiparallel) one on the other, this is not the case for the three-color-order: here, the analog of the three-dimensional spin $\boldsymbol{N} = \langle \boldsymbol{\sigma} \rangle$ is an eight-dimensional flavor vector $\boldsymbol{m} = \langle \boldsymbol{\lambda} \rangle$, and the three color order corresponds to a coplanar $120°$ configuration of the flavor vectors on the three sublattices of the triangular lattice. If we fix the vector on one sublattice, the remaining two vectors can lie in any plane that contains the first vector. To accomodate this residual freedom, one can endow the order parameter vector with a sublattice index, introducing $\boldsymbol{m}_i, i \in \{1, 2, 3\}$, subject to a $(120°)$ constraint. A fractionalization similar to the $SU(2)$ case then leads to $\boldsymbol{m}_i = \bar{z}_{\alpha i} \boldsymbol{\lambda}_{\alpha\beta} z_{\beta i}$, with three-component and sublattice dependent spinors $z_{\alpha i}$; the $(120°)$-constraint maps to orthogonality of the spinors on the three sublattices.

○ This mapping also entails a threefold $U(1)$ gauge invariance. Therefore, one can guess a low-energy theory of the phase transition which contains spinors $z_{\alpha i}$ coupled to sublattice dependent non-compact $U(1)$ gauge fields $\mathcal{A}_i$. This corresponds to three coupled NCCP^2 models subject to an orthogonality constraint, and is a rare example of a deconfined criticality transition outside of the NCCP^N universality class.

○ This theory can be semi-rigorously derived from both sides of the phase transition. On the color-ordered side, one starts from a generalized non-linear sigma model, and obtains the gauge field from manipulations similar to Eqs. (4.8), (4.9). We assume that the Berry phase term renders instantons irrelevant, and the gauge theory is non-compact. This can be motivated from the VBS side: here, one can consider Z_6 vortices, which carry a free flavor in their core. Since in the corresponding XY model a sixthfold anisotropy is even more irrelevant than a fourfold one, duality suggests that instantons are indeed irrelevant.

○ The VBS vortices have a quite intriguing structure: while a vortex can be drawn simply on each sublattice, there is no way to draw an elementary antivortex with winding number (-1)! This may be related to the fact that, if a vortex field $z_{\alpha i}$ tranforms under an elementary $SU(3)$, the corresponding antivortex field which likewise transforms under an elementary $SU(3)$ (recall our discussion on page 62) and has negative gauge charge is given by $\bar{z}_{\alpha i} = \epsilon_{ijk} \epsilon_{\alpha\beta\gamma} z_{\beta j} z_{\gamma k}$ – antivortices can therefore be seen as composite objects. A gauge theory which accomodates this vortex structure is precisely the one obtained from the three color side.

○ To analyze our critical gauge theory, we have softened the orthogonality and normalization constraints on the spinors with an appropriate potential, which depends on a quadratic coupling m and three quartic couplings $\rho_n, n \in \{1, 2, 3\}$. We then used an FRG analysis in the background field formalism of Sec. 4.1.3, and a fairly simple truncation – in particular, we approached the phase transition from the symmetric side where $\langle \mathbf{z} \rangle = 0$. In this way, we do not find a critical fixed point for three-flavor spinors; however, if we generalize to N flavors, a critical point is indeed obtained for fairly high N. We expect that this critical point will survive the limit $N \searrow 3$ for a more involved truncation – at least, such a situation occurs in the $SU(2)$ case. In total, this lends strong support to the deconfined criticality scenario in our new model.

Deconfined Criticality in $SU(3)$ antiferromagnets on the triangular lattice

by

D. Pimenov,[1] and M. Punk[1]

[1] Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for

NanoScience, and Munich Center for Quantum Science and Technology (MCQST),

Ludwig-Maximilians-Universität München, 80333 München, Germany

Deconfined quantum criticality in SU(3) antiferromagnets on the triangular lattice

Dimitri Pimenov[*] and Matthias Punk

Physics Department, Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience,
Ludwig-Maximilians-University Munich, 80333 Munich, Germany
(Received 9 March 2017; published 25 May 2017)

We propose field theories for a deconfined quantum critical point in SU(3) antiferromagnets on the triangular lattice. In particular we consider the continuous transition between a magnetic, three-sublattice color-ordered phase and a trimerized SU(3) singlet phase. Starting from the magnetically ordered state we derive a critical theory in terms of fractional bosonic degrees of freedom, in close analogy to the well-developed description of the SU(2) Néel—valence bond solid (VBS) transition on the square lattice. Our critical theory consists of three coupled CP^2 models and we study its fixed point structure using a functional renormalization group approach in a suitable large N limit. We find a stable critical fixed point and estimate its critical exponents, thereby providing an example of deconfined criticality beyond the universality class of the CP^N model. In addition we present a complementary route towards the critical field theory by studying topological defects of the trimerized SU(3) singlet phase.

I. INTRODUCTION

Deconfined criticality is a concept that has emerged in recent years to describe quantum phase transitions beyond the Landau-Ginzburg paradigm [1–3]. Its basic idea is that a continuous quantum phase transition between two different symmetry broken phases is generically possible, if it is driven by the proliferation of topological defects which carry quantum numbers related to the order parameter of the other phase. Disordering one phase by condensing topological defects thus automatically leads to the appearance of the other order parameter. Such continuous transitions do not arise in the Landau-Ginzburg framework, where transitions between two different symmetry broken phases are generically of first order.

The prime example for deconfined criticality is the transition between a magnetically ordered Néel state and a valence bond solid (VBS) in SU(2) antiferromagnets on the square lattice [1–5]. The Néel state spontaneously breaks the spin rotation symmetry, whereas the VBS state spontaneously breaks lattice symmetries; the transition can be driven, e.g., by changing the relative strength of nearest-neighbor exchange and ring-exchange terms in generalized Heisenberg models [6]. On both sides of the transition elementary excitations (either spin waves in the Néel phase, or triplet excitations in the VBS phase) carry spin $S = 1$, while the relevant low energy degrees of freedom at the critical point are fractionalized (or deconfined) $S = 1/2$ spinon excitations, which are strongly coupled to an emergent $U(1)$ gauge field.

Early indications for the existence of such critical points came about by realizing that the nonlinear sigma model, describing low energy fluctuations of the Néel state in $(2 + 1)$-dimensions, breaks rotation symmetries in the paramagnetic phase [4,5]. This is due to the fact that Berry phase terms play a crucial role if singular configurations of the Néel order parameter field become important [7]. These singular configurations are topological defects known as "hedgehogs," which start to proliferate and condense at the transition out of the Néel state and can be viewed as magnetic monopoles of the dual gauge

theory. Since the monopole operator transforms nontrivially under lattice symmetries, the proliferation of monopoles automatically gives rise to VBS order. In hindsight, the fact that the paramagnetic state has to break lattice symmetries comes as no surprise. This is because unique paramagnetic ground states with an energy gap do not exist in models with one spin $S = 1/2$ per unit cell in two-dimensional systems, due to the Lieb-Schultz-Mattis-Hastings-Oshikawa theorem [8–10].

Alternatively one can understand the deconfined critical point by approaching it from the VBS phase. In this case the Néel state can be viewed as condensate of vortices on the VBS side, which carry spin $S = 1/2$ and are the electric charges of the dual gauge theory [11]. A crucial point in the theory of deconfined criticality is that the density of monopoles vanishes at the critical point and the vortices are thus deconfined.

An action for the critical theory can be readily derived by fractionalizing the Neel order parameter $\hat{N}$ in terms of bosonic spinor variables z_α ($\alpha \in \{1,2\}$) as [1,2,12,13]

$$\hat{N} = \bar{z}_\alpha \boldsymbol{\sigma}_{\alpha\beta} z_\beta ,\tag{1}$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. Physically, the fields z_α can be identified with XY-type vortices in the VBS phase which carry spin-1/2 and thus transform as spinors under SU(2). One can now construct the most general action by an expansion in powers and gradients of z_α that are allowed by symmetries. In addition, the $U(1)$ gauge redundancy $z_\alpha \rightarrow e^{i\phi} z_\alpha$ in the mapping (1) has to be incorporated by introducing a gauge field $\mathcal{A}$. The resulting theory is the celebrated euclidean CP^1 model in $2 + 1$ dimensions

$$S_{CP^1} = \int d^3x \left[\sum_{\alpha,\mu} |(\partial_\mu - i\mathcal{A}_\mu)z_\alpha|^2 + m z_\alpha \bar{z}_\alpha \right.$$
$$\left. + \rho(z_\alpha \bar{z}_\alpha)^2 + \frac{1}{4e^2}\mathcal{F}_{\mu\nu}^2 \right],\tag{2}$$

where the last term is the usual Maxwell term for the gauge field. Note that the gauge field in Eq. (2) is not compact, i.e., monopoles are irrelevant and the gauge theory is deconfined. The theory in Eq. (2) is strongly coupled and reliable results only exist in the large N limit of generalized CP^{N-1} models,

---
[*]D.Pimenov@physik.lmu.de

where the field z_α has N components. Nevertheless, extensive theoretical work indicates the presence of a stable critical fixed point at the relevant value $N = 2$, suggesting that a second order phase transition indeed exists [12,14–18]. By contrast, numerical studies of the SU(2) Heisenberg model with ring-exchange terms have not reached a consensus yet. While some early works claimed evidence for deconfined criticality [6,19], later strong corrections to scaling were found [20–22], while other works claim that the transition is weakly first order [23,24]. The situation is much clearer for particular $SU(N)$ generalizations of the Heisenberg model, where deconfined critical points in the universality class of the CP^{N-1} model have been found for $N > 4$ [25–27].

So far, most deconfined critical points in magnets that have been discussed in the literature are in the universality class of the CP^{N-1} model. In this work we are going to study a scenario for a deconfined critical point in a different universality class. In particular we consider SU(3) antiferromagnets in two dimensions, where a spin in the fundamental representation of SU(3) is placed on each site of a triangular lattice (note that this is in contrast to the $SU(N)$ generalizations in Ref. [25], where spins on the two different sublattices of the bipartite square lattice transform under fundamental and conjugate representations, respectively). Such SU(3) antiferromagnets appear at a specific parameter point of the more general spin-1 bilinear-biquadratic (BBQ) model [28]. Moreover, they can be realized in systems of ultracold atoms, where they arise in the strong coupling Mott limit of SU(3) symmetric Hubbard models with three flavors of fermions. The physics of such $SU(N)$ magnets, which host a multitude of novel states, have been realized in several cold atom setups in recent years [29–33]. In our work we study possible continuous transitions between SU(3) analogues of the magnetically ordered Néel phase and the VBS phase. We argue that the critical theory can be written in terms of three coupled CP^2 models, which features a new critical fixed point.

The rest of this paper is outlined as follows: In Sec. II we introduce the model and discuss the two symmetry broken phases of interest. A critical theory in terms of a CP parametrization is constructed in Sec. III starting from both a nonlinear sigma model description and by fractionalizing an appropriate order parameter. We briefly discuss the mean field phase diagram as well. In Sec. IV we perform a one-loop renormalization group study using the framework of functional renormalization group (FRG), where we treat the gauge sector within the background field formalism. Conclusions are presented in Sec. V, and in Appendix A we outline details of the FRG computation. Finally, in Appendix B we present a complementary route to obtain the critical field theory, by analyzing topological excitations in the trimerized SU(3) singlet phase.

II. SU(3) ANTIFERROMAGNETS

We consider an antiferromagnetic Heisenberg model on the triangular lattice with a spin in the fundamental representation of SU(3) on each lattice site. Its Hamiltonian is given by

$$H = J \sum_{\langle i,j \rangle} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j , \quad J > 0, \quad (3)$$
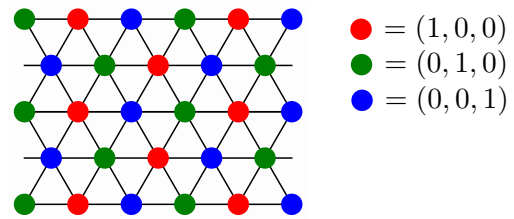

FIG. 1. A realization of the three-sublattice color-ordered state of SU(3) spins on the triangular lattice.

where $\boldsymbol{\lambda}_i$ is the eight-dimensional vector of Gell-Mann matrices, which are the generators of SU(3), and the sum extends over nearest neighbors. The Hilbert space at lattice site i is the projective space CP^2 of three-dimensional complex normalized vectors $\mathbf{z}_i$ defined up to a phase (for brevity, we will call them "spinors" in the following). Defining $\mathbf{m}_i = \langle \boldsymbol{\lambda}_i \rangle$, the mean field ground state is the well known $120°$ ordered state where

$$\sum_{i \in \triangle} \mathbf{m}_i = 0, \quad \|\mathbf{m}_i\| = \frac{2}{\sqrt{3}} , \quad (4)$$

for every elementary plaquette $\triangle$ of the triangular lattice. Its mean field energy is found to be $-2J$ [34]. The SU(3) flavor vectors $\mathbf{m}_i$ on the three sublattices are coplanar and span $120°$ angles, while the corresponding complex spinors $\mathbf{z}_i$ on the three sublattices are mutually orthogonal. Note that this configuration is the direct analog of the SU(2) Néel state on the square lattice. Indicating the basis vectors of CP^2 with colors red, green, blue, one possible realization of this color-ordered state is pictorially shown in Fig. 1. In analogy to the staggered magnetization for ordinary square lattice SU(2) antiferromagnets, we can define a scalar order parameter for the color-ordered phase by

$$m_c = \left\| \sum_{i \in A} \mathbf{m}_i + e^{i2\pi/3} \sum_{i \in B} \mathbf{m}_i + e^{i4\pi/3} \sum_{i \in C} \mathbf{m}_i \right\|, \quad (5)$$

where A, B, C are the three sublattices. One can straightforwardly show that m_c is maximized in the color-ordered state out of all possible states.

Numerical studies of the Hamiltonian in Eq. (3) show that the exact ground state indeed exhibits three-sublattice color order [35]. Generalized Hamiltonians with additional ring-exchange terms, which arise naturally from higher order terms in the usual strong-coupling expansion of the $SU(N)$ Hubbard model, have been studied in Refs. [36,37], where a variety of nontrivial paramagnetic ground states were found. Among several quantum spin liquid states, a trimerized SU(3) singlet state was found in Ref. [37]. This state is an analog of the VBS state in SU(2) spin systems. In the trimerized state the lattice is covered with SU(3) singlets $|S\rangle$ formed by three spins on an elementary triangle

$$|S\rangle = \epsilon_{\alpha\beta\gamma} z_{\alpha 1} z_{\beta 2} z_{\gamma 3} , \quad (6)$$

where $\varepsilon_{\alpha\beta\gamma}$ is the fully antisymmetric tensor of SU(3) and a summation convention is used for greek indices. The singlets order in a specific pattern, thereby breaking lattice translation and rotation symmetries. Here, we will focus on the most
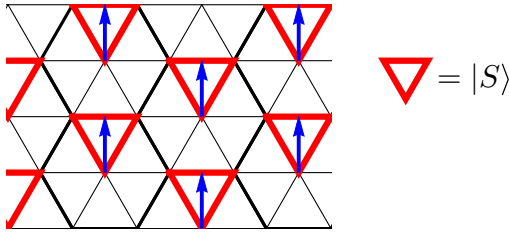
FIG. 2. A particular realization of the trimerized SU(3) singlet state. The arrows represent the $\mathbb{Z}_6$-clock order parameter.

simple singlet configurations, which can be indexed by a $\mathbb{Z}_6$-clock order parameter. They correspond to a sixfold degenerate paramagnetic ground state. A pictorial representation of a trimer state is shown in Fig. 2.

In this paper we want to address the question whether a direct second order quantum phase transition between the color-ordered and the trimerized phase is possible, and study its properties. This transition would be a direct analog of the deconfined critical point for the Néel-VBS transition in SU(2) antiferromagnets on the square lattice.

III. CRITICAL THEORY

A. Path-integral derivation of the critical theory

Our critical theory will be based on the nonlinear sigma model for the color ordered state derived by Smerald and Shannon [38], which we briefly review in the following. The starting point is the SU(3)-symmetric point of the bilinear-biquadratic model, given by

$$H_{\text{BBQ}}^{SU(3)} = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + (\mathbf{S}_i \cdot \mathbf{S}_j)^2 , \qquad (7)$$

where $\mathbf{S}_i$ are spin-1 operators. Up to a constant shift and rescaling of J, this Hamiltonian is equivalent to Eq. (3) [39]. In terms of spin-1 coherent states Eq. (7) can then be rewritten as

$$H_{\text{BBQ}}^{SU(3)} = J \sum_{\langle i,j \rangle} |\mathbf{d}_i \cdot \bar{\mathbf{d}}_j|^2 , \qquad (8)$$

where $\mathbf{d}_i$ is a three-dimensional complex normalized vector. In fact, the overall phase of $\mathbf{d}_i$ is fixed in Ref. [38], but the precise phase choice is immaterial at this stage of the analysis. In the next step, the fluctuating fields $\mathbf{d}_i$ are expanded around a generic color-ordered configuration, which is parametrized by three mutually orthogonal, complex vector fields $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$, where 1,2,3 correspond to the three sublattices. These fields fulfill

$$\bar{\mathbf{z}}_i \cdot \mathbf{z}_j = \delta_{ij}. \qquad (9)$$

Deviations from this color order are parametrized in terms of small "canting" fields $\mathbf{l}$, which can be integrated out at the quadratic level. Changing to a continuum description and introducing a kinetic term (which arises in the standard way from the path-integral construction), the resulting partition function in the zero temperature limit is given by

$$Z \sim \int \prod_i \tilde{\mathcal{D}}\mathbf{z}_i \prod_{j \geqslant i} \delta(\bar{\mathbf{z}}_i \cdot \mathbf{z}_j - \delta_{ij}) \exp(-S), \qquad (10)$$

where the measure $\tilde{\mathcal{D}}\mathbf{z}_i$ contains a gauge fixing of the phase of $\mathbf{z}_i$ to avoid double counting of physical degrees of freedom. The euclidean action S appearing in (10) reads

$$S = \int d^3x \, \alpha \sum_i \bar{\mathbf{z}}_i \cdot \partial_\tau \mathbf{z}_i + \sum_{\mu, i \neq j} |\bar{\mathbf{z}}_i \cdot \partial_\mu \mathbf{z}_j|^2 , \qquad (11)$$

where $\mu \in \{0,1,2\}$, and $\tau = x_0$ denotes the imaginary time direction. Here α is a numerical coefficient which depends on J and the details of the continuum limit and will not be of importance to us.

An important property of the action in Eq. (11) is its manifest invariance under *sublattice-dependent U(1) gauge transformations* of the form

$$\mathbf{z}_i(\mathbf{x}) \rightarrow e^{i\theta_i(\mathbf{x})} \mathbf{z}_i(\mathbf{x}). \qquad (12)$$

For the second term in Eq. (11) this invariance follows from the orthogonality constraint contained in Eq. (9), while the first term only picks up a total time derivative under gauge transformation by virtue of the normalization constraint.

The first term in Eq. (11) corresponds to a topological Berry phase term. Only singular field configurations should give a nonzero Berry-phase contribution. By analogy to SU(2), we can expect these configurations to be hedgehog events, where an appropriately defined soliton winding number jumps in time. In the SU(2) case on the square lattice, inclusion of Berry phase terms renders these events dangerously irrelevant at the quantum critical point, but relevant in the paramagnetic phase [1]. In the gauge language, the hedgehogs correspond to magnetic monopoles, and their irrelevance makes the resulting $U(1)$ gauge theory noncompact. For SU(3), the soliton structure on the color-ordered side was recently studied in Ref. [40] by a homotopy analysis of the ground state manifold, giving rise to a $\mathbb{Z} \times \mathbb{Z}$ winding number classification. We will not perform an analysis of the corresponding hedgehog events here and disregard the Berry-phase terms altogether, assuming that their only role is to render the $U(1)$ gauge field noncompact as in the SU(2) case.

Following Refs. [40,41] we can bring the remaining action in another form by introducing the following real functions of the $\mathbf{z}_i$ fields

$$\tilde{\mathcal{A}}_\mu^i = -\frac{i}{2}[\bar{\mathbf{z}}_i \partial_\mu \mathbf{z}_i - (\partial_\mu \bar{\mathbf{z}}_i)\mathbf{z}_i]. \qquad (13)$$

Under the gauge transformation in Eq. (12) $\tilde{\mathcal{A}}_\mu^i$ transforms as

$$\tilde{\mathcal{A}}_\mu^i \rightarrow \tilde{\mathcal{A}}_\mu^i + \partial_\mu \theta_i. \qquad (14)$$

With help of these fields, and the identities in Eq. (9), the Lagrangian is now rewritten as (c.f. [40])

$$\sum_{\mu, i \neq j} |\bar{\mathbf{z}}_i \cdot \partial_\mu \mathbf{z}_j|^2 = \sum_{i,\mu} |\partial_\mu \mathbf{z}_i|^2 - \left(\tilde{\mathcal{A}}_\mu^i\right)^2. \qquad (15)$$

Following Ref. [41] we may trade the $\mathbf{z}$-dependent gauge fields $\tilde{\mathcal{A}}$ for $\mathbf{z}$-independent gauge fields $\mathcal{A}$ with help of a Hubbard-Stratonovich transformation of the form

$$\exp\left[\tilde{\mathcal{A}}_\mu^i{}^2\right] \sim \int_{-\infty}^{\infty} d\mathcal{A}_\mu^i \exp\left(-\mathcal{A}_\mu^i + 2\mathcal{A}_\mu^i \tilde{\mathcal{A}}_\mu^i\right). \qquad (16)$$

For Eq. (16) to hold after gauge transformations, the fields $\mathcal{A}$ must inherit the transformation properties of the fields $\tilde{\mathcal{A}}$ given

by (14). Inserting (16) into the partition function, we finally arrive at

$$Z = \int \prod_{i,\mu} \mathcal{D}\mathbf{z}_i \tilde{\mathcal{D}} \mathcal{A}_\mu^i \prod_{j \geqslant i} \delta(\bar{\mathbf{z}}_i \cdot \mathbf{z}_j - \delta_{ij})$$

$$\times \exp\left\{ -\int d^3x \sum_{i,\mu} \left| (\partial_\mu - i\mathcal{A}_\mu^i)\mathbf{z}_i \right|^2 \right\}. \quad (17)$$

In writing Eq. (17) we tacitly performed the following manipulation: the gauge fixing term for the phase of the $\mathbf{z}$ fields, which was contained in the measure $\tilde{\mathcal{D}}\mathbf{z}$ in Eq. (10), is carried over to an equivalent gauge fixing condition in the measure $\tilde{\mathcal{D}}\mathcal{A}$ in Eq. (17), which will be made explicit by introducing a standard gauge fixing term later.

The action of Eq. (17) bears close resemblance to the CP^1 action of Eq. (2). In fact, one may imagine a derivation of Eq. (2) for SU(2) largely analogous to the one presented above, with two-dimensional $\mathbf{z}_i$ fields, and two sublattices only. However, for $\mathbf{z}_i \in \mathbb{C}^2$, the orthogonality constraint contained in Eq. (17) fully determines $\mathbf{z}_2$ as a function of $\mathbf{z}_1$ (or vice versa), up to a phase. This can be made explicit by writing

$$z_{\alpha 2} \sim \epsilon_{\alpha\beta}\bar{z}_{\beta 1}. \quad (18)$$

Inserting this into (17), the gauge fields couple to the same $\mathbf{z}$ in identical fashion and are indistinguishable. Softening the unit length constraint on $\mathbf{z}$, one therefore recovers (2) (up to the Maxwell term, which is generated during the RG flow, see below).

B. Fractionalizing the order parameter

For spin-1/2 models, a common shortcut in deriving critical actions is a fractional parametrization of an appropriate real order parameter [13,42], as shown in Eq. (1) for the Néel state. We can proceed accordingly, identifying the triple of physical 8-D flavor vectors $\mathbf{m}_i$ with $120°$ order from Eq. (4) as order parameter. The vectors $\mathbf{m}_i$ parametrize the manifold of classical ground states, which are, in fact, product states. Therefore, we can re-express them as

$$\mathbf{m}_i = \bar{z}_{\alpha i}\boldsymbol{\lambda}_{\alpha\beta}z_{\beta i}, \quad i,\alpha,\beta \in \{1,2,3\}, \quad (19)$$

where the fields $\mathbf{z}_i$ precisely fulfill the normalization and orthogonality constraints of Eq. (9). The phase ambiguity in Eq. (19) entails a threefold gauge invariance, and an action in terms of $\mathbf{z}_i$ can be readily derived by expanding in covariant derivatives. This reasoning immediately gives us our critical theory in Eq. (17). One should note that, while the above argumentation seems to be limited to the description of classical order parameter fluctuations, the previous path integral formulation explicitly shows that the critical theory does include quantum fluctuations as well.

C. Softening the constraints

The critical theory in Eq. (17) is rather inconvenient to handle due to the delta-function constraints. We can proceed by softening the constraints and replace them by an appropriate potential $V(\mathbf{z})$, which must obey the following properties: (i) invariance under global SU(3)-rotations $\mathbf{z}_i \to \hat{U}\mathbf{z}_i$; (ii) invariance under lattice symmetries, which simply permute
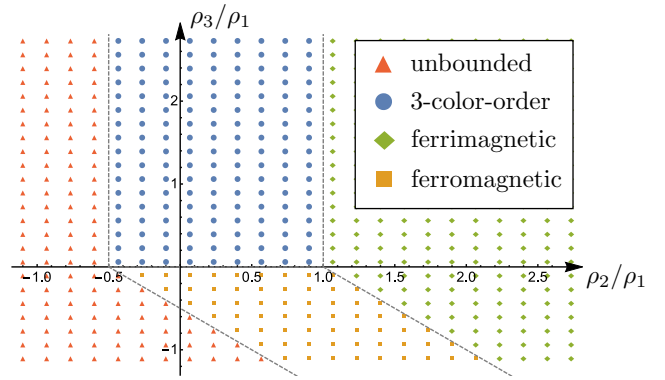

FIG. 3. Mean field phase diagram corresponding to the potential $V(\mathbf{z})$ from Eq. (20). The dashed lines indicate the phase boundaries.

the sublattice indices (e.g., under rotations with base point on sublattice 1: $\mathbf{z}_1 \to \mathbf{z}_1, \mathbf{z}_2 \leftrightarrow \mathbf{z}_3$); (iii) $U(1)$ gauge invariance. Expanding up to quartic terms, the resulting general potential has the form

$$V(\mathbf{z}) = \sum_i \{m(\mathbf{z}_i \cdot \bar{\mathbf{z}}_i) + \rho_1(\mathbf{z}_i \cdot \bar{\mathbf{z}}_i)^2\}$$

$$+ 2\sum_{i \neq j} \{\rho_2(\mathbf{z}_i \cdot \bar{\mathbf{z}}_i)(\mathbf{z}_j \cdot \bar{\mathbf{z}}_j) + \rho_3(\mathbf{z}_i \cdot \bar{\mathbf{z}}_j)(\mathbf{z}_j \cdot \bar{\mathbf{z}}_i)\},$$

$$(20)$$

where the factor of 2 is introduced for later convenience, and $m, \rho_1, \rho_2, \rho_3$ are real coupling constants. Our resulting critical theory therefore reads

$$S = \int d^3x \sum_{i,\mu} \left| (\partial_\mu - i\mathcal{A}_\mu^i)\mathbf{z}_i \right|^2 + V(\mathbf{z}), \quad (21)$$

which features three CP^2 models coupled via quartic interaction terms.

To gain some insight into the structure of V, let's perform a mean field analysis, restricting ourselves to the ordered phase where $m < 0, \rho_1 > 0$. It is obvious that the term $\sim \rho_3$ is the only one which depends on the relative direction of the spinors: When $\rho_3 > 0$, the spinors are orthogonal on the mean field level and parallel for $\rho_3 < 0$. Some easy algebra then yields the following mean field boundaries: First, for $\rho_3 > 0$:

(1) $\rho_2 < -\frac{1}{2}\rho_1$: For this (unphysical) parameter choice, the potential is not bounded below.

(2) $-\frac{1}{2}\rho_1 < \rho_2 < \rho_1$: These values correspond to a well-defined three color order, with nonzero expectation values and mutual orthogonality for all $\mathbf{z}_i$.

(3) $\rho_2 > \rho_3$: One finds a "ferrimagnetic" phase, where the expectation value of two spinors $\mathbf{z}_i$ is zero.

Second, for $\rho_3 < 0$:

(1) $\rho_3 < -\frac{1}{2}\rho_1 - \rho_2$: The potential is unbounded.

(2) $-\frac{1}{2}\rho_1 - \rho_2 < \rho_3 < \rho_1 - \rho_2$: Corresponds to a "ferromagnetic" phase, where all spinors have a nonzero expectation value and point in the same direction.

(3) $\rho_1 - \rho_2 > \rho_3$: Ferrimagnetic phase.

The above phases are summarized in Fig. 3.

IV. FRG ANALYSIS

A. General FRG setup

We can analyze the possible phase transitions predicted by the critical action S by looking for its RG fixed points. To derive the flow equations, our method of choice is functional renormalization group (FRG), employing conventions from Refs. [17,43]. The backbone of this analysis is the functional *Wetterich equation*, which describes the running of the scale-dependent Legendre effective action Γ_Λ under variation of the momentum cutoff scale Λ. At the initial UV scale Λ_0, Γ_{Λ_0} reduces to the bare action S; for $\Lambda \to 0$, Γ_Λ becomes the Legendre transform of the true generating functional of connected Green's functions. This is achieved by successively integrating out UV degrees of freedom via inclusion of regulator terms $\mathcal{R}_\Lambda$, which suppress IR fluctuations. Taking functional derivatives, the flow of Γ_Λ can then be projected on the flow of the coupling constants.

While the FRG treatment of the scalar sector is very straightforward, technical difficulties arise upon including gauge degrees of freedom. Several workarounds are available [44]; following previous treatments of CP^n models [16,17,45–47], we will employ the *background field* formalism, introduced by Reuter and Wetterich. Its main idea is to work with an effective action Γ_Λ which is manifestly gauge invariant, while at the same time containing a gauge-fixing term necessary for well-defined functional integrals in the first place. This gauge-invariant formulation allows us to choose a meaningful truncation of Γ_Λ, as necessary to make any technical progress.

To implement this idea, one first expands the dynamical gauge fields $\mathbf{A}$ appearing in the bare action around some fixed background field $\bar{\mathbf{A}}$ (we suppress indices for now), which gauge transforms in the standard way. The effective action Γ_Λ, obtained via an appropriate Legendre transform of S, then depends on $\mathbf{A} = \langle \mathcal{A} \rangle$, $\phi = \langle z \rangle$, and $\bar{\mathbf{A}}$ (averages are taken w.r.t. to S along with sources and regulator terms), and is gauge invariant under gauge transformations of both $\bar{\mathbf{A}}, \mathbf{A}$, and ϕ. However, since arbitrary powers of $(\bar{\mathbf{A}} - \mathbf{A})$ are gauge invariant, using $\Gamma_\Lambda[\mathbf{A}, \bar{\mathbf{A}}, \phi]$ is still inconvenient, and one needs to eliminate the field $\bar{\mathbf{A}}$. This can be achieved by identifying it with $\mathbf{A}$. In doing so, one picks up spurious functional derivatives, which can be partially accounted for by an appropriate gauge-invariant counterterm $\mathcal{C}_\Lambda(\mathbf{A})$. This term will modify the flow equation of the gauge coupling only. Ultimately, defining an appropriate effective action, one arrives at the following approximate flow equation

$$\frac{\partial}{\partial l} \Gamma_\Lambda[\phi, \mathbf{A}]$$
$$\simeq \frac{1}{2} \mathrm{Tr}\left[\frac{\partial}{\partial l} \mathcal{R}_\Lambda[\mathbf{A}] \left(\Gamma_\Lambda^{(2)}[\phi, \mathbf{A}] + \Gamma_{\mathrm{gf}}^{(2)} + \mathcal{R}_\Lambda[\mathbf{A}] \right)^{-1} \right]$$
$$+ \frac{\partial}{\partial l} \mathcal{C}_\Lambda[\mathbf{A}]. \tag{22}$$

Here l is the logarithmic RG scale, connected to the momentum cutoff-scale Λ by

$$\Lambda = \Lambda_0 \exp(-l). \tag{23}$$

$\Gamma_\Lambda^{(2)}$ and $\Gamma_{\mathrm{gf}}^{(2)}$ are second derivatives of Γ_Λ and the gauge-fixing term Γ_{gf} (we choose the Lorenz gauge) w.r.t. the fields $\mathbf{A}, \phi$. All objects on the r.h.s. of (22) are matrix valued in $(\phi, \bar{\phi}, \mathbf{A})$ space, and the trace involves a summation in this space as well as over all internal indices.

To proceed, we need to specify an ansatz for an effective action Γ_Λ. We choose it to be equal to the bare action, with running couplings, a standard wave-function renormalization term Z_Λ, and lattice-dependent Maxwell terms coupled by a running gauge charge e_Λ. Explicitly writing out the index structure and employing Einstein sums we have

$$\Gamma_\Lambda[\mathbf{A}, \phi] = \int d^3x \left\{ Z_\Lambda \left| (\partial_\mu - i A_\mu^i) \phi_{\alpha i} \right|^2 + m_\Lambda \phi_{\alpha i} \bar{\phi}_{\alpha i} \right.$$
$$\left. + \hat{\rho}_{ijkl}^\Lambda \phi_{\alpha i} \bar{\phi}_{\alpha j} \phi_{\beta k} \bar{\phi}_{\beta l} + \frac{1}{4 e_\Lambda^2} \left(\mathcal{F}_{\mu\nu}^i \right)^2 \right\}, \tag{24}$$

where $\hat{\rho}^\Lambda$ is a compact notation for the quartic potential terms of the form

$$\hat{\rho}_{ijkl}^\Lambda = \begin{cases} \rho_1^\Lambda, & i = j = k = l \\ \rho_2^\Lambda, & i = j, \quad k = l, \quad i \neq k \\ \rho_3^\Lambda, & i \neq j, \quad i = l, \quad k = j \\ 0 & \text{else.} \end{cases} \tag{25}$$

Roman letters always denote sublattice indices, α, β, γ $SU(3)$ indices, and μ, ν spacetime indices. One should note that all gauge charges are equal, as required by translational invariance, which permutes the gauge field sublattice indices. Also, we disregard terms which couple the gauge fields on different sublattices to each other, since they renormalize on two-loop level only. Specifying the regulator $\mathcal{R}_\Lambda[\mathbf{A}]$ and the counterterm $\mathcal{C}_\Lambda[\mathbf{A}]$, we can obtain the one-loop flow equation of the couplings from the central equation (22) by taking appropriate functional derivatives w.r.t. the fields. A pictorial representation of these flow equations is shown in Fig. 4.

Unlike earlier studies of CP^n models, we take the functional derivatives at zero scalar fields for simplicity, approaching the fixed point from the symmetric phase. This usually leads to numerically less precise results for the critical exponents, but is sufficient to determine the fixed point structure of our theory. To correct for these truncation related errors, we derive the flow equations for general $SU(N)$, i.e., we extend the summations over the three SU(3) indices α, β in Eq. (24) to a summation which ranges from 1 to N. We then study the behavior of the flow equations in the large N limit. In the CP^{N-1}-case, this was shown to yield qualitatively correct results in Ref. [16].

Technical details of the derivation are given in Appendix A. We phrase the flow equations in terms of dimensionless quantities

$$\tilde{m} = \frac{m_\Lambda}{\Lambda^2 Z_\Lambda}, \quad \tilde{\rho}_i = \frac{\rho_i^\Lambda}{\Lambda Z_\Lambda^2}, \quad \tilde{e}^2 = \frac{e_\Lambda^2}{\Lambda}. \tag{26}$$

Furthermore, we introduce the anomalous dimension of the scalar fields as

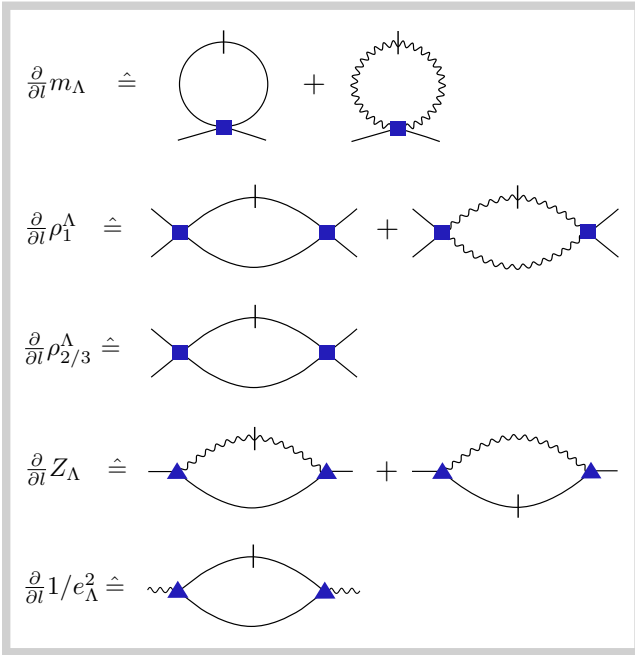$$\eta \equiv \frac{\partial}{\partial l} \log(Z_\Lambda). \tag{27}$$

FIG. 4. Diagrammatic representation of the flow equations. Full (wiggly) lines denote scalar (gauge) field propagators, while blue square (triangles) denote four-point (three-point) vertices. Vertical lines represent regulator insertions.

The flow equations obtained this way read

$$\frac{\partial}{\partial l}\tilde{m} = \tilde{m} \cdot (2 - \eta)$$
$$+ \frac{\tilde{e}^2(\tilde{m}+1)^2 + 2((N+1)\tilde{\rho}_1 + 2(N\tilde{\rho}_2 + \tilde{\rho}_3))}{3\pi^2(\tilde{m}+1)^2}$$

(28)

$$\frac{\partial}{\partial l}\tilde{\rho}_1 = \tilde{\rho}_1 \cdot (1 - 2\eta)$$
$$- \frac{4(\tilde{e}^4(\tilde{m}+1)^3 + (N+4)\tilde{\rho}_1^2 + 2(N\tilde{\rho}_2^2 + 2\tilde{\rho}_3\tilde{\rho}_2 + \tilde{\rho}_3^2))}{3\pi^2(\tilde{m}+1)^3}$$

(29)

$$\frac{\partial}{\partial l}\tilde{\rho}_2 = \tilde{\rho}_2 \cdot (1 - 2\eta)$$
$$- \frac{4((N+2)\tilde{\rho}_2^2 + 2\tilde{\rho}_1((N+1)\tilde{\rho}_2 + \tilde{\rho}_3) + 2\tilde{\rho}_3\tilde{\rho}_2 + \tilde{\rho}_3^2)}{3\pi^2(\tilde{m}+1)^3}$$

(30)

$$\frac{\partial}{\partial l}\tilde{\rho}_3 = \tilde{\rho}_3 \cdot (1 - 2\eta) - \frac{4\tilde{\rho}_3((N+1)\tilde{\rho}_3 + 2\tilde{\rho}_1 + 4\tilde{\rho}_2)}{3\pi^2(\tilde{m}+1)^3}$$ (31)

$$\frac{\partial}{\partial l}\tilde{e}^2 = \frac{\tilde{e}^2(12\pi^2(\tilde{m}+1) - N\tilde{e}^2)}{12\pi^2(\tilde{m}+1)}$$
$$+ \frac{\tilde{e}^4 N(\tilde{m}\sqrt{\tilde{m}+2} - 2(\tilde{m}+1)\coth^{-1}(\sqrt{\tilde{m}+2}))}{12\pi^2(\tilde{m}+1)(\tilde{m}+2)^{3/2}}$$

(32)

$$\eta = -\frac{8\tilde{e}^2(\tilde{m}+2)}{9\pi^2(\tilde{m}+1)^2}.$$ (33)

B. Fixed point structure

Let's analyze the fixed point structure of the above flow equations. The fixed points are obtained by numerically solving for the zeros of the beta functions. Linearizing the beta functions around the fixed points and determining the eigenvalues of the resulting coefficient matrix, one can then derive the stability properties. Note that the coefficient matrix is in generally not symmetric (s.t. the left and right eigenvectors do not coincide), but always found to be diagonalizable.

As a sanity check, one can set $\tilde{\rho}_2 = \tilde{\rho}_3 = 0$ [which is of course a solution to Eqs. (30) and (31)]. Then, our model simply reduces to three copies of the standard CP^{N-1} model, and we can compare the fixed point structure to prior treatments, in particular to the large N analysis of Bergerhoff *et al.* [16]. For small or moderate N, as can be deduced from Eq. (32), no fixed point at nonzero gauge charge is found, and the flow equations reduce to the ones of the usual $O(2N)$ model: The gauge field fluctuations are not strong enough to renormalize the scalar sector. As a result, there are just two fixed points: the Gaussian and the Wilson-Fisher fixed point, which has one additional unstable direction corresponding to the gauge charge. Since the scalar field anomalous dimension is exclusively generated by gauge field fluctuations within our treatment ($\eta \sim \tilde{e}^2$), in this regime we find $\eta = 0$.

The picture changes for $N \geqslant 353$, where two further fixed points at nonzero gauge charge appear, corresponding to additional zeros of the gauge charge beta function (32). The analytical structure of these additional fixed points is transparent in the limit $N \to \infty$: First, one observes that the gauge flow equation (32) has the form

$$\frac{\partial}{\partial l}\tilde{e}^2 = \tilde{e}^2 \cdot (1 - Nf(\tilde{m}) \cdot \tilde{e}^2),$$ (34)

where $f(\tilde{m})$ is some function; for $\tilde{m}$ sufficiently larger than -1, which is always fulfilled for meaningful fixed points, $f(\tilde{m})$ is of order 1. Therefore, the fixed point value $\tilde{e}_\star^2$ scales as $1/N$. Linearizing (34), the corresponding RG eigenvalue is -1.

Since $\tilde{e}_\star^2 \sim 1/N$, the leading N behavior of the scalar sector near the fixed points then completely decouples from the gauge sector, and the scalar flow equations reduce to:

$$\frac{\partial}{\partial l}\tilde{m} = 2\tilde{m} + \frac{2}{3\pi^2}N \cdot \frac{\tilde{\rho}_1}{(1+\tilde{m})^2} + \mathcal{O}\left(\frac{1}{N}\right)$$ (35)

$$\frac{\partial}{\partial l}\tilde{\rho}_1 = \tilde{\rho}_1 - \frac{4}{3\pi^2}N \cdot \frac{\tilde{\rho}_1^2}{(1+\tilde{m})^3} + \mathcal{O}\left(\frac{1}{N^2}\right).$$ (36)

These equations have the usual Gaussian and Wilson-Fisher fixed points. At $\tilde{e}_\star^2 > 0$, the Gaussian fixed point of the scalar sector gives rise to the "tricritical fixed point" [48] of the full theory, while the Wilson-Fisher fixed point of the scalar sector determines the critical fixed point of the full theory; the leading large N values of the latter read $\tilde{m}_\star = -1/5$, $\tilde{\rho}_{1\star} = 48\pi^2/(125N)$, with corresponding eigenvalues $\sqrt{5/2}, -\sqrt{5/2}$ [49]. Focusing on the critical fixed point, we can also recover the subleading terms in $1/N$ from a numerical evaluation of the full flow equations. The relevant fixed point values are shown in Table I (right column), and the RG eigenvalues obtained are presented in Table II

TABLE I. Fixpoint values of the stable fixed point (one relevant operator) for the full theory of the flow equations (28)–(32) (left column), and the CP^{N-1} theory obtained by setting $\tilde{\rho}_2 = \tilde{\rho}_3 = 0$ (right column), to next to leading order in $1/N$. The last two quantities, being strongly truncation dependent, are given to leading order only.

	Full theory	CP^{N-1} theory
$\tilde{m}_\star$	$-1/5 + 14/N$	$-1/5 + 13/N$
$\tilde{\rho}_{1\star}$	$48\pi^2/125 \cdot 1/N - 110/N^2$	$48/125\pi^2 \cdot 1/N - 120/N^2$
$\tilde{\rho}_{2\star}$	$-144\pi^2/125 \cdot 1/N^2$	0
$\tilde{\rho}_{3\star}$	$48\pi^2/125 \cdot 1/N + 290/N^2$	0
$\tilde{e}_\star^2$	$54/N$	$54/N$
η	$-14/N$	$-14/N$

(right column). Note that the RG eigenvalues correspond to scaling operators which are linear combinations of the original couplings.

While the qualitative N dependence of all relevant quantities coincides with the findings of Ref. [48], their numerical values are rather distinct. E.g., for the anomalous dimension η, our result $\simeq -14/N$ is quite different from the result $-0.31/N$ given by Ref. [48]. This can be attributed to the fact that we disregarded any contribution to η arising from the scalar sector, since we approach the fixed point from the symmetric phase. As we merely want to clarify if a stable fixed point exists for our full theory, we will overlook these numerical deviations.

Having discussed the CP^{N-1} case we now return to the full theory in question, where $\tilde{\rho}_2, \tilde{\rho}_3 \neq 0$ in general. Solving for zeros of the flow equations numerically, we find a quite similar fixed point structure as before: For small N, four unstable fixed points with vanishing gauge charge are found. When N is increased (above $N \gtrsim 20$), additional unstable zero gauge fixed points are found. For $N > 353$, unstable fixed points at nonzero gauge charge appear. Finally, for $N \geqslant 382$, a stable critical fixed point is found, out of a total of $\mathcal{O}(15)$ fixed points. We can determine its large N properties (semi)analytically as follows: First, we perform a numerical large N scaling analysis, which reveals the same scaling behavior as in the CP^{N-1} case for $\tilde{m}_\star, \tilde{\rho}_{1\star}, \tilde{e}_\star^2$, and furthermore $\tilde{\rho}_{2\star} \sim 1/N^2, \tilde{\rho}_{3\star} \sim 1/N$. We then insert this behavior back into the flow equations and keep the terms that are leading in $1/N$ only. This yields the following result: As before, the beta functions for $\tilde{m}, \tilde{\rho}_1$ decouple and take the forms (35),

TABLE II. RG eigenvalues of the critical fixed point, in next to leading order in $1/N$.

Full theory	CP^{N-1} theory
$\sqrt{5/2} + 20/N$	$\sqrt{5/2} + 18/N$
$-\sqrt{5/2} + 190/N$	$-\sqrt{5/2} + 180/N$
$-1 + 90/N$	$-1 + 80/N$
$-1 - 24/N$	
$-1 + 90/N$	

(36). The relevant solution is the Wilson-Fisher fixed point, with the same leading behavior of the fixed point values and RG eigenvalues as before. The remaining large N form of the flow of $\tilde{\rho}_2, \tilde{\rho}_3$ reads, upon inserting the critical values $\tilde{m}_\star, \tilde{\rho}_{1\star}$:

$$\frac{\partial}{\partial l} \tilde{\rho}_2 = -\tilde{\rho}_2 - \frac{2\tilde{\rho}_3}{N} - \frac{125\tilde{\rho}_3^2}{48\pi^2} + \mathcal{O}\left(\frac{1}{N^3}\right) \quad (37)$$

$$\frac{\partial}{\partial l} \tilde{\rho}_3 = \tilde{\rho}_3 - \frac{125N\tilde{\rho}_3^2}{48\pi^2} + \mathcal{O}\left(\frac{1}{N^2}\right), \quad (38)$$

with relevant fixed point solution

$$\tilde{\rho}_{2\star}^\Lambda = -\frac{144\pi^2}{125N^2}, \quad \tilde{\rho}_{3\star} = \frac{48\pi^2}{125N}, \quad (39)$$

and eigenvalues $(-1, -1)$. The subleading behavior is then again determined numerically and is shown in Tables I and II (left columns). As a result, we find that the scaling properties of the previous critical point of the CP^{N-1} model are only slightly modified by the presence of the two additional irrelevant couplings $\tilde{\rho}_{2\star}, \tilde{\rho}_{3\star}$. Finally we can estimate the correlation length exponent ν, which corresponds to the inverse of the relevant RG eigenvalue and takes the value $\nu = \sqrt{2/5} - \mathcal{O}(1/N) \simeq 0.63$ in the large N limit.

In our simple truncation a stable fixed point only appears for sufficiently large N. This is a well known problem in the RG treatment of gauge theories, which are often plagued by runaway RG flows [14]. However, since the stable CP^{N-1} fixed point does survive in the limit $N \to 2$ when a more involved truncation is used [16,17], we conjecture that the same holds true for our theory in the interesting limit $N \to 3$.

C. Interpretation of the critical fixed point

As stated in the introduction, there is solid evidence that the critical point of the CP^1 model describes the phase transition between the Néel phase and the valence bond solid in square lattice systems. By analogy, it seems natural to associate the critical fixed point found above with a continuous phase transition between the three-color-ordered state and a paramagnet, possibly the trimer state discussed in Sec. II. However, this immediate interpretation is hindered by the fact that our phase diagram in the magnetically ordered phase (see Fig. 3) allows for three different magnetically ordered phases.

As an attempt to resolve this conundrum, we give the following argument: At least in the large N limit, the values of the critical fixed point fulfill the relations (see Table I)

$$\tilde{m}_\star < 0, \quad \tilde{\rho}_{1\star} = \tilde{\rho}_{3\star} \gg |\tilde{\rho}_{2\star}|. \quad (40)$$

These parametric relations will also carry over to the non-rescaled coupling values [see Eq. (26)], at every finite value of the cutoff scale Λ. Comparing with the phase boundaries given in Sec. III 3., we therefore see that, at least on the mean field level, the bulk phase "adjacent" to the critical fixed point is indeed the three-color-ordered phase.

At this point our analysis doesn't make any statement about the structure of the paramagnetic state. In particular, it is not obvious why the paramagnet should be of the trimerized SU(3) singlet type. While for SU(2) the ground state cannot be a trivial disordered paramagnet due to the

Lieb-Schultz-Mattis-Hastings-Oshikawa theorem [8–10], to our knowledge no direct generalization of this theorem to SU(3) magnets in two dimensions is available. In addition, the detailed analysis of Berry phase effects in the SU(2) case shows that the paramagnetic phase breaks lattice symmetries as expected in the VBS phase, which is a strong argument in favor of the dQCP scenario. Even though we do not present an analysis of Berry phase effects for the SU(3) problem in this work, we give a complementary derivation of our critical theory starting from the paramagnetic, trimerized VBS phase in Appendix B. The fact that the same critical theory describes the transition out of both ordered phases provides a strong argument that our theory indeed provides the correct description of the deconfined quantum critical point.

V. CONCLUSIONS AND OUTLOOK

This paper explored the possibility of a deconfined quantum critical point in SU(3) magnets on the triangular lattice. Guided by the analogy to SU(2) magnets on the square lattice, we constructed a critical theory for the continuous transition between a magnetically ordered three-sublattice color-ordered phase (the analog of the Néel phase) and a trimerized SU(3) singlet phase (the analog of the VBS phase). This theory consists of three CP^2 models coupled by quartic interaction terms. Employing the functional renormalization group method in a suitable large N limit, we located a stable critical fixed point, which is not in the universality class of the CP^n model.

Our derivation of the critical field theory starts from the magnetically ordered phase. One drawback of this approach is that the properties of the paramagnetic state are encoded in subtle Berry phase effects, which we did not analyze in this work. However, we provide a strong argument that our field theory correctly describes the deconfined critical point by presenting an alternative derivation of the same critical field theory starting from the paramagnetic VBS phase (details can be found in App. B). Nevertheless, a detailed analysis of Berry phase effects in the spirit of Haldane's work [7], checking that they result in a sextupling of hedgehog events, implying the noncompactness of our critical theory and eventually giving rise to the sixfold degenerate trimer phase, would further substantiate our claim and we leave this problem open for future study.

Additionally, in order to substantiate our understanding of the critical action, it would be interesting to analyze the constraint softening that leads to its final polynomial form. One possible way to achieve this would be to combine the three mutually orthogonal spinors to an SU(3) matrix order parameter. The resulting theory is then a matrix field theory where the orthogonality constraint is rigorously incorporated (see also Ref. [38]).

A more advanced RG study of our critical field theory would be worthwhile as well. Even though our fRG analysis shows that a new stable critical fixed point exists, the fact that it appears only at rather large N is clearly a shortcoming of our simple truncation scheme. Similar problems are well known in the context of the CP^n model, where simple RG approaches give rise to runaway RG flows at small N. Nevertheless, we expect that the large N critical fixed point survives in the limit $N \to 2$. Within fRG this could be analyzed using a

more sophisticated truncation scheme. Finally, one can easily generalize our approach to other interesting $SU(N)$ magnets in two dimensions, such as $SU(4)$ on the square lattice, where a dQCP between a magnetically ordered phase and a quadrimerized singlet phase is possible.

APPENDIX A: DERIVATION OF THE FLOW EQUATIONS

In this Appendix, we present the derivation of the flow equations (28)–(33). Let us first explicitly denote the gauge fixing term by

$$\Gamma_{\text{gf}} = \frac{1}{2\alpha_{\text{gf}}} \sum_i \int d^D x \left(\partial_\mu A_\mu^i\right)^2 . \tag{A1}$$

We will work in the limit $\alpha_{\text{gf}} \to 0$, which fixes the transversal gauge.

Next, we specify the regularization procedure. Following [17], we employ the Litim regulator $\mathcal{R}_\Lambda[\mathbf{A}]$, which will allow for simple analytic evaluations since momentum integrals are rendered trivial. When evaluated at zero gauge field, its scalar and gauge field components in momentum space take the form

$$R_\Lambda^\phi(\mathbf{k}) = Z_\Lambda(\Lambda^2 - k^2)\theta(\Lambda^2 - k^2) \tag{A2}$$

$$R_\Lambda^A(\mathbf{k}) = \frac{1}{e_\Lambda^2}(\Lambda^2 - k^2)\theta(\Lambda^2 - k^2). \tag{A3}$$

At nonzero gauge fields, we need to replace ordinary derivatives by covariant ones in the real-space version of the scalar regulator (A2), but this will only be of relevance for the flow of gauge coupling, to be discussed therein.

Having specified all ingredients to the Wetterich equation (22) (except for the counterterm $\mathcal{C}_\Lambda$, see below), we can compute the running of the couplings. The running of the mass term reads

$$\frac{\partial}{\partial l} m_\Lambda = \frac{1}{2}\text{Tr}\left[\frac{\partial}{\partial l}\mathcal{R}_\Lambda\left(-\hat{G}\frac{\delta^2 \hat{\Gamma}_\Lambda^{(2)}}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})}\hat{G}\right)\right]\Bigg|_{\phi=0,\mathbf{A}=0}, \tag{A4}$$

where the argument of $\phi_{11}, \bar{\phi}_{11}$ denotes zero momentum. The propagator matrix $\hat{G}$ has the following structure in field derivative space

$$\hat{G} = \begin{array}{c} \delta\phi \\ \delta\bar{\phi} \\ \delta A \end{array}\begin{pmatrix} \overset{\delta\phi}{0} & \overset{\delta\bar{\phi}}{G_\phi} & \overset{\delta A}{0} \\ G_\phi & 0 & 0 \\ 0 & 0 & G_A \end{pmatrix}, \tag{A5}$$

with scalar and gauge field components in the transversal gauge

$$G_\phi(\mathbf{p},\beta,j;\mathbf{k},\alpha,i) = \delta_{ij}\delta_{\alpha\beta}\delta(\mathbf{k} - \mathbf{p}) \cdot \frac{1}{Z_\Lambda k^2 + m_\Lambda + R_\Lambda^\phi(k)} \tag{A6}$$

$$G_A(\mathbf{p},\nu,j;\mathbf{q},\mu,i) = \delta_{ij}\delta(\mathbf{q} + \mathbf{p}) \cdot \frac{\delta_{\mu\nu} - q_\mu q_\nu/q^2}{q^2/e_\Lambda^2 + R_\Lambda^A(q)}. \tag{A7}$$

The same field space structure applies to the regulator matrix $\mathcal{R}_\Lambda$, with components as given in (A2). Performing the trace over field space in (A4) results in

$$\frac{\partial}{\partial l} m = \text{tr} \int_{\mathbf{k}} \left[\frac{\partial}{\partial l} R_\Lambda^\phi(\mathbf{k}) \left(-G_\phi(\mathbf{k}) \frac{\delta^2 \Gamma_{\bar{\phi}\phi}(\mathbf{k})}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})} G_\phi(\mathbf{k}) \right) + \frac{1}{2} \frac{\partial}{\partial l} R_\Lambda^A(\mathbf{k}) \left(-G_A(-\mathbf{k}) \frac{\delta^2 \Gamma_{AA}(\mathbf{k})}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})} G_A(-\mathbf{k}) \right) \right] \Bigg|_{\substack{\phi=0 \\ \mathbf{A}=0}},$$

(A8)

where tr denotes the sum over all discrete indices,

$$\int_{\mathbf{k}} \equiv \int \frac{d^3 k}{(2\pi)^3},$$

and $\Gamma_{\bar{\phi}\phi}, \Gamma_{AA}$ are the field space components of the vertex operator $\hat{\Gamma}_\Lambda^{(2)}$ in obvious notation.

It should be noted that diagrams for the mass flow involving three-point vertices are absent in the transversal gauge. The running of the quartic couplings $\rho_{1,2,3}^\Lambda$ can be obtained in analogous manner, using

$$\partial_l \rho_1^\Lambda = \frac{1}{4} \frac{\delta^4 \partial_l \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})} \Bigg|_{\phi=0, \mathbf{A}=0}$$

$$\partial_l \rho_2^\Lambda = \frac{1}{2} \frac{\delta^4 \partial_l \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})\delta\bar{\phi}_{22}(\mathbf{0})\delta\phi_{22}(\mathbf{0})} \Bigg|_{\phi=0, \mathbf{A}=0}$$

$$\partial_l \rho_3^\Lambda = \frac{1}{2} \frac{\delta^4 \partial_l \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{12}(\mathbf{0})\delta\bar{\phi}_{22}(\mathbf{0})\delta\phi_{21}(\mathbf{0})} \Bigg|_{\phi=0, \mathbf{A}=0}.$$

(A9)

For general external indices the required fourfold derivative reads

$$\frac{\delta^4 \partial_l \Gamma_\Lambda}{\delta\bar{\phi}_{\alpha_4 i_4}(\mathbf{0})\delta\phi_{\alpha_3 i_3}(\mathbf{0})\delta\bar{\phi}_{\alpha_2 i_2}(\mathbf{0})\delta\phi_{\alpha_1 i_1}(\mathbf{0})} \Bigg|_{\phi=0, \mathbf{A}=0} = \frac{1}{2} \text{Tr} \left[\frac{\partial}{\partial l} \mathcal{R}_\Lambda \left(\hat{G} \frac{\delta^2 \hat{\Gamma}_\Lambda^{(2)}}{\delta\bar{\phi}_{\alpha_4 i_4}(\mathbf{0})\delta\phi_{\alpha_3 i_3}(\mathbf{0})} \hat{G} \frac{\delta^2 \hat{\Gamma}_\Lambda^{(2)}}{\delta\bar{\phi}_{\alpha_2 i_2}(\mathbf{0})\delta\phi_{\alpha_1 i_1}(\mathbf{0})} \hat{G} + \text{permutations} \right) \right],$$

(A10)

where all possible permutations of the external field derivatives acting on $\hat{\Gamma}_\Lambda^{(2)}$ need to be taken into account. Performing the trace over field space yields, for the flow of ρ_1^Λ, a similar structure as Eq. (A8), see Fig. 4. By contrast, the flow equations for $\rho_{2/3}^\Lambda$ do not pick up any contributions from the gauge sector, which is a consequence of the lattice dependence of the gauge fields.

Similarly, the running of the wave-function renormalization can be extracted from

$$\frac{\partial}{\partial l} Z_\Lambda = \frac{\partial}{\partial p^2} \frac{\delta^2 \partial_l \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{p})\delta\phi_{11}(\mathbf{p})} \Bigg|_{p=0} = \frac{\partial}{\partial p^2} \frac{1}{2} \text{Tr} \left[\frac{\partial}{\partial l} \mathcal{R}_\Lambda \left(\hat{G} \frac{\delta \hat{\Gamma}_\Lambda^{(2)}}{\delta\bar{\phi}_{11}(\mathbf{p})} \hat{G} \frac{\delta \hat{\Gamma}_\Lambda^{(2)}}{\delta\phi_{11}(\mathbf{p})} \hat{G} + \phi_{11} \leftrightarrow \bar{\phi}_{11} \right) \right] \Bigg|_{p=0}$$

$$= \frac{\partial}{\partial p^2} \frac{1}{2} \text{tr} \int_{\mathbf{k}} \left[G_A(\mathbf{k}) \left[\frac{\delta \Gamma_{A\phi}}{\delta\bar{\phi}_{11}(\mathbf{p})} \right](-\mathbf{k}, \mathbf{p}+\mathbf{k}, \mathbf{p}) \, G_\phi(\mathbf{p}+\mathbf{k}) \left[\frac{\delta \Gamma_{\bar{\phi}A}}{\delta\phi_{11}(\mathbf{p})} \right](\mathbf{k}, \mathbf{p}, \mathbf{p}+\mathbf{k}) \, G_A(\mathbf{k})\partial_l R_\Lambda^A(-\mathbf{k}) + \phi_{11} \leftrightarrow \bar{\phi}_{11} \right] \Bigg|_{p=0},$$

(A11)

where the second term in (A11), obtained by permuting the fields, has a similar momentum structure as the first one.

To evaluate the flow equations (A8), (A10), (A11), one needs to insert the appropriate vertex terms. They read

$$\frac{\delta^4 \Gamma_\Lambda}{\delta\bar{\phi}_{\alpha_4 i_4}(\mathbf{0})\delta\phi_{\alpha_3 i_3}(\mathbf{0})\delta\bar{\phi}_{\alpha_2 i_2}(\mathbf{k})\delta\phi_{\alpha_1 i_1}(\mathbf{p})} = \left(2\hat{\rho}_{i_1 i_2 i_3 i_4}^\Lambda \delta_{\alpha_1\alpha_2}\delta_{\alpha_3\alpha_4} + 2\hat{\rho}_{i_1 i_4 i_3 i_2}^\Lambda \delta_{\alpha_2\alpha_3}\delta_{\alpha_1\alpha_4} \right)\delta(\mathbf{k}-\mathbf{p})$$

(A12)

$$\frac{\delta^4 \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{0})\delta\phi_{11}(\mathbf{0})\delta A_\mu^1(\mathbf{k})\delta A_\nu^1(\mathbf{p})} = 2Z_\Lambda \delta_{\mu\nu}\delta(\mathbf{k}+\mathbf{p})$$

(A13)

$$\frac{\delta^3 \Gamma_\Lambda}{\delta\bar{\phi}_{11}(\mathbf{p})\delta\phi_{\alpha i}(\mathbf{q})\delta A_\mu^j(\mathbf{k})} = -Z_\Lambda \delta_{i1}\delta_{j1}\delta_{\alpha 1}\delta(\mathbf{p}-\mathbf{q}-\mathbf{k}) \cdot (p_\mu + q_\mu).$$

(A14)

Inserting these vertices, and the Litim cutoff, all momentum integrations are rendered trivial, and the flow equations are readily computed—see main text, Eqs. (28)–(33).

The flow of the inverse gauge coupling ($1/e_\Lambda^2$) could be derived in a similar manner, by taking an appropriate momentum derivative of the diagram shown in Fig. 4. However,

to avoid ambiguities arising from the sharp Litim cutoff, we instead follow the recipe presented by Reuter and Wetterich in Refs. [46,47]. Adapted to our lattice-dependent gauge field setup, its main idea is as follows: We start from Γ_Λ in real space and evaluate it at a field configuration where $\phi, \bar{\phi}, \mathbf{A}^2, \mathbf{A}^3 = 0$ and the gauge field of sublattice 1 is such that it corresponds

to a constant magnetic field B. Then

$$\Gamma_\Lambda = \int d^3x \frac{1}{4e_\Lambda^2} \cdot 2B^2 = \frac{1}{e_\Lambda^2} \cdot \frac{1}{2} B^2 \Omega, \qquad (A15)$$

where Ω is the system volume. Now, the flow of $1/e_\Lambda^2$ can be obtained by evaluating the right hand side of the Wetterich equation (22) in the field configuration described above, and singling out the coefficient proportional to $B^2\Omega$. By construction of the effective action, on the r.h.s. of (22) the vector potential $\mathbf{A}$ enters only via the squared covariant derivative $-D(\mathbf{A})^2$, where $D(\mathbf{A})_\mu = \partial_\mu - iA_\mu$. Fortunately, the spectrum of $-D(\mathbf{A})^2$ in the given field configuration is explicitly known; it is related to a summation over Landau levels. Using the Euler-McLaurin summation formula to evaluate the sum, the flow of the gauge coupling is then derived as

$$\frac{\partial}{\partial l}\left(\frac{1}{e_\Lambda^2}\right) = \frac{1}{24\pi^2} \sum_{\alpha=1}^N \int_0^\infty dx \, x^{-1/2} \frac{d}{dx}\left(\partial_l R_\Lambda^\phi(x)\right.$$
$$\left. \times \frac{1}{Z_\Lambda x + m_\Lambda + R_\Lambda^\phi(x)}\right) + \frac{\partial}{\partial l}\left(\frac{1}{e_\Lambda^2}\right)\Big|_{\mathcal{C}_\Lambda}, \qquad (A16)$$

where the second summand denotes the contribution from $\mathcal{C}_\Lambda$ specified below. To compute the integral, it is convenient to use the prescription

$$\lim_{\epsilon\to 0}\lim_{\delta\searrow 0} \frac{N}{24\pi^2} \int_\delta^\infty dx \, x^{-1/2} \frac{d}{dx}\left(\frac{\partial_l R_\epsilon(x)}{Z_\Lambda x + m_\Lambda + R_\epsilon(x)}\right), \qquad (A17)$$

where R_ϵ is a version of the Litim cutoff smoothened over a small momentum range ϵ. Then, integrating by parts, the integral is easily solved, and the limits can be taken without problems.

Let's now consider the extra term $\mathcal{C}_\Lambda$. Following [46], an appropriate form of $\mathcal{C}_\Lambda$, which cancels the spurious background-field derivatives from the Wetterich equation, is given by

$$\mathcal{C}_\Lambda[\mathbf{A}] \simeq -\frac{1}{2}\mathrm{Tr}_\phi\left[\log\left(1 + \frac{R_\Lambda^\phi[\mathbf{A}]}{Z_\Lambda\Lambda^2 + m}\right)\right]. \qquad (A18)$$

The notation Tr_ϕ indicates that the trace only extends over the scalar sector in field space; in general, $\mathcal{C}_\Lambda$ also has contributions from the gauge sector, but these are not of relevance for the flow equations at our level of truncation.

One should note that the ansatz Eq. (A18) works best for regulators whose l derivatives are strongly peaked near momenta $\simeq\Lambda$. This does not hold true for the Litim regulator. However, as already mentioned, the error made this way is of two-loop order. To be on the safe side, we have also checked that a Litim adaption of Eq. (A18) with an extra parameter does not modify the qualitative analytical structure of the stable fixed point we are after.

To compute the $\mathcal{C}_\Lambda$ contribution to the running of $1/e_\Lambda^2$, following the logic above we need to evaluate the l derivative of (A18) in the field configuration corresponding to the constant magnetic field on sublattice 1. Again using the Landau-level
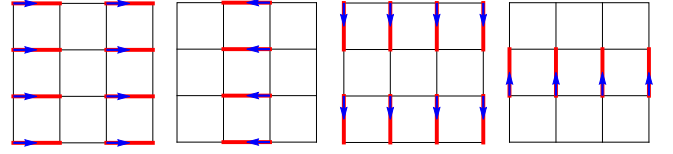
FIG. 5. Sketch of the four degenerate ground states of the SU(2) VBS. The red lines indicate SU(2) singlets, the blue arrows represent the $\mathbb{Z}_4$ order. This picture was adapted from Ref. [11].

summation, this leads to

$$\frac{\partial}{\partial l}\left(\frac{1}{e_\Lambda^2}\right)\Big|_{\mathcal{C}_\Lambda}$$
$$= -\frac{N}{24\pi^2} \frac{\partial}{\partial l} \int_0^\infty dx \, x^{-1/2} \frac{\partial}{\partial x} \log\left(1 + \frac{R_\Lambda^\phi(x)}{Z_\Lambda\Lambda^2 + m_\Lambda}\right). \qquad (A19)$$

This form can easily be evaluated analytically. Rephrasing the so-derived flow equations in terms of dimensionless quantities, one obtains the flow equations as given in the main text, Eqs. (28)–(33).

APPENDIX B: FIELD THEORY FROM THE TRIMER SIDE

In this Appendix we present an alternative construction of the critical theory in Eq. (21) by starting from the trimerized VBS phase. To this end, let us first recapitulate the approach by Levin and Senthil for the SU(2) case on the square lattice [11], which we will follow here.

The columnar SU(2) VBS order can be described by a $\mathbb{Z}_4$ clock order parameter and admits four degenerate ground states, sketched in Fig. 5. The natural field theory that captures this $\mathbb{Z}_4$ order is an XY model with a quartic anisotropy term $\cos(4\theta)$, which is known to be dual to a compact $U(1)$ gauge theory with magnetic monopoles by means of the particle vortex duality [15]. The topological excitations of this model are $\mathbb{Z}_4$ vortices, where four VBS domain walls merge at a single point. On the gauge theory side, they correspond to electrically charged bosons. Proliferating these vortices destroys the VBS order, consequently they are a natural starting point to construct a theory for the deconfined QCP. Levin and Senthil now made the following crucial observations: (i) the vortices always carry a free spin at the core (see the sketch in Fig. 6). Therefore, the corresponding field theory is a modified XY model with an additional spinor structure, or equivalently a $U(1)$ gauge theory of charged bosons z_α with charge $q = 1$, which transform as
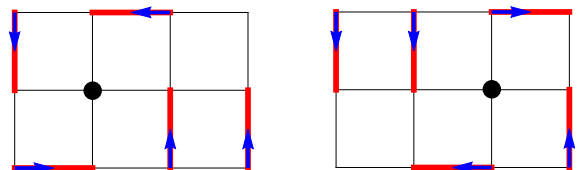

FIG. 6. Sketch of the vortex and antivortex in the SU(2) case. The black circle indicates the vortex core, which carries a free spin. A translation by a lattice unit vector takes a vortex into an antivortex.

spinors under SU(2)

$$z_\alpha \xrightarrow{SU(2)} U_{\alpha\beta} z_\beta \quad z_\alpha \xrightarrow{U(1)} e^{i\phi} z_\alpha. \tag{B1}$$

(ii) the $\mathbb{Z}_4$ structure of the vortices is encompassed in the $\cos(4\theta)$ term. For the standard XY model, this term is known to be irrelevant at the quantum critical point, and it is reasonable to expect that this will be the case for the modified XY model as well. (iii) expanding the action in the spinor fields z_α and their derivatives, one therefore immediately arrives at the standard CP^1 model in Eq. (2).

The Higgs phase of the CP^1 model can then be identified with the Néel state as follows. Vortices are always centered on one sublattice, while antivortices are centered on the other sublattice. A lattice translation by one unit vector thus takes a vortex into an antivortex (see Fig. 6). Both vortices and antivortices must transform under the fundamental representation of SU(2), since both carry a free spin. Furthermore, the antivortex must have $U(1)$ gauge charge $q = -1$. As the vortex is represented by z_α, the antivortex must be represented by $\epsilon_{\alpha\beta}\bar{z}_\beta$, which fulfills both these requirements. Therefore, a translation operation acts as

$$z_\alpha \to \epsilon_{\alpha\beta}\bar{z}_\beta, \tag{B2}$$

flipping the vortex spin. This immediately shows that condensation of vortices results in antiferromagnetic Néel order.

The CP^1 model in Eq. (2), obtained as critical continuum theory from this elegant reasoning, does not make any explicit reference to the sublattice structure. This structure is hidden in the transformation property of the spinors (B2). Note that a lattice translation is equivalent to time reversal in this case (which makes sense since it turns a vortex into an antivortex), and so also flips the spatial components of the gauge field $\mathcal{A}_j \to -\mathcal{A}_j$.

Let's now make the sublattice structure explicit by assigning the field $z_{\alpha 1}$ to the vortex (which sits on sublattice 1) and the field $z_{\alpha 2}$ to the antivortex (which sits on sublattice 2). As a lattice translation takes $z_{\alpha 1}$ to $z_{\alpha 2}$, Eq. (B2) translates to

$$z_{\alpha 2} = \epsilon_{\alpha\beta}\bar{z}_{\beta 1}. \tag{B3}$$

This relation can also be understood from the simple reasoning that it must be possible to break up a SU(2) singlet in the VBS phase into a vortex-antivortex pair or vice versa. In the Lagrangian this process corresponds to a term of the form

$$\varepsilon_{\alpha\beta}z_{\alpha 1}z_{\beta 2} \equiv \bar{z}_{\alpha 1}z_{\alpha 1} \tag{B4}$$

which is indeed an SU(2) singlet and charge neutral under $U(1)$, as required.

When we derived the SU(3) critical theory from the three-color-ordered side in Sec. III 1., we asserted that a SU(2) derivation along the same lines was clearly possible; this derivation results in Eq. (18): $z_{\alpha 2} \sim \epsilon_{\alpha\beta}\bar{z}_{\beta 1}$. That is, one obtains the same relation as in Eq. (B3), but without fixed relative phase. The fact that the relative phase is not fixed will *a priori* lead to two different, sublattice dependent gauge fields $\mathcal{A}_\mu^i$. However, this enlarged gauge redundancy does not lead to different properties of the dQCP in the SU(2) case, because the two $U(1)$ gauge fields couple to the field z_α in precisely the same way after using Eq. (18). To summarize, for the SU(2) case the difference between our derivation and
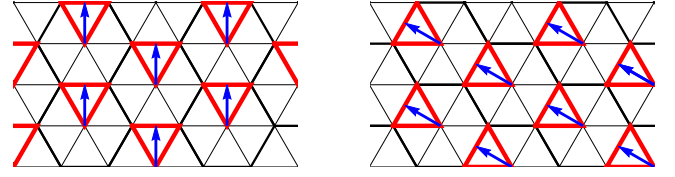

FIG. 7. Two out of six degenerate ground states of the trimer state. The red triangles indicate SU(3) singlets; the blue arrows represent the $\mathbb{Z}_6$ order.

the VBS derivation by Levin and Senthil lies solely in the fact that the relative phase of the fields $z_{\alpha i}$ is fixed by the latter.

Let us now attempt a derivation of our critical theory for the SU(3) case from the trimer side. We will only consider coverings of the triangular lattice with SU(3) singlets which have the same unit cell as the three-color-ordered state (three times the unit cell of the triangular lattice). There are six such coverings, corresponding to six degenerate paramagnetic ground states. Two of them are shown in Fig. 7. Consequently a natural description of this trimer state is a $\mathbb{Z}_6$ clock model, respectively an XY model with sixfold anistropy $\cos(6\theta)$. Again, the anisotropy term is irrelevant at the critical point, so we're not going to consider it further.

In analogy to the SU(2) case it is easy to see that $\mathbb{Z}_6$ vortices carry a free SU(3) spin, as shown in Fig. 8. Interestingly, it is not possible to draw a corresponding antivortex, however. As we will argue below, this is due to the fact that an antivortex transforms under the conjugate representation in the SU(3) case and thus cannot carry a single free SU(3) spin at its core.

The fundamental degrees of freedom in our theory are thus bosonic fields $z_{\alpha j}$ which describe a vortex on sublattice j transforming under the fundamental representation of SU(3), coupled to a $U(1)$ gauge field

$$z_{\alpha j} \xrightarrow{SU(3)} U_{\alpha\beta} z_{\beta j} \quad z_{\alpha j} \xrightarrow{U(1)} e^{i\phi_j} z_{\alpha j}. \tag{B5}$$

As a next step, we need to find an analog of Eq. (B3) for the SU(3) case in order to determine how antivortices are represented in our theory. By the above rationale we should
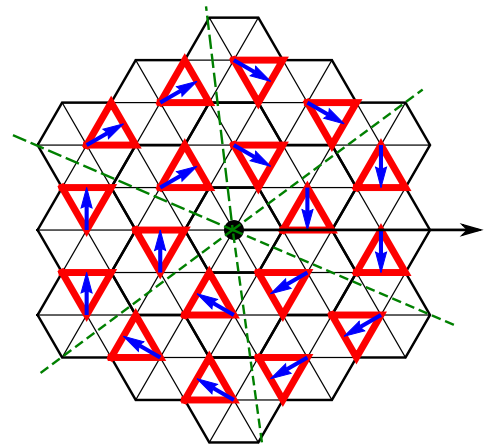

FIG. 8. A vortex $z_{\alpha 1}$ centered on sublattice 1. The black circle represents the free SU(3) spin and the dashed green lines indicate the domain walls.
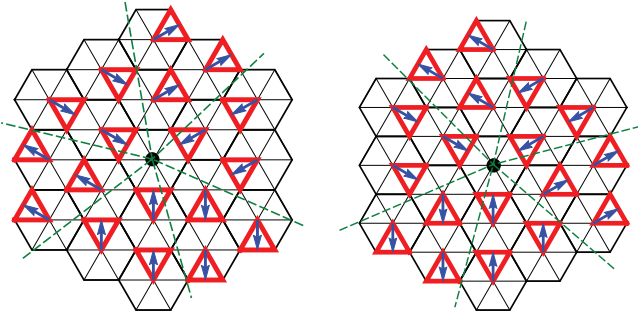
FIG. 9. Two composite vortices centered on sublattice 1, which we represent by the fields $x_{\alpha 1}$ (left) and $y_{\alpha 1}$ (right), respectively. We identify vortices which are obtained from $x_{\alpha i}, y_{\alpha i}$ by space reflection along the axis indicated in Fig. 8 with $x_{\alpha i}, y_{\alpha i}$, accordingly.
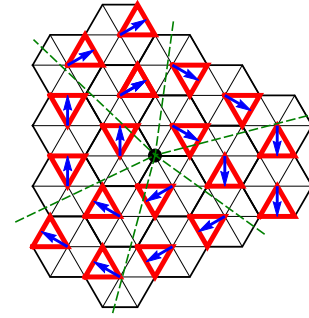

FIG. 10. A $z_{\alpha 2}$ vortex centered on sublattice 2, obtained by a shift of a vortex $y_{\alpha 1}$ on sublattice 1 (see Fig. 9) to the right by one unit vector.

aim at a theory with a fixed relative phase. As we argue below, the direct generalization of (B3) to our case of interest is given by (summation convention over repeated indices is implied)

$$\bar{z}_{\alpha i} = \epsilon_{ijk}\epsilon_{\alpha\beta\gamma} z_{\beta j} z_{\gamma k}, \tag{B6}$$

which only works if we simultaneously demand $|z_{\alpha i}| = 1$ for all three sublattices i. Again, this equation can be understood from the fact that it must be possible to break up a SU(3) singlet in the VBS into three vortices on the respective neighboring sublattice sites or vice versa. The corresponding term in the Lagrangian is then

$$\varepsilon_{ijk}\varepsilon_{\alpha\beta\gamma} z_{\alpha i} z_{\beta j} z_{\gamma k} \equiv \bar{z}_{\alpha i} z_{\alpha i}, \tag{B7}$$

which is clearly an SU(3) singlet, as required. More importantly, this SU(3) singlet also needs to be charge neutral under $U(1)$ transformations. Obviously, this is only possible if we allow for sublattice dependent $U(1)$ gauge transformations. Equation (B6) thus enforces a partial gauge fixing condition $\phi_1 + \phi_2 + \phi_3 = 0$. Note that Eq. (B6) ensures that the spinors on the three sublattices are mutually orthogonal. The condensation of vortices thus automatically leads to a three-sublattice color ordered phase, as discussed in the main text. Moreover, Eq. (B6) is in accordance with SU(3) transformation properties of the spinors, i.e., $\bar{z}_{\alpha j}$ transforms under the conjugate representation of SU(3).

In the following we present a more microscopic justification of Eq. (B6) by deriving it from the lattice transformation properties of all topological excitations which can be represented pictorially. In addition to the vortex of Fig. 8, we can draw two "composite vortices" on each sublattice, which we denote by $x_{\alpha i}, y_{\alpha i}$, respectively. An example is depicted in Fig. 9. Due to the discrete nature of the $\mathbb{Z}_6$ order parameter, one cannot assign a unique vorticity (or gauge charge) q to these objects. The most natural choice seems to be $q = -2$, but this assignment is ambiguous as, e.g., the vorticity is not reversed if the vortex is encircled clockwise.

Analogous pictures can be drawn for vortices centered on sublattices 2,3, which can be obtained from Fig. 9 by translation operations. One such operation is sketched in Fig. 10. A highly interesting feature of these pictures is that, opposed to the SU(2) case, one can never draw an antivortex as an elementary object carrying a free SU(3) spin.

Let us now try to parametrize the composite vortices x, y in terms of the elementary one. Up to irrelevant constant prefactors, the only meaningful ansatz is

$$x_{\alpha i} = y_{\alpha i} = \epsilon_{ijk}\epsilon_{\alpha\beta\gamma} \bar{z}_{\beta j} \bar{z}_{\gamma k}, \tag{B8}$$

which is consistent with transformations under the fundamental representation of SU(3).

Next, we need to take into account the transformation properties of the vortices. Drawing all possible pictures, they read

$\hat{R}$ (Rot. by $\pi/3$ with base point on sublattice 1): $\qquad$ (B9)

$$z_1 \leftrightarrow z_1, \quad x_1 \leftrightarrow y_1, \quad z_2 \leftrightarrow z_3, \quad x_2 \leftrightarrow y_3, \quad y_2 \leftrightarrow x_3$$

$\mathcal{T}_1$ (transl. along the first lattice direction): $\qquad$ (B10)

$$z_1 \to x_2 \to y_3 \to z_1$$
$$x_1 \to y_2 \to z_3 \to x_1$$
$$y_1 \to z_2 \to x_3 \to y_1$$

$\mathcal{T}_2$ (transl. along the second lattice direction): $\qquad$ (B11)

$$z_1 \to y_3 \to x_2 \to z_1$$
$$x_1 \to z_3 \to y_2 \to x_1$$
$$y_1 \to x_3 \to z_2 \to y_1.$$

Here, we have supressed the SU(3) indices. Along with reflection along the x axis, which acts trivially on the vortices per definition (see caption of Figs. 8 and 9), the operations $\hat{R}, \mathcal{T}_{1/2}$ span the symmetry group of the triangular lattice. Note that the above transformations hold up to constant prefactors, which cannot be deduced from the pictures. One must now reconcile the transformation properties ((B9)–(B11)) with the definition of the composite vortices (B8). It is found that the only way to do so is precisely to demand the orthogonality constraint with fixed relative phase (B6). Writing down a gauge

theory consistent with (B6), one then recovers Eq. (17) along with the relative gauge fixing constraint. Again, softening the

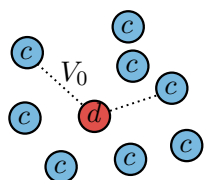unit length constraint for the fields $z_{\alpha j}$ we arrive at the theory in Eq. (21).

---

[1] T. Senthil, L. Balents, S. Sachdev, A. Vishwanath, and M. P. A. Fisher, Phys. Rev. B **70**, 144407 (2004).

[2] T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, and M. P. A. Fisher, Science **303**, 1490 (2004).

[3] S. Sachdev, Nat. Phys. **4**, 173 (2008).

[4] N. Read and S. Sachdev, Phys. Rev. Lett. **62**, 1694 (1989).

[5] N. Read and S. Sachdev, Phys. Rev. B **42**, 4568 (1990).

[6] A. W. Sandvik, Phys. Rev. Lett. **98**, 227202 (2007).

[7] F. D. M. Haldane, Phys. Rev. Lett. **61**, 1029 (1988).

[8] E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. **16**, 407 (1961).

[9] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).

[10] M. Oshikawa, Phys. Rev. Lett. **84**, 1535 (2000).

[11] M. Levin and T. Senthil, Phys. Rev. B **70**, 220403 (2004).

[12] O. I. Motrunich and A. Vishwanath, Phys. Rev. B **70**, 075104 (2004).

[13] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, Cambridge, 2011).

[14] B. Halperin, T. Lubensky, and S.-K. Ma, Phys. Rev. Lett. **32**, 292 (1974).

[15] C. Dasgupta and B. I. Halperin, Phys. Rev. Lett. **47**, 1556 (1981).

[16] B. Bergerhoff, D. F. Litim, S. Lola, and C. Wetterich, Int. J. Mod. Phys. A **11**, 4273 (1996).

[17] L. Bartosch, Phys. Rev. B **88**, 195140 (2013).

[18] Y. Huh, P. Strack, and S. Sachdev, Phys. Rev. Lett. **111**, 166401 (2013).

[19] R. G. Melko and R. K. Kaul, Phys. Rev. Lett. **100**, 017203 (2008).

[20] A. W. Sandvik, Phys. Rev. Lett. **104**, 177201 (2010).

[21] A. Banerjee, K. Damle, and F. Alet, Phys. Rev. B **82**, 155139 (2010).

[22] H. Shao, W. Guo, and A. W. Sandvik, Science **352**, 213 (2016).

[23] A. B. Kuklov, M. Matsumoto, N. V. Prokof'ev, B. V. Svistunov, and M. Troyer, Phys. Rev. Lett. **101**, 050405 (2008).

[24] K. Chen, Y. Huang, Y. Deng, A. B. Kuklov, N. V. Prokof'ev, and B. V. Svistunov, Phys. Rev. Lett. **110**, 185701 (2013).

[25] R. K. Kaul and A. W. Sandvik, Phys. Rev. Lett. **108**, 137201 (2012).

[26] K. Harada, T. Suzuki, T. Okubo, H. Matsuo, J. Lou, H. Watanabe, S. Todo, and N. Kawashima, Phys. Rev. B **88**, 220408 (2013).

[27] R. K. Kaul and M. S. Block, J. Phys.: Conf. Ser. **640**, 012041 (2015).

[28] N. Papanicolaou, Nucl. Phys. B **305**, 367 (1988).

[29] C. Honerkamp and W. Hofstetter, Phys. Rev. Lett. **92**, 170403 (2004).

[30] A. Gorshkov, M. Hermele, V. Gurarie, C. Xu, P. Julienne, J. Ye, P. Zoller, E. Demler, M. Lukin, and A. Rey, Nat. Phys. **6**, 289 (2010).

[31] S. Taie, Y. Takasu, S. Sugawa, R. Yamazaki, T. Tsujimoto, R. Murakami, and Y. Takahashi, Phys. Rev. Lett. **105**, 190401 (2010).

[32] S. Taie, R. Yamazaki, S. Sugawa, and Y. Takahashi, Nat. Phys. **8**, 825 (2012).

[33] C. Hofrichter, L. Riegger, F. Scazza, M. Höfer, D. R. Fernandes, I. Bloch, and S. Fölling, Phys. Rev. X **6**, 021030 (2016).

[34] A. Läuchli, F. Mila, and K. Penc, Phys. Rev. Lett. **97**, 087205 (2006).

[35] B. Bauer, P. Corboz, A. M. Läuchli, L. Messio, K. Penc, M. Troyer, and F. Mila, Phys. Rev. B **85**, 125116 (2012).

[36] S. Bieri, M. Serbyn, T. Senthil, and P. A. Lee, Phys. Rev. B **86**, 224409 (2012).

[37] H.-H. Lai, Phys. Rev. B **87**, 205131 (2013).

[38] A. Smerald and N. Shannon, Phys. Rev. B **88**, 184430 (2013).

[39] M. Aguado, M. Asorey, E. Ercolessi, F. Ortolani, and S. Pasini, Phys. Rev. B **79**, 012408 (2009).

[40] H. T. Ueda, Y. Akagi, and N. Shannon, Phys. Rev. A **93**, 021606 (2016).

[41] A. Auerbach, *Interacting Electrons and Quantum Magnetism*, 1st ed. (Springer, New York, 1994).

[42] A. V. Chubukov, S. Sachdev, and T. Senthil, Nucl. Phys. B **426**, 601 (1994).

[43] P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the Functional Renormalization Group*, Lect. Notes Phys. 798 (Springer, Heidelberg, 2010).

[44] H. Gies, in *Renormalization Group and Effective Field Theory Approaches to Many-Body Systems* (Springer, Berlin, 2012), pp. 287–348.

[45] M. Reuter and C. Wetterich, Nucl. Phys. B **391**, 147 (1993).

[46] M. Reuter and C. Wetterich, Nucl. Phys. B **427**, 291 (1994).

[47] M. Reuter and C. Wetterich, Nucl. Phys. B **408**, 91 (1993).

[48] B. Bergerhoff, F. Freire, D. F. Litim, S. Lola, and C. Wetterich, Phys. Rev. B **53**, 5734 (1996).

[49] Note that the textbook eigenvalues $(-1, 1)$ [13] found via ϵ expansion are recovered if we expand all equations in $\tilde{m}$.

5 Molecules, Polarons, and Polaritons

5.1 Introduction

5.1.1 The Fermi polaron problem

Let us consider the following innocent looking problem in two dimensions: a single impurity, represented by operators $d, d^\dagger$, is immersed in a Fermi sea of majority particles, with operators $c, c^\dagger$. The two particle species interact via a short-ranged attractive interaction of strength V_0. This setup is sketched below, and the impurity Hamiltonian is readily written down:

$$H_{\text{imp}} = \sum_{\mathbf{k}} \left(\epsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + E_{\mathbf{k}} d_{\mathbf{k}}^\dagger d_{\mathbf{k}} \right) - \frac{V_0}{\mathcal{S}} \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}-\mathbf{q}} d_{\mathbf{p}}^\dagger d_{\mathbf{p}+\mathbf{q}} . \qquad (5.1)$$

Here, $\mathcal{S}$ is the system area, and we assume parabolic dispersions for the particles, $\epsilon_{\mathbf{k}} = \mathbf{k}^2/2m$, $E_{\mathbf{k}} = \mathbf{k}^2/2M + E_0$, with an energy offset (gap) E_0 for the impurity. Note that the condition that there is one impurity only is not yet apparent from the Hamiltonian (5.1), but can easily be imposed when computing correlation functions as discussed below.

The impurity problem is characterized by the competition of two energy scales: the kinetic energy that is determined by the chemical potential μ of the c-particles, and the potential energy. The latter is most conveniently parametrized by the energy $-E_{\text{b}}$ of the bound state between a c and d particle in the two-particle limit, which microscopically depends on V_0 and a UV cutoff. Such a bound state can always be formed in dimensions $d = 2$ for an attractive interaction potential, no matter the value of V_0. By tuning the ratio μ/E_{b}, qualitatively different states of the system can be accessed, as indicated in Fig. 5.1:
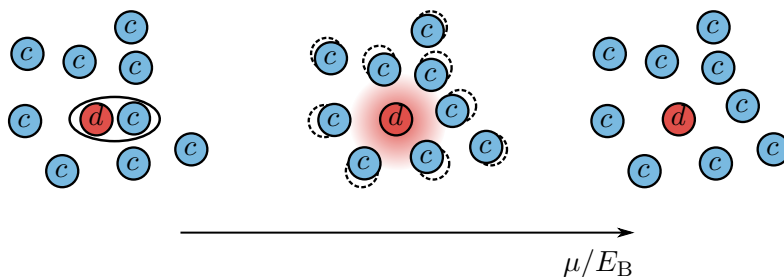
Figure 5.1 Evolution of the many-body ground state from the molecular to the polaronic and free-particle state. Fig. is adapted from Ref. [SWSZ09].

For $\mu/E_{\text{b}} \ll 1$, the impurity tightly binds a majority particle, forming a "molecule". For increasing values of μ/E_{b}, the molecule eventually dissolves, and the impurity is surrounded by a dressing cloud of c-particles. The renormalized impurity state formed in this way is called "polaron" (a polarization quasiparticle) by analogy to the standard Froehlich-type polarons formed when electrons are coupled to phonons in a solid (see, e.g., [Mah00]). Finally, for $\mu/E_{\text{b}} \gg 1$, interactions are insignificant, and the impurity is essentially free. While

the polaronic state is smoothly connected to the free particle state, the evolution towards the molecule potentially occurs via an interesting quantum phase transition. Again, this transition defies a simple order parameter description, as will be discussed in Sec. 5.4.

In the context of this thesis, we will be interested in two different experimental realizations of the Hamiltonian (5.1), which can both tune the ratio μ/E_b in their own way:

○ **Optically excited semiconductors**: In a doped semiconductor illuminated with visible light, a valence band hole ($\hat{=} d$) can be created upon absorption of a single optical photon. This hole interacts with the conduction band Fermi sea via a screened Coulomb interaction ($\hat{=} V_0$). By applying a gate voltage, one can modify the chemical potential μ in the conduction band, and thus tune μ/E_b. Such experiments can e.g. be peformed on 2D GaAS quantum wells [SWH$^+$14]. Furthermore, if the Coulomb interaction is very strong, the valence band hole tightly binds a conduction band electron, forming an exciton [WCG$^+$18]. In the transition metal-dichalcogenides (TMDs), which are layered 2D materials, this exciton has an enormous binding energy $\simeq .5\text{eV}$. In contrast, the relevant Fermi energy in the doped systems is $\simeq 30$ meV [SBC$^+$17]. Therefore one can view the exciton as rigid impurity with a residual interaction with the conduction band Fermi sea.

As an interesting experimental extension, the semiconductor structure can be inserted into a microcavity [SWH$^+$14, GPC$^+$07, SBC$^+$17], and strong coupling to a cavity photon mode gives rise to a new light-matter quasiparticle, the polariton (see Sec. 5.2).

○ An alternative to the solid state experiments are **ultracold gas** setups [SWSZ09, KZJ$^+$12, KPV$^+$12, CJL$^+$16, SVM$^+$17, DORB$^+$19], where the impurity and majority particles are typically two distinct hyperfine states of an atom. Both two and three-dimensional systems have been studied, as well as different ratios of impurity/majority mass m/M. In these experiments, the scattering length a (strictly positive in $2D$) and hence the binding energy $E_b \propto 1/a^2$ can be adjusted by tuning through a Feshbach-resonance [CGJT10]. For "broad" Feshbach resonances, the momentum dependence of the effective interaction can be disregarded and the systems is adequately described by a contact interaction. Ultracold gases show a particularly clean molecule-to-polaron evolution in the spectral features, typically observed by RF-spectroscopy (see, e.g., [SKI$^+$18]).

Of course, in both these experimental platforms factually a finite number of impurities is created, but the ratio impurity/majority is still small enough to warrant a a single-impurity description. This simplifies the theoretical analysis; still, the problem is a many-body one, and in general a rigorous understanding of the impurity behaviour is out of reach apart from the trivial cases $\mu/E_b \to \infty$ or $\mu/E_b \to 0$. However, one interesting solvable limit of the problem does exist: An infinitely heavy impurity with $M \to \infty$ has no internal structure (no momentum quantum numbers). As a result, the infinite mass impurity is described by a single particle scattering problem, and correlations can be obtained exactly in scattering theory language. This was first accomplished in the late sixties in a seminal work by Nozières and de Dominicis [NDD69], who studied the impurity problem in the context of the X-ray absorption in metals. In addition, a diagrammatic solution of the infinite mass problem based on the summation of parquet diagrams exists, also due to Nozières and coworkers [RGN69, NGR69].

In the two publications that mainly constitute this chapter [PvDGG17, PG18], we have taken the diagrammatic infinite mass solution as a starting point, and generalized it to large but finite masses. This is done in two different physical and technical contexts: In the first paper (see Sec. 5.2), which was initiated during my master's thesis [Pim15], we computed the two-particle propagator $\propto \langle cdd^\dagger c^\dagger \rangle$ in order to understand the optical mesurements in GaAs [SWH$^+$14], where the ratio of conduction and valence band mass m/M is indeed small

($\simeq 0.14$). Polariton formation in a microcavity is discussed as well. In the second publication (Sec. 5.3), we computed the impurity single-particle spectrum $\langle dd^\dagger \rangle$, with both the ultracold gases and the TMD setup in mind. While the limit of heavy mass is not quite fulfilled in these experiments, we nethertheless believe that our analysis captures the main physics, without relying on uncontrolled approximations as previous works.

Both these publications mostly deal with analytically controlled limits $\mu/E_b \ll 1$ or $\mu/E_b \gg 1$, and cannot make robust claims about the molecule-polaron transition. A progress report on ongoing work in this direction is presented in Sec. 5.4.

5.1.2 Finite mass modification of edge singularities

As a guide to the publications, let us give a non-technical overview over the impurity correlations in the infinite mass limit, and the finite mass modifications thereof.

For $M \to \infty$, the momentum labels of the impurity operators can be dropped, and the Hamiltonian reduces to

$$H[M = \infty] = \sum_{\mathbf{k}} \left(\epsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + E_0 d^\dagger d \right) - \frac{V_0}{\mathcal{S}} \sum_{\mathbf{k}, \mathbf{p}} c_{\mathbf{k}}^\dagger c_{\mathbf{p}} d^\dagger d \ . \tag{5.2}$$

To begin with, we assume that the energy of the bound state $(-E_b)$ is negligible, i.e. $E_b \ll \mu$. We can therefore consider an expansion of the correlation functions directly in V_0, or rather the dimensionless parameter $g = \rho V_0$ where $\rho = m/2\pi$ is the 2D density of states of the conduction band electrons. For the contact potential $E_b \simeq \mu \exp(-1/g),$[1] and thus $g \simeq 1/\log(\mu/E_b)$. For illustration, we now consider the retarded optical response function $\Pi(t) = -i\theta(t) \langle 0|d(t)c(t)c^\dagger(0)d^\dagger(0)|0\rangle$ in real time and $T = 0$, or respectively its Fourier transform $\Pi(\omega)$. Here, $|0\rangle$ is the ground state of the system without impurity (Fermi sea), which explains why the response function does not involve a commutator.

When constructing diagrams, a key ingredient is the automatic retardation of the impurity propagator: since there is no impurity Fermi sea, impurities can only propagate forward in time (corresponding to retarded impurity propagators in frequency space). The diagrams for Π constructed with this recipe up to second order in g, but without self energy insertions, are shown in Fig. 5.2.
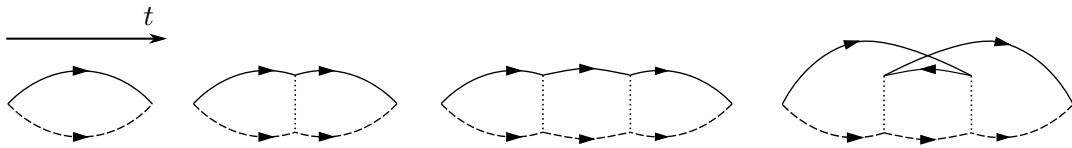
Figure 5.2 Diagrams for the optical response funciton Π. Full (dashed) lines indicated majority (impurity) particles, and dotted lines the interaction.

As easily shown, in frequency space these diagrams are IR divergent as a function of the external frequency ω (note that the diagrams are momentum independent for $M \to \infty$), yielding different powers of $L = -\log(|\omega|/\mu) \gg 1$.[2] The divergences arise from the parallel and antiparallel "bubbles" that make up the diagrams, with the latter appearing first in the diagram on the very right.[3] These bubbles effectively each contribute a factor of gL. The "parquet" diagrams that contain all leading divergences can be constructed by succesively

---

1 If the UV cutoff is of order μ.

2 The correct energy thresholds can e.g. be found with Fumi's theorem [Mah00], but here we will not discuss this further.

3 Note that the labels "parallel" and "antiparallel" interchange if one employs impurity electron operators instead of the impurity hole operators we have tacitly been using here.

replacing all interaction vertices by any of the bubbles [RGN69] and their proper resummation thus yields a consistent approximation to Π. Non-parquet diagrams, e.g. self-energy insertions, scale as $g^n L$ with $n > 1$, and are thus subleading when evaluating Π in the small coupling limit $g \ll 1$.

Taking into account the parquet diagrams only, one can write down two coupled Bethe-Salpeter equations for the interaction vertex, which can be solved with "logarithmic accuracy" by transformation to differential equations. Importantly, one finds that the logarithmic singularities of the two bubble channels partially cancel each other. Other solution possibilities besides the Bethe-Salpeter approach exist as well, for instance the recently introduced numerical multiloop functional renormalization group method [KvD18]. In any case, the imaginary part of Π obtained this way reads

$$\text{Im}[\Pi](\omega) \propto - \left(\frac{\omega}{\mu}\right)^{-2g} \theta(\omega) = -\theta(\omega) \left(1 - 2g \log\left(\frac{\omega}{\mu}\right) + \mathcal{O}(g^2)\right) . \tag{5.3}$$

This prototypical asymmetric power law divergence is known as "Fermi-edge singularity". Ultimately, this singularity roots in the "Anderson Orthogonality Catastrophe" [And67], i.e. the orthogonality of the many-body groundstate wavefunctions with and without impurity. We have also indicated the expansion of the power law in the weak coupling limit; the shown two terms arise from the first two diagrams of Fig. 5.2.

The exact solution of the corresponding (time-dependent) scattering problem [NDD69] yields a modified power law exponent $(\delta(\mu)/\pi)^2 - 2\delta(\mu)/\pi$, where $\delta(\mu)$ is the scattering phase shift of Fermi level electrons in reaction to the point-like impurity potential g. One can show that $\delta(\mu)/\pi \simeq g$ and thus $(\delta(\mu)/\pi)^2 - 2\delta(\mu)/\pi \simeq -2g$ at weak coupling, where the diagrammatic and exact solutions agree.

For the impurity Green's function $D(t) = -i\theta(t) \langle 0|d(t)d^\dagger(0)|0\rangle$, a similar, but somewhat more complicated summation technique [NGR69] or alternatively a "linked cluster method" ($\hat{=}$ reexponentiation, see [Mah00]) leads to

$$\text{Im}[D](\omega) \propto - \left(\frac{\omega}{\mu}\right)^{g^2-1} \theta(\omega) = -\theta(\omega)\frac{\mu}{\omega} \left(1 + g^2 \log\left(\frac{\omega}{\mu}\right) + \mathcal{O}(g^4)\right), \tag{5.4}$$

while in the exact solution g is replaced by $\delta(\mu)/\pi$. One can call this dressed impurity a polaron, albeit without a quasi-particle weight in the infinite mass case.

For $g \ll 1$, the power law (5.3) is a "weak one", converging to a step-function as $g \to 0$, while the "strong" power law in (5.4) becomes a delta function (if some approriate prefactors $\propto g$ are attached). Representative sketches are shown in Fig. 5.3 (a), (b), with $\alpha \simeq g$. Note that we do not know the precise form of the non-universal UV behaviour $\omega \simeq \mu$.

What happens for large but finite impurity masses, $\beta \equiv m/M \ll 1$? The new physical ingredient is the **impurity recoil**, which effectively provides an IR cutoff to the logarithms via a self-energy effect [GNRC69]. The "master formula" is

$$\boxed{\log\left(\frac{|\omega|}{\mu}\right) \xrightarrow{M<\infty} \log\left(\frac{|\omega|}{\mu} + \beta\right)} . \tag{5.5}$$

Replacing the IR divergent logarithms in the expressions (5.3), (5.4) allows to understand the gist of the finite mass behaviour, in agreement with results obtained with various other methods by several authors [GNRC69, RSR87, US90, Haw91, Pro93, Noz94, RK95, Ros99]. Note that Eq. (5.5) also shows that one cannot properly describe the finite mass problem by expanding in β around $\beta = 0$, since this produces bogus results for the interesting range $\omega \lesssim \beta\mu$.
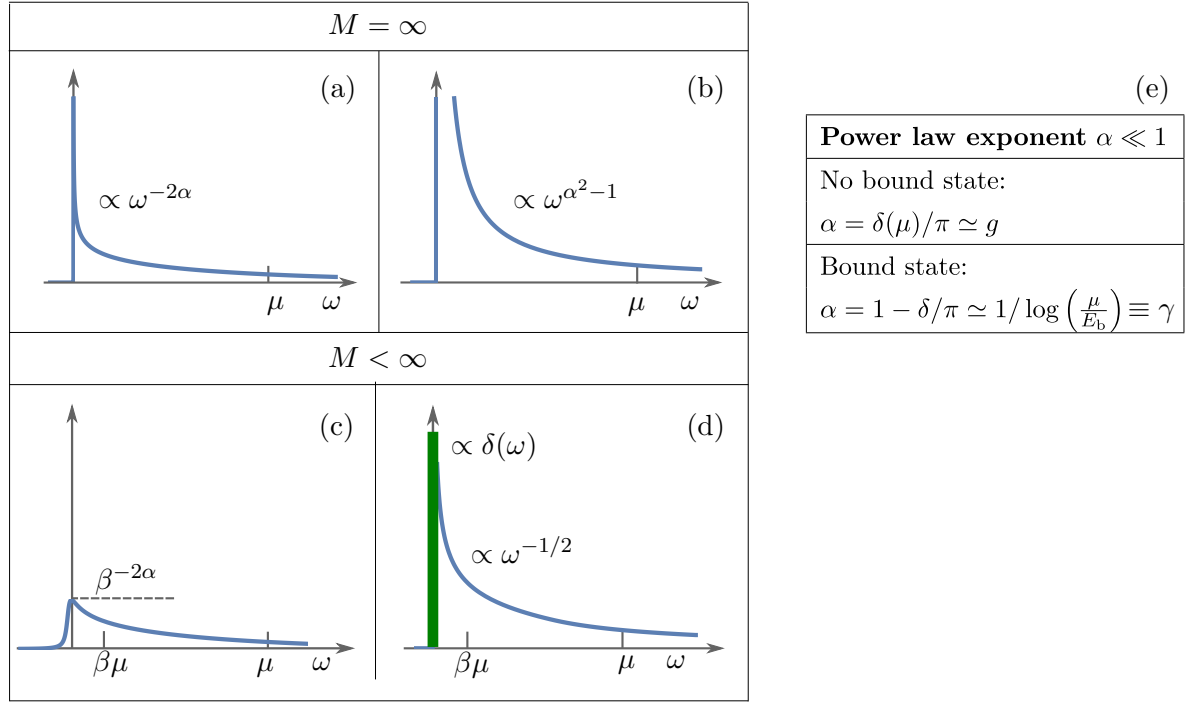
Figure 5.3 Most important finite mass modifications of the edge singularities. Figs. (a) and (b) show "weak" and "strong" infinite mass power laws. Fig. (c) (Fig. (d)) shows the finite mass modifcation of Fig. (a) (Fig. (b)). The table in Fig. (e) relates the weak coupling coefficient α to physical quantities in the absence or presence of a bound state. For more details, see main text.

First, let us apply our master formula to Π, (5.3): If the logarithms are cut off in this equation, no divergence remains at all, and the edge singularity becomes a step-like spectral feature with maximal height $\propto \beta^{-2g}$. A sketch is shown in Fig. 5.3(c): for small energies $0 < \omega < \beta\mu$, the infinite mass behaviour is modified, and the singularity is cut off. For much larger energies $\omega \gg \beta\mu$, in all edge singularity problems the infinite mass behaviour is recovered. Note that this general structure is not dependent on the dimensionality (as is the infinite mass problem, which is effectively zero-dimensional). Some more universal details of the spectrum (dependent on the dimensionality) can also be extracted from phase space arguments, as shown in our publications [PvDGG17, PG18].

For the strong singularity of Eq. (5.4), the finite mass effects are rather different: As can be seen from the expansion in small g, when the logarithms are cut off the singularity becomes even stronger, ω^{-1} instead of ω^{g^2-1}! In fact, a more careful analysis reveals that the impurity spectrum has a delta peak with a finite weight of order β^{g^2}, on top of an incoherent background which scales a $\omega^{-1/2}$ in 2D and $\simeq$ const. in 3D [Ros99]. This behaviour originates from an impurity self-energy of order $\omega^{3/2}(\omega^2$ in 3D), giving rise to a delta peak due to the frequency power that is larger than one, compare also our discussion around Eq. (3.30) in Sec. 3.1.5. A sketch of the finite mass spectrum is shown in Fig. 5.3(d). Note that the finite mass effect again pushes the spectrum towards its non-interacting limit $g \to 0$, where one would obtain a pure delta function.

Now, let us reinstall the bound state. This is both important in the semiconductor context, where the bound state corresponds to the exciton, and to desribe ultracold gases, which is done in terms of a renormalized interaction (parametrized by the scattering length a) in lieu of the bare interaction V_0 we have used so far. For the non-perturbative infinite mass methods [CN71, SKI$^+$18], the transition to the bound state can be made straigthforwardly. The main findings are as follows: First, the bound state appears in the spectra at negative

energies, and so there are two divergent spectral features now (a primary and a secondary one). Second, the value of the phase shifts $\delta(\mu)$ that control the power laws change (see, e.g., [RDS90]). In particular, according to Levinson's theorem, in the presence of a single bound state the scattering phase shift of a low-energy scattered state is shifted by π.

Unfortunately, the non-perturbative methods cannot be generalized to finite mass, because they sensitively rely on the zero-dimensional character of the infinite mass impurity. One must again appeal to diagrammatics, but this turns out be tricky: to resolve the bound state with energy $(-E_{\rm b})$, we can no longer expand in g as done in Fig. 5.2, but must resum all "ladder diagrams" instead, i.e. all the diagrams containing parallel bubbles only (see Fig. 5.2). This is in accordance with the usual regularization of the bare potential V_0 via the scattering length a by summing the "T-matrix" [Pun10]. Thus, we must reorder the diagrammatic expansion, and identify a new weak coupling parameter. As we show in [PvDGG17, PG18], this can be achieved in the limit $E_{\rm b} \gg \mu$, where the phase shift fullfills $\delta(\mu) \lesssim \pi$. For example, the primary divergence of the impurity correlator in the infinite mass limit scales as [CN71] $\omega^{(\delta(\mu)/\pi)^2-1} = \omega^{(1-\delta(\mu)/\pi-1)^2-1} = \omega^{(1-\delta(\mu)/\pi)^2-2(1-\delta(\mu)/\pi)}$. For $E_{\rm b} \gg \mu$, one finds $1 - \delta(\mu)/\pi \simeq 1/\log(E_{\rm b}/\mu) \equiv \gamma \ll 1$ [RDS90], which mirrors the previous case of $\mu \gg E_{\rm b}$ where we had $g = 1/\log(\mu/E_{\rm b}) \ll 1$. Thus, we should utilize γ as expansion parameter. While γ does not directly appear in the Hamiltonian, it is small in $\mu/E_{\rm b}$, which does suggest a meaningful diagrammatic expansion: for every conduction band hole in a Feynman diagram, one picks up an integration over μ. Thus, we must systematically sum diagrams with an increasing number of holes as shown in Fig. 5.4.
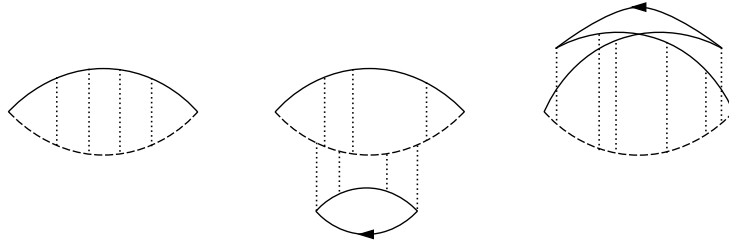
Figure 5.4 Important optical response diagrams with up to one hole, which is indicated by the backwards arrow. Interaction lines connect impurity and majority electrons (but not the majority hole) in arbitrary fashion.

Although the computation of these diagrams is more involved due to the complex interaction structure, it is feasible, and one can indeed extract the leading logarithms of the infinite mass power laws $\propto \gamma \log(\omega/\mu)$. This computation can then be carried over to finite mass, and the modifications essentially parallel the ones in the limit $\mu \gg E_{\rm b}$, only with a modified interaction parameter (γ instead of g).

In short, one can therefore summarize the finite mass modifications of edge singularities in 2D as follows: there are two controlled limits, $\mu \gg E_{\rm b}$, with control parameter $g \simeq 1/\log(\mu/E_{\rm b})$, and $E_{\rm b} \gg \mu$, with control parameter $\gamma \simeq 1/\log(E_{\rm b}/\mu)$. In both cases, the infinite mass spectra resemble the trivial limts $g, \gamma \to 0$ if finite masses are switched on. We expect that this is fulfilled in 3D as well, although the precise form of the control parameter might change [PG18].

In our publications, we show that this simple rule of thumb has profound physical implications, e.g. calling the standard view on the "attractive polaron" into question (see Sec. 5.3).

5.2 Exciton-Polaritons and the Fermi-edge singularity

5.2.1 Overview

In the 2D semiconducting systems, not only the Fermi impurity physics on its own with all the associated spectral manifestations of Fig. 5.3 can be realized, but there is another physical dimension to be added in the experiments: a strong coupling of the matter excitations to light in a microcavity. The basic setup is sketched in Fig. 5.5.
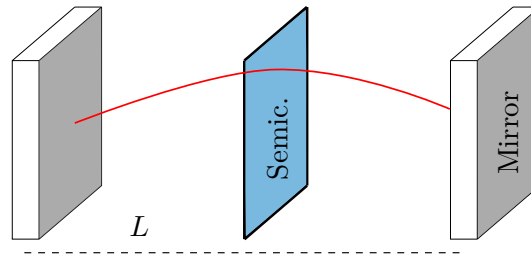

Figure 5.5 In an optical cavity composed of two mirrors at a distance L, standing light waves (cavity photons) exist as eigenstates. The electrical field strength of the one lowest cavity eigenmode is indicated by the red line. The cavity photons strongly interact with the matter excitations of the 2D semiconductor. Fig. is adapted from my master's thesis [Pim15].

With some standard (rotating wave and dipole-) approximations, a Hamiltonian for the combined light-matter system in 2D is easily derived (see, e.g. , [H. 09, Pim15]):

$$H_{\text{light-matter}} = H_{\text{imp}} + \omega_c(L)a^\dagger a - iM_0 \sum_{\mathbf{p}} ac^\dagger_{-\mathbf{p}} d^\dagger_{\mathbf{p}} \quad + \quad \text{h.c.} \tag{5.6}$$

Here, H_{imp} is the impurity Hamiltonian from Eq. (5.1), M_0 is the dipole matrix element, and a are the operators representing the lowest cavity eigenmode with vanishing 2D in-plane momentum. Other cavity modes can mostly be neglected – they are far off in energy due to their steep dispersion. The energy of the relevant mode $\omega_c(L)$ can be tuned by modifying the effective length of the cavity L.

When close in energy, cavity photons can hybridize with a matter resonance, resulting in an avoided crossing of two "polariton" branches. This is straightforward if the matter excitation in question is a simple featureless pole (e.g. an exciton), and the light-matter physics is just a mixing of two discrete levels [H. 09]. Things are more interesting if the matter resonance is of complex many-body origin. In experiments, "Fermi-edge polaritons" [GPC+07, SWH+14] as well as "polaron-polaritons" [SBC+17] have been observed.

Let us turn to the heavy-impurity singularities of the previous section. For an arbitrary light-matter coupling M_0, our previous analysis is no longer valid, since one must allow for photon absorption/emission processes within the bubbles characterizing the optical response Π, Figs. 5.2, 5.4. However, when the matrix element M_0 is much smaller than the Fermi energy, electronic processes happen way faster than photon absorption/emission. In effect, we can view Π as a photon self-energy valid to second order in M_0, and thus resolve the polaritons as dressed photons. This is done in the publication reprinted below, with an eye for the experimental results of Ref. [SWH+14], where both a low-mobility ($\hat{=}$ infinite mass) and high-mobility ($\hat{=}$ finite mass, $m/M \simeq 0.14$: see introductory section of [Pim15]) samples were used. To summarize the main results:

○ First, we compute Π in the limits $E_{\text{b}} \gg \mu$ and $E_{\text{b}} \ll \mu$, for the first time analyzing the finite mass modification of edge singularities involving a bound state as discussed in Sec. 5.1.2.

○ We then find that the polariton spectra essentially inherit the main attributes of the matter excitations: for $E_b \gg \mu$, the response function Π is sharp for infinite mass (since it is a two-particle correlator that displays the two-body excitonic bound state), and becomes even sharper for finite mass (recall Fig. 5.3(b),(d)). These trends are also reflected in the polaritons. By contrast, for $\mu \gg E_b$, the resonance of Π that is coupled to light is the weak Fermi-edge singularity, which is cut off and broadened for finite mass ((recall Fig. 5.3(a),(c)). The polaritons are broadened accordingly, and the light-matter avoided crossing becomes less pronounced. This faithfully reflects the experimental trends for the "Fermi-edge" polaritons seen in [SWH+14].

Fermi-edge exciton-polaritons in doped semiconductor microcavities with finite hole mass

by

D. Pimenov,[1] J. von Delft,[1] L. Glazman,[2] and M. Goldstein[3]

[1] Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for NanoScience, and Munich Center for Quantum Science and Technology (MCQST), Ludwig-Maximilians-Universität München, 80333 München, Germany

[2]Department of Physics, Yale University, New Haven, Connecticut 05520, USA

[3]Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel

Fermi-edge exciton-polaritons in doped semiconductor microcavities with finite hole mass

Dimitri Pimenov[*] and Jan von Delft

Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-University Munich, 80333 Munich, Germany

Leonid Glazman

Departments of Physics, Yale University, New Haven, Connecticut 06520, USA

Moshe Goldstein

Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel

The coupling between a 2D semiconductor quantum well and an optical cavity gives rise to combined light-matter excitations, the exciton-polaritons. These were usually measured when the conduction band is empty, making the single polariton physics a simple single-body problem. The situation is dramatically different in the presence of a finite conduction-band population, where the creation or annihilation of a single exciton involves a many-body shakeup of the Fermi sea. Recent experiments in this regime revealed a strong modification of the exciton-polariton spectrum. Previous theoretical studies concerned with nonzero Fermi energy mostly relied on the approximation of an immobile valence-band hole with infinite mass, which is appropriate for low-mobility samples only; for high-mobility samples, one needs to consider a mobile hole with large but finite mass. To bridge this gap, we present an analytical diagrammatic approach and tackle a model with short-ranged (screened) electron-hole interaction, studying it in two complementary regimes. We find that the finite hole mass has opposite effects on the exciton-polariton spectra in the two regimes: in the first, where the Fermi energy is much smaller than the exciton binding energy, excitonic features are enhanced by the finite mass. In the second regime, where the Fermi energy is much larger than the exciton binding energy, finite mass effects cut off the excitonic features in the polariton spectra, in qualitative agreement with recent experiments.

I. INTRODUCTION

When a high-quality direct semiconductor 2D quantum well (QW) is placed inside an optical microcavity, the strong coupling of photons and QW excitations gives rise to a new quasiparticle: the polariton. The properties of this fascinating half-light, half-matter particle strongly depend on the nature of the involved matter excitations.

If the Fermi energy is in the semiconductor band gap, the matter excitations are excitons. This case is theoretically well understood [1,2], and the first observation of the resulting microcavity exciton-polaritons was already accomplished in 1992 by Weisbuch *et al.* [3]. Several studies on exciton-polaritons revealed remarkable results. For example, exciton-polaritons can form a Bose-Einstein condensate [4] and were proposed as a platform for high-T_c superconductivity [5].

The problem gets more involved if the Fermi energy is above the conduction-band bottom, i.e., a conduction-band Fermi sea is present. Then the matter excitations have a complex many-body structure, arising from the complementary phenomena of Anderson orthogonality [6] and the Mahan exciton effect, entailing the Fermi-edge singularity [7–11]. An experimental study of the resulting "Fermi-edge polaritons" in a GaAs QW was first conducted in 2007 by Gabbay *et al.* [12], and subsequently extended by Smolka *et al.* [13] (2014). A similar experiment on transition metal dichalcogenide monolayers was recently published by Sidler *et al.* [14] (2016).

From the theory side, Fermi-edge polaritons have been investigated in Refs. [15,16]. However, in these works, only the case of infinite valence-band hole mass was considered, which is the standard assumption in the Fermi-edge singularity or x-ray edge problem. Such a model is valid for low-mobility samples only and thus fails to explain the experimental findings in [13]: there, a high-mobility sample was studied, for which an almost complete vanishing of the polariton splitting was reported. Some consequences of a finite hole mass for polaritons were considered in a recent treatment [17], but without fully accounting for the so-called crossed diagrams that describe the Fermi sea shakeup, as we further elaborate below.

The aim of the present paper is therefore to study the effects of both finite mass and Fermi-edge singularity on polariton spectra in a systematic fashion. This is done analytically for a simplified model involving a contact interaction, which nevertheless preserves the qualitative features of spectra stemming from the finite hole mass and the presence of a Fermi sea. In doing so, we distinguish two regimes, with the Fermi energy μ being either much smaller or much larger than the exciton binding energy E_B. For the regime where the Fermi energy is much larger than the exciton binding energy, $\mu \gg E_B$, several treatments of finite-mass effects on the Fermi-edge singularity alone (i.e., without polaritons) are available, both analytical and numerical. Without claiming completeness, we list Refs. [18–22]. In our work, we have mainly followed the approach of Ref. [18], extending it by going from 3D to 2D and, more importantly, by addressing the cavity coupling which gives rise to polaritons. For infinite hole mass, the sharp electronic spectral feature caused by the

---

[*]D.Pimenov@physik.lmu.de

Fermi edge singularity can couple with the cavity mode to create sharp polariton-type spectral peaks [15,16]. We find that the finite hole mass cuts off the Fermi edge singularity and suppresses these polariton features.

In the opposite regime of $\mu \ll E_B$, where the Fermi energy is much smaller than the exciton binding energy, we are not aware of any previous work addressing the modification of the Fermi-edge singularity due to finite mass. Here, we propose a way to close this gap using a diagrammatic approach. Interestingly, we find that in this regime the excitonic singularities are not cut off, but are rather enhanced by finite hole mass, in analogy to the heavy valence-band hole propagator treated in [23].

This paper has the following structure. First, before embarking into technical details, we will give an intuitive overview of the main results in Sec. II. Detailed computations will be performed in subsequent sections. In Sec. III, the full model describing the coupled cavity-QW system is presented. The key quantity that determines its optical properties is the cavity-photon self-energy Π, which we will approximate by the electron-hole correlator in the absence of a cavity. Section IV shortly recapitulates how Π can be obtained in the regime of vanishing Fermi energy, for infinite and finite hole masses. Then we turn to the many-body problem in the presence of a Fermi sea in the regimes of small (Sec. V) and large Fermi energy (Sec. VI). Using the results of the previous sections, polariton properties are addressed in Sec. VII. Finally, we summarize our findings and list several possible venues for future study in Sec. VIII.

II. SUMMARY OF RESULTS

In a simplified picture, polaritons arise from the hybridization of two quantum excitations with energies close to each other, the cavity photon and a QW resonance [1,2]. The resulting energy spectrum consists of two polariton branches with an avoided crossing, whose light and matter content are determined by the energy detuning of the cavity mode from the QW mode.

While the cavity photon can be approximated reasonably by a bare mode with quadratic dispersion and a Lorentzian broadening due to cavity losses, the QW resonance has a complicated structure of many-body origin. The QW optical response function is rather sensitive to nonzero density of conduction-band (CB) electrons. Roughly, it tends to broaden QW spectral features, which contribute to the spectral width of polariton lines.

A more detailed description of the polariton lines requires finding first the optical response function $\Pi(\mathbf{Q},\Omega)$ of the QW alone (without polaritons). Here, $\mathbf{Q}$ and Ω are, respectively, the momentum and the energy of an incident photon probing the optical response. The imaginary part of $\Pi(\mathbf{Q},\Omega)$, $A(\mathbf{Q},\Omega) = -\text{Im}[\Pi(\mathbf{Q},\Omega)]/\pi$, defines the spectral function of particle-hole excitations in the QW. In the following, we discuss the evolution of $A(\mathbf{Q},\Omega)$ as the chemical potential μ is varied, concentrating on the realistic case of a finite ratio of the electron and hole masses. We assume that the temperature is low, and consider the zero-temperature limit in the entire work. In addition, we will limit ourselves to the case where the
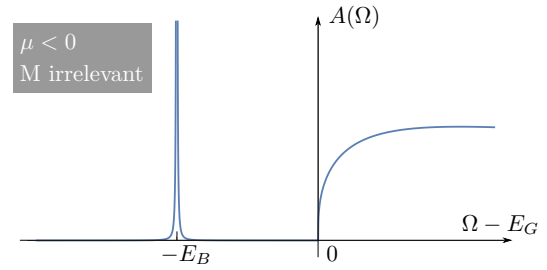

FIG. 1. Absorption spectrum for short-range electron-hole interaction and $\mu < 0$, given by the imaginary part of Eq. (19).

photon is incident perpendicular to the QW, i.e., its in-plane momentum is zero, and study $A(\Omega) \equiv A(Q=0,\Omega)$.

In the absence of free carriers (μ is in the gap), a CB electron and a hole in the valence-band (VB) create a hydrogenlike spectrum of bound states. In the case of a QW it is given by the 2D Elliot formula (see, e.g., Ref. [24]). Being interested in the spectral function close to the main exciton resonance, we replace the true Coulomb interaction by a model of short-ranged interaction potential of strength g [see Eqs. (10) and (12)]. As a result, there is a single bound state at an energy $E_G - E_B(g)$, which we identify with the lowest-energy exciton state. Here, E_G is the VB-CB gap, and energies are measured with respect to the minimum of the conduction band. A sketch of $A(\Omega)$ is shown in Fig. 1.

For $\mu > 0$, electrons start to populate the CB. If the chemical potential lies within the interval $0 < \mu \ll E_B$, then the excitonic Bohr radius r_B remains small compared to the Fermi wavelength λ_F of the electron gas, and the exciton is well defined. Its interaction with the particle-hole excitations in the CB modifies the spectral function $A(\Omega)$ in the vicinity of the exciton resonance. The limit of an infinite hole mass was considered by Nozières *et al.* [8–10]; due to particle-hole excitations of the CB Fermi sea, which can happen at infinitesimal energy cost, the exciton resonance is replaced by a power law spectrum, see inset of Fig. 2. In terms of the
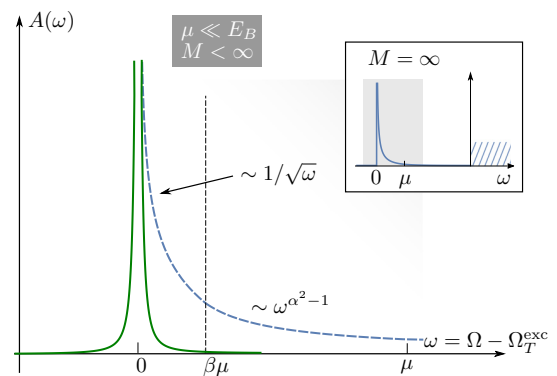

FIG. 2. Absorption for $\mu \ll E_B$ and finite hole mass, illustrating Eq. (5). The full green curve shows the delta peak (broadened for clarity), while the dashed blue line is the incoherent part. Frequencies are measured from the exciton threshold frequency $\Omega_T^{exc} = E_G + \mu - E_B$. The inset shows the infinite mass spectrum for comparison. The dashed region in the inset indicates the continuous part of the spectrum, whose detailed form is beyond the scope of this paper, as we only consider the leading singular parts of all spectra.

detuning from the exciton threshold,

$$\omega = \Omega - \Omega_T^{\mathrm{exc}}, \quad \Omega_T^{\mathrm{exc}} = E_G + \mu - E_B, \tag{1}$$

the spectral function, $A_{\mathrm{exc}}(\omega) = -\mathrm{Im}[\Pi_{\mathrm{exc}}(\omega)]/\pi$, scales as

$$A_{\mathrm{exc}}(\omega)|_{M=\infty} \sim \theta(\omega)\frac{E_B}{\omega}\left(\frac{\omega}{\mu}\right)^{\alpha^2}, \quad \omega \ll \mu. \tag{2}$$

The effective exciton-electron interaction parameter α was found by Combescot et al. [11], making use of final-state Slater determinants. In their work, α is obtained in terms of the scattering phase shift δ of Fermi level electrons off the hole potential, in the presence of a bound state, as $\alpha = |\delta/\pi - 1|$. For the system discussed here, this gives [25]

$$\alpha = 1\left/\left|\ln\left(\frac{\mu}{E_B}\right)\right|\right. \tag{3}$$

We re-derive the result for α diagrammatically (see Sec. V), in order to extend the result of Combescot et al. to the case of a small but nonzero CB electron-VB hole mass ratio β, where

$$\beta = m/M. \tag{4}$$

While the deviation of β from zero does not affect the effective interaction constant α, it brings qualitatively new features to $A(\Omega)$, illustrated in Fig. 2. The origin of these changes is found in the kinematics of the interaction of the exciton with the CB electrons. Momentum conservation for finite exciton mass results in phase-space constraints for the CB particle-hole pairs which may be excited in the process of exciton creation. As a result, the effective density of states $\nu(\omega)$ of the pairs with pair energy ω (also corresponding to the exciton decay rate) is reduced from $\nu(\omega) \sim \omega$ at $\beta = 0$ [11] to $\nu(\omega) \sim \omega^{3/2}$ when ω is small compared to the recoil energy $E_R = \beta\mu$. A smaller density of states for pairs leads to a reduced transfer of the spectral weight to the tail; therefore the delta-function singularity at the exciton resonance survives the interaction with CB electrons, i.e., $\beta > 0$ tends to restore the exciton pole, and one finds

$$A_{\mathrm{exc}}(\omega)|_{M<\infty} = A_{\mathrm{exc,incoh.}}(\omega)\theta(\omega) + \beta^{\alpha^2}E_B\delta(\omega), \tag{5a}$$

$$A_{\mathrm{exc,incoh.}}(\omega) \sim E_B \begin{cases} \frac{\alpha^2}{\sqrt{\omega\beta\mu}}\beta^{\alpha^2}, & \omega \ll \beta\mu \\ \frac{\alpha^2}{\omega}\left(\frac{\omega}{\mu}\right)^{\alpha^2}, & \beta\mu \ll \omega \ll \mu \end{cases}. \tag{5b}$$

The main features of this spectral function are summarized in Fig. 2: as expected, the exciton recoil only plays a role for small frequencies $\omega \ll \beta\mu$, while the infinite mass edge singularity is recovered for larger frequencies. The spectral weight of the delta peak is suppressed by the interaction. For $\beta \to 0$ and $\alpha \neq 0$, we recover the infinite mass result, where no coherent part shows up. If, on the opposite, $\alpha^2 \to 0$ but $\beta \neq 0$, the weight of the delta peak goes to one; the exciton does not interact with the Fermi sea, and its spectral function becomes a pure delta peak, regardless of the exciton mass. A partial survival of the coherent peak at $\alpha, \beta \neq 0$ could be anticipated from the results of Rosch and Kopp [23] who considered the motion of a heavy particle in a Fermi gas of light particles. This problem was also analyzed by Nozières [22], and the coherent peak can be recovered by Fourier transforming his time domain result for the heavy particle Green's function.

At this point, let us note the following: for $\mu > 0$, the hole can bind two electrons with opposite spin, giving rise to trion features in the spectrum. We will not focus on those, since, for weak doping, their spectral weight is small in μ (more precisely, in μ/E_T, where $E_T \ll E_B$ is the trion binding energy), and they are red detuned with respect to the spectral features highlighted in this work. In the regime of $\mu \gg E_B \gg E_T$, trions should be neglible as well. Some further discussion of trion properties can be found in Appendix C.

Upon increase of chemical potential μ, the CB continuum part (inset of Fig. 2) starts building up into the well-known Fermi-edge singularity (FES) at the Burstein-Moss [26,27] shifted threshold, $\Omega_T^{\mathrm{FES}} = E_G + \mu$. For finite mass ($\beta \neq 0$), the FES will however be broadened by recoil effects (see below). At the same time, the delta-function singularity of Eq. (5a) at the absorption edge vanishes at some value of μ. So, at higher electron densities, it is only the FES which yields a nonmonotonic behavior of the absorption coefficient, while the absorption edge is described by a converging power law with fixed exponent, see Eq. (8). This evolution may be contrasted to the one at $\beta = 0$. According to [11,21], the counterparts of the absorption edge and broadened FES are two power law nonanalytical points of the spectrum, which are present at any μ and characterized by exponents continuously evolving with μ. A more detailed discussion of the evolution of absorption spectra as μ increases from small to intermediate to large values is presented in Appendix A.

Let us now consider the limit $\mu \gg E_B$, where the FES is the most prominent spectral feature, in closer detail. In the case of infinite hole mass ($\beta = 0$), and in terms of the detuning from the FES threshold,

$$\omega = \Omega - \Omega_T^{\mathrm{FES}}, \quad \Omega_T^{\mathrm{FES}} = E_G + \mu, \tag{6}$$

the FES absorption scales as [8–10]

$$A_{\mathrm{FES}}(\omega)|_{M=\infty} \sim \theta(\omega)\left(\frac{\omega}{\mu}\right)^{-2g}, \tag{7}$$

as illustrated in the inset of Fig. 3. In the above formula, the interaction contribution to the treshold shift, which is of order $g\mu$, is implicitly contained in a renormalized gap E_G.
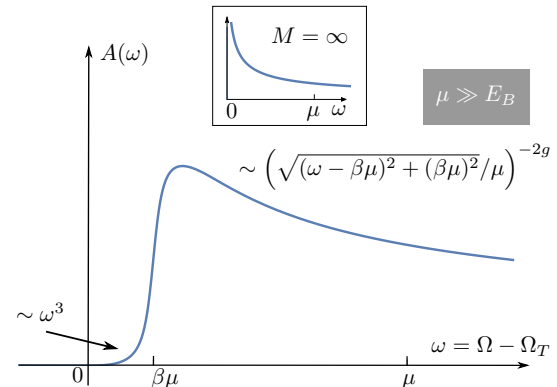

FIG. 3. Finite mass absorption in the case $E_B \ll \mu$. Frequencies are measured from $\Omega_T^{\mathrm{FES}} = E_G + \mu$. The inset shows the infinite mass case for comparison.
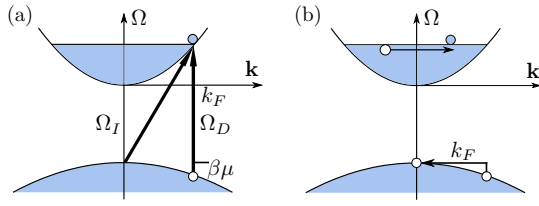
FIG. 4. (a) The direct threshold $\Omega_D = \Omega_T^{\mathrm{FES}} + \beta\mu$ and the indirect threshold $\Omega_I = \Omega_T^{\mathrm{FES}}$ (in the main text, $\omega_{D/I} = \Omega_{D/I} - \Omega_T^{\mathrm{FES}}$). (b) The VB hole can undergo inelastic processes which reduces its energy, smearing the infinite mass edge singularity.

What happens for finite mass? This question was answered in Refs. [18,21,22]; as before, the recoil comes into play, effectively cutting the logarithms contributing to (7). Notably, the relevant quantity is now the *VB hole* recoil, since the exciton is no longer a well defined entity. The FES is then replaced by a rounded feature, sketched in Fig. 3, which sets in continuously:

$$A_{\mathrm{FES}}(\omega)|_{M<\infty} \sim \begin{cases} \left(\frac{\omega}{\beta\mu}\right)^3 \beta^{-2g}\,\theta(\omega), & \omega \ll \beta\mu \\ \left(\frac{\sqrt{(\omega-\beta\mu)^2+(\beta\mu)^2}}{\mu}\right)^{-2g}, & \beta\mu \ll \omega \ll \mu \end{cases} . \tag{8}$$

Equation (8) can be obtained by combining and extending to 2D the results presented in Refs. [18,21].

The maximum of Eq. (8) is found at the so-called direct threshold, $\omega_D = \beta\mu$ [see Fig. 4(a)]. This shift is a simple effect of the Pauli principle: the photoexcited electron needs to be placed on top of the CB Fermi sea. The VB hole created this way, with momentum k_F, can subsequently decay into a zero momentum hole, scattering with conduction-band electrons [see Fig. 4(b)]. These processes render the lifetime of the hole finite, with a decay rate $\sim g^2\beta\mu$. Within the logarithmic accuracy of the Fermi edge calculations, this is equal to $\beta\mu$, the cutoff of the power law in Eq. (8) (see Sec. VI B for a more detailed discussion). As a result, the true threshold of absorption is found at the indirect threshold, $\omega_I = 0$. Due to VB hole recoil, the CB hole-electron pair density of states now scales as $\nu(\omega) \sim \omega^3$, leading to a similar behavior of the spectrum, see Fig. 3.
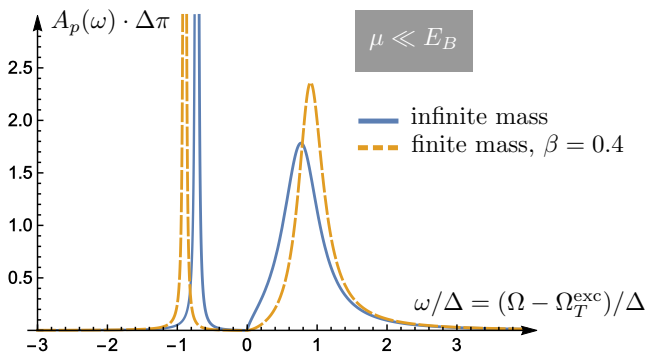
FIG. 5. Comparison of the polariton spectrum for $\mu \ll E_B$, at zero cavity detuning. Frequencies are measured from the exciton threshold, $\Omega_T^{\mathrm{exc}} = E_G + \mu - E_B$. The energy unit Δ corresponds to the half mode splitting at zero detuning in the bare exciton case ($\mu = 0$).
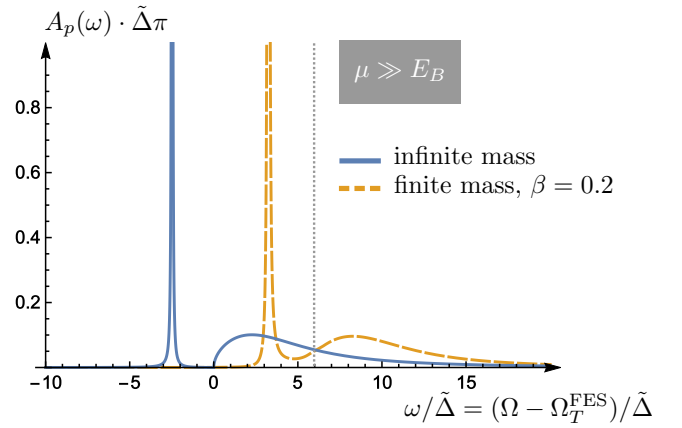

FIG. 6. Comparison of the polariton spectrum for $\mu \gg E_B$, at zero cavity detuning. Frequencies are measured from the indirect threshold, $\Omega_T^{\mathrm{FES}} = E_G + \mu$. The energy unit $\tilde{\Delta}$, which determines the polariton splitting at zero detuning, is defined in Sec. VII, Eq. (76). The dotted vertical line indicates the position of the direct threshold, $\omega_D = \beta\mu$.

We note that at finite ratio $\beta = m/M$, raising the chemical potential μ from $\mu \ll E_B$ to $\mu \gg E_B$ results in a qualitative change of the threshold behavior from a singular one of Eq. (5b), to a converging power law, see the first line of Eq. (8). Simultaneously, a broadened FES feature appears in the continuum, at $\omega > 0$. The difference in the value of the exponent in the excitonic result [Eq. (5b)], as compared to the FES low-energy behavior [Eq. (8) for $\omega \ll \beta\mu$], can be understood from the difference in the kinematic structure of the excitations. In the exciton case, the relevant scattering partners are an exciton and a CB electron-hole pair. In the FES case, one has the photoexcited electron as an additional scattering partner, which leads to further kinematic constraints and eventually results in a different low-energy power law.

In the frequency range $\beta\mu \ll \omega \lesssim \mu$, the physics is basically the same as in the infinite hole mass case ($\beta = 0$). There, the behavior near the lowest threshold (which is exciton energy for $\mu \ll E_B$ and the CB continuum for $\mu \gg E_B$) is always $\sim \omega^{(1-\delta/\pi)^2-1} = \omega^{(\delta/\pi)^2-2\delta/\pi}$. But in the first case ($\mu \ll E_B$), $\delta \sim \pi - \alpha$ is close to π (due to the presence of a bound state), so the threshold singularity is in some sense close to the delta peak, $\sim \mathrm{Im}[1/(\omega + i0^+)]$, that one would have for $\mu = 0$, whereas in the second case ($\mu \gg E_B$), $\delta \sim g$ is close to zero, so the threshold singularity is similar to a discontinuity.

Having discussed spectral properties of the QW alone, we can now return to polaritons. Their spectra $A_p(\omega)$ can be obtained by inserting the QW polarization as photon self-energy. While a full technical account will be given in Sec. VII, the main results can be summarized as follows.

In the first case of study, of $\mu \ll E_B$ and finite β, the polaritons arise from a mixing of the cavity and the sharp exciton mode. The smaller the hole mass, the more singular the exciton features, leading also to sharper polariton features. Furthermore, the enhanced exciton quasiparticle weight pushes the two polariton branches further apart. Conversely, in the singular limit of infinite hole mass, the pole in the exciton spectrum turns into the pure power law familiar from previous
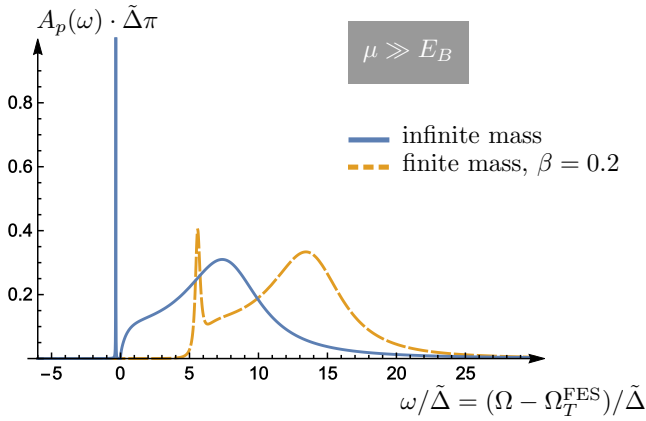
FIG. 7. Comparison of the polariton spectrum for $\mu \gg E_B$, at large positive cavity detuning. Frequencies are measured from the indirect threshold, $\Omega_T^{\mathrm{FES}} = E_G + \mu$.

work, resulting in broader polariton features. A comparison of the infinite and finite hole mass versions of the polariton spectra $A_p(\omega)$ when the cavity photon is tuned into resonance with the exciton is presented in Fig. 5. Notably, the above effects are rather weak, since the exciton is a relatively sharp resonance even for infinite hole mass.

In the second case, $\mu \gg E_B$, the matter component of the polaritons corresponds to the FES singularity, which is much less singular than the exciton. Consequently, the polaritons (especially the upper one, which sees the high-frequency tail of the FES) are strongly washed out already at $\beta = 0$. For finite hole mass, the hole recoil cuts off the FES singularity, resulting in further broadening of the polaritons. In addition, there is an overall upward frequency shift by $\beta\mu$, reflecting the direct threshold effect. Figure 6 shows the two polariton spectra at zero detuning.

The cutoff of the lower polariton for finite masses is even more drastic when the cavity is blue-detuned with respect to the threshold: Indeed, at large positive cavity detuning, the lower polariton is mostly matter-like, and thus more sensitive to the FES broadening. It therefore almost disappears, as seen in Fig. 7.

III. MODEL

After the qualitative overview in the previous section, let us now go into more detail, starting with the precise model in question. To describe the coupled cavity-QW system, we study the following 2D Hamiltonian:

$$H = H_M + H_L, \tag{9}$$

$$H_M = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} - \sum_{\mathbf{k}} (E_{\mathbf{k}} + E_G) b_{\mathbf{k}}^{\dagger} b_{\mathbf{k}}$$
$$- \frac{V_0}{\mathcal{S}} \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} a_{\mathbf{k}}^{\dagger} a_{\mathbf{p}} b_{\mathbf{k}-\mathbf{q}} b_{\mathbf{p}-\mathbf{q}}^{\dagger}, \tag{10}$$

$$H_L = \sum_{\mathbf{Q}} \omega_{\mathbf{Q}} c_{\mathbf{Q}}^{\dagger} c_{\mathbf{Q}} - i \frac{d_0}{\sqrt{\mathcal{S}}} \sum_{\mathbf{p},\mathbf{Q}} a_{\mathbf{p}+\mathbf{Q}}^{\dagger} b_{\mathbf{p}} c_{\mathbf{Q}} + \mathrm{H.c.} \tag{11}$$

Here, H_M, adapted from the standard literature on the x-ray edge problem [18], represents the matter part of the system, given by a semiconductor in a two-band approximation: $a_{\mathbf{k}}$ annihilates a conduction-band (CB) electron with dispersion $\epsilon_{\mathbf{k}} = \frac{k^2}{2m}$, while $b_{\mathbf{k}}$ annihilates a valence-band (VB) electron with dispersion $-E_{\mathbf{k}} + E_G = -\frac{k^2}{2M} + E_G$. E_G is the gap energy, which is the largest energy scale under consideration; in GaAs, $E_G \simeq 2\,\mathrm{eV}$, while all other electronic energies are on the order of meV. The energies are measured from the bottom of the conduction band. $\mathcal{S}$ is the area of the QW, and we work in units where $\hbar = 1$. Unless explicitly stated otherwise, we assume spinless electrons, and concentrate on the zero temperature limit.

When a valence-band hole is created via cavity photon absorption, it interacts with the conduction-band electrons with an attractive Coulomb interaction. Taking into account screening, we model the interaction as pointlike, with a constant positive matrix element V_0. The effective potential strength is then given by the dimensionless quantity

$$g = \rho V_0, \quad \rho = \frac{m}{2\pi}, \tag{12}$$

ρ being the 2D DOS. The appropriate value of g will be further discussed in the subsequent sections.

Interactions of CB electrons with each other are completely disregarded in Eq. (9), presuming a Fermi liquid picture. This is certainly a crude approximation. It can be justified if one is mostly interested in the form of singularities in the spectral function. These are dominated by various power laws, which arise from low-energy particle hole excitations of electrons close to the Fermi energy, where a Fermi-liquid description should be valid.

The photons are described by H_L. We study lossless modes with QW in-plane momenta $\mathbf{Q}$ and energies $\omega_{\mathbf{Q}} = \omega_c + Q^2/2m_c$, where m_c is the cavity mode effective mass. Different in-plane momenta $\mathbf{Q}$ can be achieved by tilting the light source with respect to the QW. In the final evaluations, we will mostly set $\mathbf{Q} = 0$, which is a valid approximation since m_c is tiny compared to electronic masses. The interaction term of H_L describes the process of absorbing a photon while creating an VB-CB electron hole pair, and vice versa. d_0 is the interband electric dipole matrix element, whose weak momentum dependence is disregarded. This interaction term can be straightforwardly derived from a minimal coupling Hamiltonian studying interband processes only, and employing the rotating wave and electric dipole approximations (see, e.g., Ref. [28]).

The optical properties of the full system are determined by the retarded dressed photon Green's function [16,17]:

$$D^R(\mathbf{Q},\Omega) = \frac{1}{\Omega - \omega_{\mathbf{Q}} + i0^+ - \Pi(\mathbf{Q},\Omega)}, \tag{13}$$

where $\Pi(\mathbf{Q},\Omega)$ is the retarded photon self-energy. This dressed photon is nothing but the polariton. The spectral function corresponding to (13) is given by

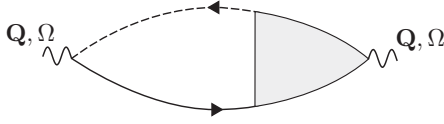$$\mathcal{A}(\mathbf{Q},\omega) = -\frac{1}{\pi}\mathrm{Im}[D^R(\mathbf{Q},\omega)]. \tag{14}$$

FIG. 8. The photon self-energy $\Pi(\mathbf{Q}, \Omega)$ in linear response. Full lines denote CB electrons, dashed lines VB electrons, and wavy lines photons. The grey-shaded area represents the full CB-VB vertex.

$\mathcal{A}(\mathbf{Q}, \omega)$ determines the absorption respectively reflection of the coupled cavity-QW system, which are the quantities typically measured in polariton experiments like [12,13].

Our goal is to determine $\Pi(\mathbf{Q}, \Omega)$. To second order in d_0, it takes the form

$$\Pi(\mathbf{Q}, \Omega) \simeq -i \frac{d_0^2}{\mathcal{S}} \int_{-\infty}^{\infty} dt \, \theta(t) e^{(i\Omega t)}$$
$$\times \sum_{\mathbf{k}, \mathbf{p}} \langle 0 | b_{\mathbf{k}}^{\dagger}(t) a_{\mathbf{k}+\mathbf{Q}}(t) a_{\mathbf{p}+\mathbf{Q}}^{\dagger}(0) b_{\mathbf{p}}(0) | 0 \rangle, \quad (15)$$

where $|0\rangle$ is the noninteracting electronic vacuum with a filled VB, and the time dependence of the operators is generated by H_M. Within this approximation, $\Pi(\mathbf{Q}, \omega)$ is given by the "dressed bubble" shown in Fig. 8. The imaginary part of $\Pi(\mathbf{Q}, \omega)$ can also be seen as the linear response absorption of the QW alone with the cavity modes tuned away.

Starting from Eq. (15), in the following we will study in detail how $\Pi(\mathbf{Q}, \omega)$ behaves as the chemical potential μ is increased, and distinguish finite and infinite VB masses M. We will also discuss the validity of the approximation of calculating Π to lowest order in d_0.

IV. ELECTRON-HOLE CORRELATOR IN THE ABSENCE OF A FERMI SEA

We start by shortly reviewing the diagrammatic approach in the case when the chemical potential lies within the gap (i.e., $-E_G < \mu < 0$). This is mainly done in order to set the stage for the more involved diagrammatic computations in the subsequent sections. In this regime of μ, Π is exactly given by the sum of the series of ladder diagrams shown in Fig. 9, first computed by Mahan [29]. Indeed, all other diagrams are absent here since they either contain VB or CB loops, which are forbidden for μ in the gap. This is seen using the following expressions for the zero-temperature time-ordered free Green's functions:

$$G_c^{(0)}(\mathbf{k}, \Omega) = \frac{1}{\Omega - \epsilon_{\mathbf{k}} + i0^+ \text{sgn}(\epsilon_{\mathbf{k}} - \mu)}, \quad (16)$$
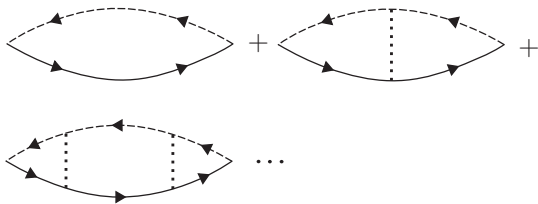

FIG. 9. The series of ladder diagrams. Dotted lines represent the electron-hole interaction.

$$G_v^{(0)}(\mathbf{k}, \Omega) = \frac{1}{\Omega + E_G + E_{\mathbf{k}} + i0^+ \text{sgn}(-E_G - E_{\mathbf{k}} - \mu)}, \quad (17)$$

where the indices c and v stand for conduction and valence band, respectively, and 0^+ is an infinitesimal positive constant. For $-E_G < \mu < 0$, CB electrons are purely retarded, while VB electrons are purely advanced. Thus no loops are possible. Higher-order terms in d_0 are not allowed as well.

One can easily sum up the series of ladder diagrams assuming the simplified interaction V_0 [18]. Let us start from the case of infinite VB mass ($\beta = 0$), and concentrate on energies $|\Omega - E_G| \ll \xi$, where ξ is an appropriate UV cutoff of order of CB bandwidth. Since the interaction is momentum independent, all integrations in higher-order diagrams factorize. Therefore the nth order diagram of Fig. 9 is readily computed:

$$\Pi_{\text{ladder}}^{(n)}(\Omega) = d_0^2 \rho(-g)^n \ln\left(\frac{\Omega - E_G + i0^+}{-\xi}\right)^{n+1}. \quad (18)$$

Here and henceforth, the branch cut of the complex logarithm and power laws is chosen to be on the negative real axis. The geometric series of ladder diagrams can be easily summed:

$$\Pi_{\text{ladder}}(\Omega) = \sum_{n=0}^{\infty} \Pi_{\text{Ladder}}^{(n)}(\Omega) = \frac{d_0^2 \rho \ln\left(\frac{\Omega - E_G + i0^+}{-\xi}\right)}{1 + g \ln\left(\frac{\Omega - E_G + i0^+}{-\xi}\right)}. \quad (19)$$

A sketch of the corresponding QW absorption $A_{\text{ladder}} = -\text{Im}[\Pi_{\text{ladder}}]/\pi$ was already shown in Fig. 1.

$\Pi_{\text{ladder}}(\Omega)$ has a pole, the so-called Mahan exciton [18,29], at an energy of

$$\Omega - E_G = -E_B = -\xi e^{-1/g}. \quad (20)$$

In the following, we will treat E_B as a phenomenological parameter. To match the results of the short-range interaction model with an experiment, one should equate E_B with E_0, the energy of lowest VB hole-CB electron hydrogenic bound state (exciton). Expanding Eq. (19) near the pole, we obtain

$$\Pi_{\text{ladder}}(\omega) = \frac{d_0^2 E_B \rho}{g^2} G_{\text{exc}}^0(\omega) + \mathcal{O}\left(\frac{\omega}{E_B}\right),$$
$$G_{\text{exc}}^0(\omega) = \frac{1}{\omega + i0^+}, \quad (21)$$

where $\omega = \Omega - E_G + E_B$, and we have introduced the bare exciton Green's function G_{exc}^0, similar to Ref. [30].

In this regime of μ, a finite hole mass only results in a weak renormalization of the energy by factors of $1 + \beta$, where $\beta = m/M$ is the small CB/VB mass ratio. Furthermore, if finite photon momenta $\mathbf{Q}$ are considered, the exciton Green's function is easily shown to be (near the pole):

$$G_{\text{exc}}^0(\mathbf{Q}, \omega) = \frac{1}{\omega + Q^2/M_{\text{exc}} + i0^+}, \quad (22)$$

with $M_{\text{exc}} = M + m = M(1 + \beta)$.

V. ELECTRON-HOLE CORRELATOR FOR SMALL FERMI ENERGY

A. Infinite VB hole mass

Let us now slightly increase the chemical potential μ, and study the resulting absorption. More precisely, we consider the regime

$$0 < \mu \ll E_B \ll \xi. \tag{23}$$

We first give an estimate of the coupling constant $g = \rho V_0$ Accounting for screening of the VB hole 2D Coulomb potential by the CB Fermi sea in the static RPA approximation, and averaging over the Fermi surface [18,29], one finds

$$g \sim \begin{cases} 1 - 8x/\pi & x \to 0, \\ \ln(x)/x & x \to \infty, \end{cases} \tag{24}$$

where $x = \sqrt{\mu/E_0}$ with E_0 being the true 2D binding energy of the lowest exciton in the absence of a CB Fermi sea. In the regime under study, we may assume $E_B \simeq E_0 \gg \mu$, and therefore $g \lesssim 1$.[1] As a result, perturbation theory in g is meaningless. Instead, we will use μ/E_B as our small parameter, and re-sum all diagrams which contribute to the lowest nontrivial order in it.

We will now restrict ourselves to the study of energies close to E_B in order to understand how a small density of CB electrons modifies the shape of the bound state resonance; we will not study in detail the VB continuum in the spectrum (cf. Fig. 2). We first compute the contribution of the ladder diagrams; as compared to Eqs. (21) and (22), the result solely differs by a shift of energies:

$$\omega = \Omega - \Omega_T^{\rm exc}, \quad \Omega_T^{\rm exc} = (E_G + \mu) - E_B. \tag{25}$$

Also, the continuum now sets in when Ω equals $\Omega_T^{\rm FES} = E_G + \mu$, which is known as the Burstein-Moss shift [26,27]. However, for finite μ, one clearly needs to go beyond the ladder approximation, and take into account the "Fermi sea shakeup." To do so, we first consider the limit of infinite M ($\beta = 0$). In this regime, the QW absorption in the presence of a bound state for the model under consideration was found by Combescot and Nozières [11], using a different approach.[2]

For finite μ, the physics of the Fermi-edge singularity comes into play: due to the presence of the CB Fermi sea, CB electron-hole excitations are possible at infinitesimal energy cost. As a result, the exciton Green's function, which we analogously to (21) define as proportional to the dressed bubble in the exciton regime,

$$\Pi_{\rm exc}(\omega) = \frac{d_0^2 E_B \rho}{g^2} G_{\rm exc}(\omega) + \mathcal{O}\left(\frac{\omega}{E_B}\right), \tag{26}$$

$$G_{\rm exc}(\omega) = \frac{1}{\omega - \Sigma^{\rm exc}(\omega)}, \tag{27}$$

---

[1]Strictly speaking, this also means $E_B \lesssim \xi$, contradicting Eq. (23). However, this clearly is a nonuniversal property, and we will not pay any attention to it in the following.

[2]In fact, their computation is in 3D, but the case of infinite hole mass is effectively 1D anyway.
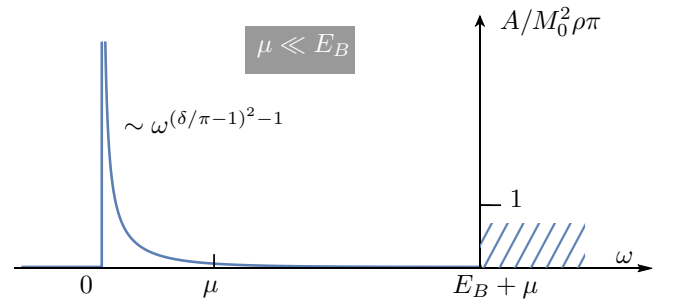

FIG. 10. QW Absorption for $\mu \ll E_B$ and $M = \infty$. The power law (28) is valid asymptotically close to the left peak. The dashed region indicates the continuous part of the spectrum, compare caption of Fig. 2.

gets renormalized by a self-energy $\Sigma^{\rm exc}(\omega)$. This self-energy turns the exciton pole turns into a divergent power law [11]:

$$G_{\rm exc}(\omega) \sim \frac{1}{\omega + i0^+} \cdot \left(\frac{\omega + i0^+}{-\mu}\right)^{(\delta/\pi - 1)^2}, \tag{28}$$

where δ is the scattering phase shift of electrons at the Fermi level off the pointlike hole potential. One should note that no delta peak will appear for $\delta/\pi \neq 1$. A sketch of the resulting absorption A is shown in Fig. 10.

Let us further discuss the result (28). It was obtained in Ref. [11] using an elaborate analytical evaluation of final state Slater determinants, and actually holds for any value of μ. A numerical version of this approach for the infinite VB mass case was recently applied by Baeten and Wouters [16] in their treatment of polaritons. In addition, the method was numerically adapted to finite masses by Hawrylak [19], who, however, mostly considered the mass effects for $\mu \gg E_B$.

However, due to the more complicated momentum structure, it seems difficult to carry over the method of [11] to finite masses analytically. Instead, we will now show how to proceed diagrammatically. Our analysis will give (28) to leading order in the small parameter μ/E_B, or, equivalently, $\alpha = \delta/\pi - 1$ (recall that by Levinson's theorem [25] $\delta = \pi$ for $\mu = 0$ due to the presence of a bound state—the exciton):

$$G_{\rm exc}(\omega) \simeq \frac{1}{\omega + i0^+}\left(1 + \alpha^2 \ln\left(\frac{|\omega|}{\mu}\right) - i\alpha^2 \pi \theta(\omega)\right). \tag{29}$$

The merit of the diagrammatical computation is twofold. First, it gives an explicit relation between α and the experimentally measurable parameters μ, E_B. Second, the approach can be straightforwardly generalized to finite masses, as we show in the next section.

Let us note that a similar diagrammatic method was also examined by Combescot, Betbeder-Matibet et al. in a series of recent papers [30–34]. Their model Hamiltonians are built from realistic Coulomb electron-hole and electron-electron interactions. As a result, they assess the standard methods of electron-hole diagrams as too complicated [30], and subsequently resort to exciton diagrams and the so-called commutation technique, where the composite nature of the excitons is treated with care. However, the interaction of excitons with a Fermi sea is only treated at a perturbative level, assuming that the interaction is small due to, e.g., spatial
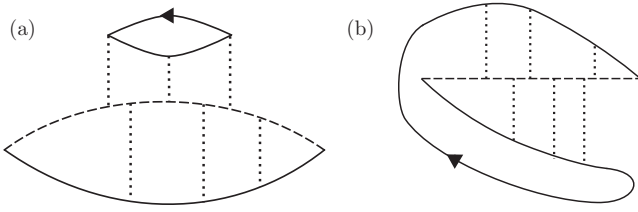
FIG. 11. Leading-order direct self-energy diagrams: (a) direct contribution D and (b) exchange contribution X.

separation [31]. This is not admissible in our model, where the interaction of the VB hole with all relevant electrons (photoexcited and Fermi sea) has to be treated on the same footing. Rather, we stick to the simplified form of contact interaction, and show how one can use the framework of standard electron-hole diagrams to calculate all quantities of interest for infinite as well as for finite VB mass. The results presented below then suggest that for $\mu \ll E_B$ the finite mass does not weaken, but rather strengthens the singularities, which is in line with results on the heavy hole found in Ref. [23].

Here, we only present the most important physical ingredients for our approach, and defer the more technical details to Appendix B. In the regime of interest, we can perform a low-density computation, employing the small parameter μ/E_B. Since all energies are close to E_B, the leading-order exciton self-energy diagrams is then the sum of all diagrams with one CB electron loop. One can distinguish two channels: direct and exchange, to be denoted by D and X, as depicted in Fig. 11. All such diagrams with an arbitrary number of interactions connecting the VB line with the CB lines in arbitrary order have to be summed. Factoring out $E_B \rho/g^2 \, G_{\text{exc}}^0(\omega)^2$, the remaining factor can be identified as the exciton self-energy diagram.

An evaluation of these diagrams is possible either in the time or in the frequency domain. Of course, both approaches must give same result. In practice, however, the time domain evaluation is more instructive and requires less approximations, which is why we will discuss it first. The frequency domain evaluation, however, is far more convenient for obtaining finite mass results, and will be discussed thereafter.

The time domain approach is similar in spirit to the classical one-body solution of the Fermi-edge problem by Nozières and de Dominicis [10]. Since the infinite-mass hole propagator is trivial, $G_v(t) = i\theta(-t)e^{iE_G t}$, the direct diagrams just describe the independent propagation of two electrons in the time-dependent hole potential. Thus, in the time domain, the sum of all direct diagrams $D(t)$ factorizes into two parts representing the propagation of these two electrons:

$$D(t) = \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} i e^{-i(E_G - \epsilon_{\mathbf{k}_1})t} B(t) C(t), \quad (30)$$

where $B(t)$ and $C(t)$ are infinite sums of convolutions (denoted by an asterisk) of the form

$$B(t) = \sum_{m=1}^{\infty} (-V_0)^m \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \cdots \int_{k_m > k_F} \frac{d\mathbf{k}_m}{(2\pi)^2}$$
$$\times \left[G_c^{0,R}(\mathbf{k}_1,) * \cdots * G_c^{0,R}(\mathbf{k}_m,) * G_c^{0,R}(\mathbf{k}_1,) \right](t), \quad (31)$$

and similarly for $C(t)$. $G_c^{0,R}$ is the retarded bare CB Green's function in the time domain. Fourier-transforming, $D(\omega)$ is then given by a convolution of $B(\omega)$ and $C(\omega)$, each of which in turn reduces to simple summations of ladder diagrams. The full convolution $D(\omega)$ is difficult to compute; one can proceed by noting that $B(\omega)$ and $C(\omega)$ have poles at $\omega \simeq 0$ and continuum contributions at $\omega \gtrsim E_B$. These are readily identified with the pole and continuum contributions of the exciton absorption, cf. Fig. 1. Combining these, there are four combinations contributing to $D(\omega)$: pole-pole, pole-continuum (two possibilities), and continuum-continuum. The imaginary part of the latter, which is of potential importance for the line shape of the exciton spectrum, can be shown to vanish in our main regime of interest, $\omega \gtrsim 0$. It is instructive to study the pole-pole combination, which corresponds to a would be "trion" (bound state of the exciton and an additional electron) and is further discussed in Appendix C. Adding to it the pole-continuum contributions we find, for small ω:

$$D(\omega) = \frac{\rho E_B}{g^2} \frac{1}{(\omega + i0^+)^2} \Sigma_{\text{exc}}^{\text{D}}(\omega). \quad (32)$$

This corresponds to a contribution to the exciton self-energy, which reads

$$\Sigma_{\text{exc}}^{\text{D}}(\omega) = -\frac{1}{\rho} \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \frac{1}{\ln\left(\frac{\omega + \epsilon_{\mathbf{k}_1} - \mu + i0^+}{-E_B}\right)}. \quad (33)$$

Before discussing this term further, we consider the contribution of the exchange diagrams, $X(\omega)$, of Fig. 11(b). Their structure is more involved compared to the direct channel, since these diagrams do not just represent the independent propagation of two electrons in the hole potential. However, relying on a generalized convolution theorem which we prove, the computation can be performed in the same vein as before (see Appendix B), leading to the following results: first, the pole-pole contribution cancels that of the direct diagrams (see Appendix C), which holds in the spinless case only (in the spinful case, the direct diagrams will come with an extra factor of two). This could be expected: trion physics is only recovered in the spinful case, where two electrons can occupy the single bound state created by the attractive potential of the hole. In a realistic 2D setup trion features will become important for large enough values of μ (see, e.g., Refs. [14,35–37]). Although we do not focus on trions here, let us stress that all standard results on trions can be recovered within our diagrammatic approach, if electrons and holes are treated as spin-1/2 particles; see Appendix C for further details.

The dominant contribution to $X(\omega)$ then arises from the pole-continuum contribution. It is given by

$$X(\omega) = -\frac{\rho E_B}{g^2} \frac{1}{(\omega + i0^+)^2} \mu. \quad (34)$$

Thus, the self-energy contribution to the exciton Green's function is simply

$$\Sigma_{\text{exc}}^{\text{X}}(\omega) = -\mu. \quad (35)$$

Since it is purely real, it will essentially just red-shift the exciton pole by μ. A discussion of this result is presented in Appendix D.

Now, it should be noted that $\Sigma_{\text{exc}}^{\text{X}}(\omega)$ is not proportional to the small parameter μ/E_B—the latter effectively canceled when factoring out the bare excitons Green's function. Thus it is inconsistent to treat $\Sigma_{\text{exc}}^{\text{X}}(\omega)$ as perturbative self-energy correction. Instead, one should repeat the calculation, but replace all ladders by ladders dressed with exchange-type diagrams. It can be expected, however, that the structure of the calculations will not change. The only change that should happen is the appearance of the renormalized binding energy $\tilde{E}_B = E_B + \mu$, in accordance with [11], as discussed in Appendix D. In the following, we will assume this is accounted for, and therefore suppress all exchange diagrams.

Let us now return to the direct self-energy contribution $\Sigma_{\text{exc}}^{\text{D}}(\omega)$, Eq. (33), writing

$$\Sigma_{\text{exc}}(\omega) = \Sigma_{\text{exc}}^{D}(\omega) \tag{36}$$

henceforth. We may apply the following asymptotic expansion for the logarithmic integral (generalized from [38]), which will also prove useful later:

$$\int_0^\omega dx \frac{x^n}{\ln^m(x)} = \frac{1}{\ln^m(\omega)} \frac{\omega^{n+1}}{(n+1)} + \mathcal{O}\left(\frac{\omega^{n+1}}{\ln(\omega)^{m+1}}\right). \tag{37}$$

This can be shown easily by integrating by parts and comparing orders. Based on this result we find, to leading logarithmic accuracy,

$$\Sigma_{\text{exc}}(\omega) \simeq -\frac{\mu}{\ln\left(\frac{\mu}{E_B}\right)} + \frac{\omega \ln\left(\frac{|\omega|}{\mu}\right)}{\ln\left(\frac{\mu}{E_B}\right)\ln\left(\frac{|\omega|}{E_B}\right)} - i\frac{\pi\omega\theta(\omega)}{\ln^2\left(\frac{|\omega|}{E_B}\right)}. \tag{38}$$

This result has several interesting features. First, we see the appearance of a small parameter $\alpha \equiv 1/|\ln(\mu/E_B)|$, which can be interpreted as follows: the scattering phase-shift at the Fermi level, δ, which determines the Anderson orthogonality power law [c.f. Eq. (28)] is approximately given by [25]

$$\delta \simeq \frac{\pi}{\ln\left(\frac{\mu}{E_B}\right)} + \pi, \tag{39}$$

which holds for small Fermi energies, where δ is close to π. Therefore δ and α are related by

$$\alpha \simeq 1 - \frac{\delta}{\pi}. \tag{40}$$

The small pole shift of order $\alpha\mu$ contained in Eq. (38) could be expected from Fumi's theorem (see, e.g., Ref. [39] and the discussion in Appendix D). We now perform an energy shift

$$\omega \rightarrow \omega + \alpha\mu. \tag{41}$$

To leading order in α, we may then rewrite $\Sigma_{\text{exc}}^{\text{D}}$ with logarithmic accuracy as

$$\Sigma_{\text{exc}}(\omega) \simeq \alpha^2 \omega \ln\left(\frac{|\omega|}{\mu}\right) - i\alpha^2 \pi\omega\theta(\omega). \tag{42}$$

Here, the imaginary part can be identified with the density of states of CB electron-hole excitations as function of ω, as discussed in Sec. II.

Upon inserting (42) into the exciton Green's function (27), we recover (28) to leading (quadratic) order in α:

$$G_{\text{exc}}(\omega) \simeq \frac{1}{\omega + i0^+}\left(1 + \alpha^2 \ln\left(\frac{|\omega|}{\mu}\right) - i\alpha^2\pi\theta(\omega)\right). \tag{43}$$
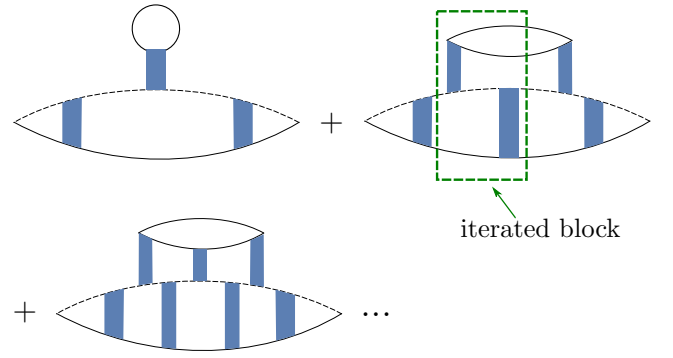

FIG. 12. Series of diagrams contributing to the direct self-energy in the frequency domain. Vertical blue bars denote interaction ladders.

As a result, our one-loop computation has given the first logarithm of the orthogonality power law, in complete analogy to the standard Fermi-edge problem (see Sec. VI). All higher loop contributions, evaluated to leading logarithmic order, should then add up to give the full power law; since we are more interested in finite mass effects here, we will not go into the details of this calculation.

To carry the diagrammatics over to finite mass, as done in the next section, it is convenient to switch to the frequency domain. A summation of all one-loop diagrams is possible by evaluating the series shown in Fig. 12.

To perform the evaluation, we make use of the following simplification. To begin with, we often encounter complicated logarithmic integrals; however, the imaginary part of the integrand is just a delta function, so, upon integration, one finds step functions. Since the integrand is retarded, it is then possible to recover the full expression from the imaginary part using the Kramers-Kronig relation; the step functions then become logarithms.

With that, the sum over diagrams appearing in Fig. 12 assumes the form

$$D(\omega) = \frac{E_B}{g^2} \frac{1}{(\omega + i0^+)^2} \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \{I + I^3 + \ldots\}, \tag{44}$$

where

$$I = \ln\left(\frac{\epsilon_{\mathbf{k}_1} + \omega - \mu + i0^+}{-E_B}\right). \tag{45}$$

Summing up the geometric series exactly reproduces the time-domain result, Eq. (32). Thus we have established how the photon self-energy can be calculated diagrammatically for the case of infinite VB mass M (to leading order in d_0).

B. Finite hole mass

We are now in a position to tackle finite VB mass M. Let us also consider a finite incoming momentum $\mathbf{Q}$. Clearly, the one-loop criterion for choosing diagrams still holds, since we are still considering the low-density limit, $\mu \ll E_B$.

We also disregard any exchange contributions for the same reasons as for the infinite mass case. As a result, we only have to recompute the series of direct diagrams of Fig 12. We start with the first one. It gives

$$I = -\frac{E_B V_0}{g} \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \frac{1}{\left(-\omega + E_B + E(\mathbf{k}_2 - \mathbf{Q}) + \epsilon_{\mathbf{k}_2} - \mu - i0^+\right)^2} \frac{1}{\ln\left(\frac{-E_B + \omega - (\mathbf{Q} - \mathbf{q})^2/2M_{\text{exc}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1} + i0^+}{-E_B}\right)}, \quad (46)$$

where $\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1$. The imaginary part of (46) reads

$$\text{Im}[I] = -\frac{V_0}{g} \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \pi \delta\left(\omega - \frac{(\mathbf{Q} - \mathbf{q})^2}{2M_{\text{exc}}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1}\right)$$
$$+ \mathcal{O}\left(\frac{\mu}{E_B}\right). \quad (47)$$

By Eq. (47), I can be rewritten in a simpler form (ensuring retardation), valid for small ω:

$$I \simeq \frac{V_0}{g} \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \frac{1}{\omega - \frac{(\mathbf{Q} - \mathbf{q})^2}{2M_{\text{exc}}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1} + i0^+}. \quad (48)$$

This form can be integrated with logarithmic accuracy, which, however, only gives $\text{Re}[I]$. Specializing to $Q \ll k_F$ for simplicity, one obtains

$$\text{Re}[I] \simeq \ln\left(\frac{\max(|\omega + \epsilon_{\mathbf{k}_1} - \mu|, \beta\mu)}{E_B}\right). \quad (49)$$

As for the infinite mass case, the higher-order diagrams of Fig. 12 give higher powers of I. Similarly to Eq. (44), one then obtains for the self-energy part, to leading logarithmic accuracy:

$$\Sigma_{\text{exc}}(\mathbf{Q}, \omega) = -\int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \frac{1}{I}. \quad (50)$$

The imaginary part, which determines the lineshape of G_{exc}, is given by

$$\text{Im}[\Sigma_{\text{exc}}(\mathbf{Q}, \omega)] \simeq -\frac{\pi V_0}{\rho g} \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2}$$
$$\times \frac{\delta\left(\omega - (\mathbf{Q} - \mathbf{q})^2/2M_{\text{exc}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1}\right)}{\ln^2\left(\frac{\max(|\omega + \epsilon_1 - \mu|, \beta\mu)}{E_B}\right)}. \quad (51)$$

We now apply the analog of the logarithmic identity, Eq. (37), for a 2D integral. Thus, in leading order, we may simply pull the logarithm out of the integral of Eq. (51) and rewrite it as

$$\text{Im}[\Sigma_{\text{exc}}](\mathbf{Q}, \omega) \simeq -\frac{\pi V_0}{\rho g} \alpha^2 \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2}$$
$$\times \delta\left(\omega - (\mathbf{Q} - \mathbf{q})^2/2M_{\text{exc}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1}\right). \quad (52)$$

The result (52) is physically transparent. It is just a phase-space integral giving the total rate of scattering of an exciton with momentum $\mathbf{Q}$ by a CB Fermi sea electron. The prefactor is determined by the scattering phase shift δ. At least for sufficiently small momenta $\mathbf{Q}$, the integral in Eq. (52) can be straightforwardly computed. For the most important case

$\mathbf{Q} = 0$, one obtains for small energies (see Appendix E)

$$\text{Im}[\Sigma_{\text{exc}}](\mathbf{Q} = 0, \omega) \sim -\alpha^2 \frac{1}{\sqrt{\beta\mu}} \theta(\omega) \omega^{3/2}, \quad \omega \ll \beta\mu, \quad (53)$$

where we suppressed an irrelevant prefactor of order one. For $\omega \gg \beta\mu$, one recovers the infinite mass case as in (42).

Compared to the infinite mass case, where $\text{Im}[\Sigma_{\text{exc}}] \sim \omega \ln(\omega)$, the self-energy (53) shows a suppression of the low-frequency scattering phase space, as seen from the higher frequency power law. Physically, the phase space suppression is understood as follows: we have found that, after accounting for the exchange diagrams, it is admissible to view the exciton as elementary particle with mass M_{exc}, which interacts with the Fermi sea with an effective interaction strength α [Eq. (40)]. As can be seen from Fig. 13, scatterings of the exciton with CB electrons involving a large momentum transfer necessarily cost a finite amount of energy (the so-called recoil energy). By contrast, in the infinite mass case such scatterings could still happen at infinitesimal energy cost, since the exciton dispersion was flat. Thus the finite-mass phase space is reduced as compared to the infinite mass case. This change eventually leads to the previously asserted reappearance of the exciton delta peak.

This phase space reduction also affects the exciton spectral function, and hence the absorption: we first restrict ourselves to the leading behavior, i.e., we disregard any small renormalizations that arise from including $\text{Re}[\Sigma_{\text{exc}}]$ or from higher-loop corrections. Inserting Eq. (53) into Eq. (27) we then obtain, for small energies ω,

$$A(\mathbf{Q} = 0, \omega) \simeq -\Delta^2 \frac{\text{Im}[\Sigma(\omega)]}{\omega^2} \sim \Delta^2 \alpha^2 \frac{\theta(\omega)}{\sqrt{\beta\mu\,\omega}}, \quad (54)$$

with

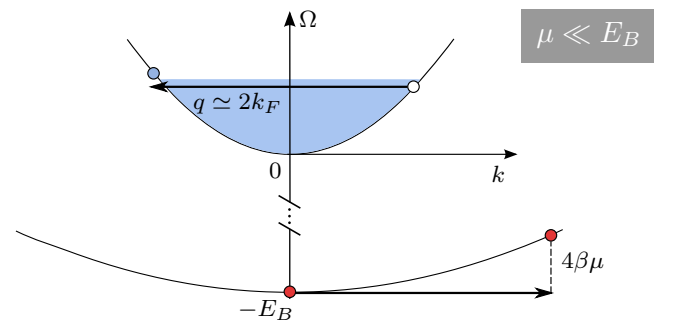$$\Delta^2 = \frac{d_0^2 \rho E_B}{g^2}. \quad (55)$$

FIG. 13. Scattering process of an exciton by a VB electron with large momentum transfer. The lower band represents the exciton dispersion. The scattering significantly increases the exciton energy.

The factor Δ (with units of energy) determines the polariton splitting at zero detuning, and will be discussed in Sec. VII. The $1/\sqrt{\omega}$ divergence seen in (54) was also found by Rosch and Kopp using a path-integral approach [23] for a related problem, that of a heavy hole propagating in a Fermi sea. In addition, Rosch and Kopp find a quasiparticle delta peak with a finite weight. This peak can also be recovered within our approach upon inclusion of the correct form of $\text{Re}[\Sigma_{\text{exc}}]$. From Eqs. (49) and (50), we may infer it to be

$$\text{Re}[\Sigma_{\text{exc}}(\mathbf{Q} = 0, \omega)] = \alpha^2 \omega \ln \left(\frac{\sqrt{\omega^2 + (\beta\mu)^2}}{\mu} \right), \quad (56)$$

where we have rewritten the maximum-function with logarithmic accuracy using a square root. This cutoff of logarithmic singularities (which are responsible for edge power laws) by recoil effects is a generic feature of our model, and will reoccur in the regime of $\mu \gg E_B$ presented in Sec. VI. In qualitative terms, this is also discussed in Ref. [22] (for arbitrary dimensions). Our results are in full agreement with this work.

We may now deduce the full photon self-energy Π_{exc} as follows. In the full finite-mass version of the power law (28), the real part of the logarithm in the exponent will be replaced by the cutoff logarithm from Eq. (56). The imaginary part of this logarithm will be some function $f(\omega)$, which continuously interpolates between the finite-mass regime for $\omega \ll \beta\mu$ [given by Eq. (53) times ω^{-1}], and the infinite mass regime for $\omega \gg \beta\mu$. Therefore we arrive at

$$\Pi_{\text{exc}}(\mathbf{Q} = 0, \omega)$$
$$= \frac{\Delta^2}{\omega + i0^+} \exp \left[\alpha^2 \left(\ln \left(\frac{\sqrt{\omega^2 + (\beta\mu)^2}}{\mu} \right) - i f(\omega) \right) \right], \quad (57)$$

where

$$f(\omega) = \begin{cases} \pi \sqrt{\frac{\omega}{\beta\mu}} \theta(\omega), & \omega \ll \beta\mu \\ \pi, & \omega \gg \beta\mu \end{cases}. \quad (58)$$

It is seen by direct inspection that (57) has a delta peak at $\omega = 0$ with weight $\Delta^2 \beta^{\alpha^2}$.

One can also asses the weight of the delta peak by comparing the spectral weights of the exciton spectral function in the infinite and finite mass cases: the weight of the delta peak must correspond to the difference in spectral weight as the absorption frequency power law is changed once β becomes finite. In the infinite mass case, the absorption scales as

$$A_\infty(\omega) \sim \frac{\Delta^2 \alpha^2}{\omega} \left(\frac{\omega}{\mu} \right)^{\alpha^2} \theta(\omega), \quad (59)$$

as follows from Eq. (28) above. Thus the spectral weight in the relevant energy region is given by

$$\int_0^{\beta\mu} d\omega A_\infty(\omega) = \Delta^2 \beta^{\alpha^2}. \quad (60)$$

In contrast, using Eq. (53), the spectral weight of the finite mass case is

$$\int_0^{\beta\mu} d\omega A(\mathbf{Q} = 0, \omega) = \Delta^2 \alpha^2. \quad (61)$$
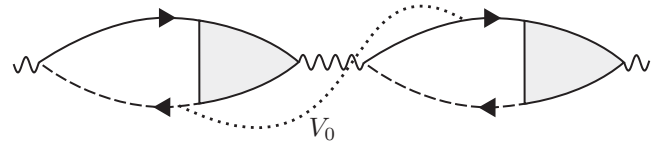

FIG. 14. Two dressed bubbles, connected by one electron-hole interaction (dotted line). This is an example of a photon self-energy diagram that is not contained in our approximation for $\Pi(\mathbf{Q}, \omega)$.

For scattering phase shifts δ close to π (i.e., $\alpha \to 0$), and for finite mass, $\beta > 0$, a pole with weight proportional to β^{α^2} [Eq. (60)] at $\omega = 0$ should be present in the spectrum, if β is not exponentially small in α. This weight is exactly the same as for the heavy hole when computed in a second-order cumulant expansion [23].

The full imaginary part of $\Pi_{\text{exc}}(\mathbf{Q} = 0, \omega)$ was already given explicitly in Eqs. (5a) and (5b), and plotted in Fig. 2. That plot illustrates the main conclusion of this section: for finite mass, Fermi sea excitations with large momentum transfer are energetically unfavorable, and are therefore absent from the absorption power law. As a result, the polelike features of the absorption are recovered.

C. Validity of the electron-hole correlator as a photon self-energy

Let us now assess the validity of the expressions for the CB electron-VB hole correlator [Eqs. (28) and (57)] as a photon self-energy. Using them, one assumes that only electron-hole interactions within one bubble are of relevance, and electron-hole interactions connecting two bubbles (an example is shown in Fig. 14) can be disregarded.

The regime where such an approximation is valid may be inferred from the following physical argument: electronic processes (i.e., electron-hole interactions) happen on the time scale of Fermi time $1/\mu$. On the other hand, the time scale for the emission and reabsorption of a photon (which is the process separating two bubbles) is given by $1/\rho d_0^2$ (where d_0 is the dipole matrix element). If the second scale is much larger than the first one, electrons and holes in distinct bubbles do not interact. Thus the our approach is valid as long as

$$\rho d_0^2 \ll \mu. \quad (62)$$

Under this condition, the following physical picture is applicable: an exciton interacts with the Fermi sea, giving rise to a broadened exciton, which in turn couples to the cavity photons. When Eq. (62) is violated, one should think in different terms: excitons couple to photons, leading to exciton-polaritons. These then interact with the Fermi sea. The second scenario is, however, beyond the scope of this paper. The above discussion is likewise valid for the regime of large Fermi energy, which is studied below.

VI. ELECTRON-HOLE CORRELATOR FOR LARGE FERMI ENERGY

We now switch to the opposite regime, where $\mu \gg E_B$, and excitons are not well-defined. For simplicity, we also assume that μ is of the order of the CB bandwidth. Hence
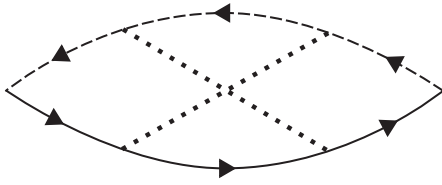
FIG. 15. Lowest-order crossed diagram contributing to the FES.

$E_B \ll \mu \simeq \xi$. Within our simplified model, the finite mass problem in 3D was solved in Ref. [18]. This treatment can be straightforwardly carried over to 2D [40]. To avoid technicalities, we will, however, just show how to obtain the 2D results in a "Mahan guess" approach [7], matching known results from [21]. To this end, we will first recapitulate the main ingredients of the infinite mass solution.

A. Infinite hole mass

The FES builds up at the Burstein-Moss shifted threshold $\Omega_T^{\text{FES}} = E_G + \mu$. Its diagrammatic derivation relies on a weak-coupling ansatz: the parameter $g = \rho V_0$ is assumed to be small. As seen from Eq. (24), this is indeed true for $\mu \gg E_0$. In principle, below the FES there will still be the exciton peak; however, this peak will be broadened into a weak power law, and thus merge with the FES. For finite mass (see below), the position of the would-be exciton may even be inside FES continuum, which makes the exciton disappear completely. What is more, the exciton weight, being proportional to E_B, is exponentially small in g (since $\mu \simeq \xi$). We may therefore safely disregard the exciton altogether (see also discussion in Appendix A).

To leading order in $g \ln(\omega/\mu)$, the dominant contribution comes from the so called "parquet" diagrams, containing all possible combinations of ladder and crossed diagrams [8,9]. The value of the pure ladder diagrams is given by Eq. (18), with $\Omega - E_G$ replaced by $\omega = \Omega - \Omega_T^{\text{FES}}$. The lowest-order crossed diagram is shown in Fig. 15. With logarithmic accuracy the contribution of this diagram is easily computed:

$$\Pi_{\text{crossed}} = -\tfrac{1}{3} d_0^2 \rho g^2 [\ln(\omega/\mu)]^3. \tag{63}$$

This is $-\tfrac{1}{3}$ times the contribution of the second order ladder diagram, cf. Eq. (18). Thus, the ladder and crossed channels partially cancel each other, a feature which persists to all orders. This also shows that the FES is qualitatively different from the broadened exciton discussed in the previous section: now the exciton effects (ladder diagrams) and the Fermi sea shakeup (crossed diagrams) have to be treated on equal footing.

In his original paper, Mahan computed all leading diagrams to third order and guessed the full series from an exponential ansatz [7]. The corresponding result for the photon self-energy $\Pi_{\text{FES}}(\omega)$ reads

$$\Pi_{\text{FES}}(\omega) = \frac{d_0^2 \rho}{2g} \left\{ 1 - \exp\left[-2g \ln\left(\frac{\omega + i0^+}{-\mu} \right) \right] \right\}. \tag{64}$$

Relying on coupled Bethe-Salpeter equations in the two channels (ladder and crossed), Nozières *et al.* then summed all parquet diagrams, where a bare vertex is replaced by (anti-)parallel bubbles any number of times [8,9]. The result corresponds exactly to Mahan's conjecture, Eq. (64).

By the standard FES identification $\delta/\pi = g + \mathcal{O}(g^3)$, the power law in Eq. (64) coincides with the one given in Eq. (28); the phase shift is now small. One should also point out that the peaks in the spectra in the regimes of small μ (Fig. 2) and large μ (Fig. 3) are not continuously connected, since the FES arises from the continuous threshold, whereas the exciton does not. Let us finally note that since μ is a large scale, Eq. (64) should be a good approximation for the photon self-energy, since the condition (62) is easily satisfied.

B. Finite hole mass

As in the regime of the exciton, in the finite mass case the result (64) will be modified due to the recoil energy $\beta \mu$. However, it will now be the *VB hole* recoil (or the hole lifetime, see below) instead of the exciton recoil—the latter is meaningless since the exciton is not a well defined entity anymore. This is most crucial: since CB states with momenta smaller than k_F are occupied, VB holes created by the absorption of zero-momentum photons must have momenta larger than k_F. Therefore the hole energy can actually be lowered by scatterings with the Fermi sea that change the hole momenta to some smaller value, and these scattering processes will cut off the sharp features of $\Pi_{\text{FES}}(\omega)$. The actual computation of the photon self-energy with zero photon momentum, $\Pi_{\text{FES}}(\mathbf{Q} = 0, \omega)$, proceeds in complete analogy to the 3D treatment of Ref. [18]. Limiting ourselves to the "Mahan guess" for simplicity, the main steps are as follows.

The first major modification is the appearance of two thresholds; as easily seen by the calculation of the ladder diagrams, the finite mass entails a shift of the pole of the logarithm from $\omega = 0$ to $\beta \mu$, which is the minimal energy for direct transitions obeying the Pauli principle. Correspondingly, $\omega_D = \beta \mu$ is called the direct threshold. Near this threshold, logarithmic terms can be large, and a nonperturbative resummation of diagrams is required. However, the true onset of 2DEG absorption will actually be the indirect threshold $\omega_I = 0$. There, the valence-band hole will have zero momentum, which is compensated by a low-energy conduction electron-hole pair, whose net momentum is $-k_F$. The two thresholds were shown in Fig. 4. It should be noted that for $E_B < \beta \mu$ the exciton energy $\approx \omega_D - E_B$, is between ω_I and ω_D. Hence, in this case, the exciton overlaps with the continuum and is completely lost.

Near ω_I, the problem is completely perturbative. In leading (quadratic) order in g, the absorption is determined by two diagrams only. The first one is the crossed diagram of Fig. 15. The second one is shown in Fig. 16. When summing these two diagrams, one should take into account spin, which will simply multiply the diagram of Fig. 16 by a factor of two (if the spin is disregarded, the diagrams will cancel in leading order). Up to prefactors of order one, the phase-space restrictions then result in a 2DEG absorption (see Ref. [21] and Appendix E):

$$A(\mathbf{Q} = 0, \omega) = d_0^2 g^2 \left(\frac{\omega}{\beta \mu} \right)^3 \theta(\omega). \tag{65}$$

The phase space power law ω^3 is specific to 2D. Its 3D counterpart has a larger exponent, $\omega^{7/2}$ [21], due to an additional restriction of an angular integration.

Let us now turn to the vicinity of ω_D, where one has to take into account the logarithmic singularities and the finite
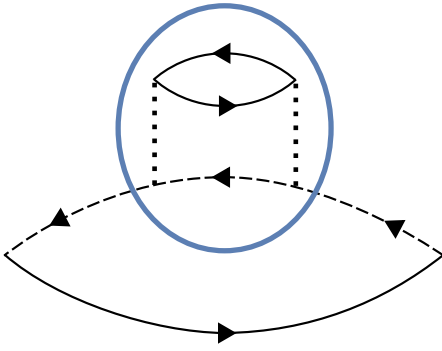
FIG. 16. Second diagram (in addition to Fig. 15) contributing to the absorption at the indirect threshold ω_I. The blue ellipse marks the VB self-energy insertion used below.

FIG. 17. The CB electron-VB hole bubble, with the hole propagator dressed by the self-energy, Eq. (67).

hole lifetime in a consistent fashion. Regarding the latter, one can dress all VB lines with self-energy diagrams as shown in Fig. 16. The self-energy insertion at the dominant momentum $k = k_F$ reads

$$\mathrm{Im}[\Sigma_{\mathrm{VB}}(k_F,\omega)] = \frac{1}{\sqrt{3}}\theta(\omega)g^2\beta\mu\frac{\omega^2}{(\beta\mu)^2}, \quad \omega \ll \beta\mu. \quad (66)$$

As can be shown by numerical integration, this expression reproduces the correct order of magnitude for $\omega = \beta\mu$, such that it can be safely used in the entire interesting regime

$\omega \in [0,\beta\mu]$. The power law in Eq. (66) is again specific to 2D. In contrast, the order of magnitude of the inverse lifetime is universal,

$$\mathrm{Im}[\Sigma_{\mathrm{VB}}(k_F,\beta\mu)] \sim g^2\beta\mu. \quad (67)$$

Disregarding the pole shift arising from $\mathrm{Re}[\Sigma]$, the self-energy (67) can be used to compute the "dressed bubble" shown in Fig. 17. With logarithmic accuracy, the dressed bubble can be evaluated analytically. In particular, its real part reads

$$\mathrm{Re}[\Pi_{\mathrm{db}}](\omega) \simeq \rho d_0^2 \ln\left(\frac{\sqrt{(\omega - \beta\mu)^2 + (g^2\beta\mu)^2}}{\mu}\right). \quad (68)$$

This is just a logarithm whose low-energy divergence is cut by the VB hole life time, in full analogy to Eq. (56), and in agreement with Ref. [22].

For the computation of polariton spectra later on, it turns out to be more practical to obtain both the real and the imaginary parts of $\Pi_{\mathrm{db}}(\omega)$ by numerically integrating the approximate form [40]:

$$\Pi_{\mathrm{db}}(\omega) \simeq \frac{d_0^2}{(2\pi)^2}\int_{k>k_F} d\mathbf{k}\frac{1}{\omega - (\epsilon_{\mathbf{k}} - \mu) - \frac{k^2}{2M} + i\mathrm{Im}[\tilde{\Sigma}_{\mathrm{VB}}(\omega - \epsilon_{\mathbf{k}} + \mu)]},$$

$$\mathrm{Im}[\tilde{\Sigma}_{\mathrm{VB}}(x)] = \begin{cases} \frac{g^2}{\sqrt{3}}\theta(x)\frac{x^2}{(\beta\mu)}, & x < \beta\mu \\ \frac{g^2}{\sqrt{3}}\beta\mu, & x > \beta\mu \end{cases},$$

(69)

to avoid unphysical spikes arising from the leading logarithmic approximation. A corresponding plot of $-\mathrm{Im}[\Pi_{\mathrm{db}}]$ is shown in Fig. 18. The numerical expression $-\mathrm{Im}[\Pi_{\mathrm{db}}]$ simplifies to the correct power law (65) in the limit $\omega \to 0$, and approaches the infinite mass value $d_0^2\rho\pi$ for large frequencies.

Higher-order diagrams will contain higher powers of the rounded logarithm (68). The parameter controlling the leading log scheme now reads

$$l \equiv g \ln(\beta g^2). \quad (70)$$

One can distinguish different regimes of l. The simplest is $l \ll 1$, which holds in the limit $g \to 0$ (or, put differently, if β is not exponentially small in g). In this limit, no singularity is left. The large value of the Fermi energy (small g) and the large value of the hole decay $\beta\mu$ have completely overcome all interaction-induced excitonic effects. A decent approximation to the 2-DEG absorption is then already given by the imaginary part of the dressed bubble. Figure 18 shows the corresponding absorption.

The more interesting regime corresponds to $g \ln(\beta g^2) \gtrsim 1$, where arbitrary numbers of conduction-band excitations
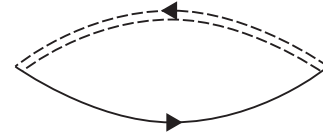
contribute to the absorption alike.[3] A nonperturbative summation is needed, which is, however, obstructed by the following fact: As found by straightforward computation, the crossed diagrams are not only cut by $g^2\beta\mu$ due to the hole decay, but also acquire an inherent cutoff of order $\beta\mu$ due to the hole recoil. A standard parquet summation is only possible in a regime where these two cutoffs cannot be distinguished with logarithmic accuracy, i.e., where $\beta \ll g^2$. For small enough g this will, however, always be the case in the truly nonperturbative regime where β must be exponentially small in g.

As a result of these considerations, the logarithms of the parquet summation have to be replaced by the cut-off logarithms (68), with $g^2\beta\mu$ replaced by $\beta\mu$. The imaginary part of the logarithm is then given by the function plotted in Fig. 18. The resulting full photon self-energy in the

_____

[3]The regime of $g \ln(\beta g^2) \gg 1$ is out of reach for the methods used in Ref. [18]. To study it, a consistent treatment of the divergences is needed, similar to Ref. [9]. We will not attempt this here.
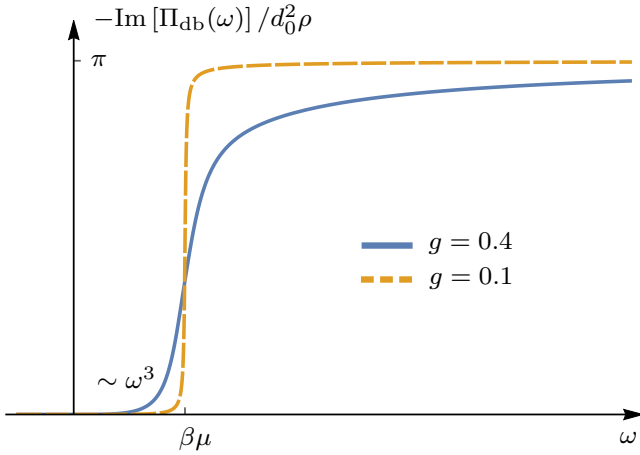
FIG. 18. Imaginary part of the dressed bubble for two values of g, obtained from numerical integration of Π_{db}, using the hole self-energy insertion of (66).

nonperturbative FES regime reads

$$\Pi_{\mathrm{FES}}(\mathbf{Q}=0,\omega) \simeq -\frac{d_0^2 \rho}{2g}\left\{ \exp\left[-2g\left(\frac{\Pi_{\mathrm{db}}(\omega)}{\rho d_0^2}\right)\right] - 1 \right\}. \tag{71}$$

A sketch of Im[Π_{FES}] is shown in Fig. 3.

VII. POLARITON PROPERTIES

When the cavity energy ω_c is tuned into resonance with the excitonic 2DEG transitions, the matter and light modes hybridize, resulting in two polariton branches. We will now explore their properties in the different regimes.

A. Empty conduction band

To gain some intuition, it is first useful to recapitulate the properties of the exciton-polariton in the absence of a Fermi sea. Its (exact) Green's function is given by Eq. (13), with $\omega_{\mathbf{Q}=\mathbf{0}} = \omega_c$ and $\Pi(\omega) = \Delta^2/(\omega + i0^+)$, where Δ is a constant (with units of energy) which determines the polariton splitting at zero detuning. In terms of our exciton model, one has $\Delta = \sqrt{d_0^2 \rho E_B/g^2}$. ω is measured from the exciton pole. A typical density plot of the polariton spectrum $A_p = -\mathrm{Im}[D^R(\omega,\omega_c)]/\pi$, corresponding to optical (absorption) measurements as, e.g., found in Ref. [13], is shown in Fig. 19. A finite cavity photon linewidth $\Gamma_c = \Delta$ is used. The physical picture is transparent: the bare excitonic mode (corresponding to the vertical line) and the bare photonic mode repell each other, resulting in a symmetric avoided crossing of two polariton modes.

For analytical evaluations, it is more transparent to consider an infinitesimal cavity linewidth Γ_c. The lower and upper polaritons will then appear as delta peaks in the polariton spectral function, at positions

$$\omega_\pm = \tfrac{1}{2}\left(\omega_c \pm \sqrt{\omega_c^2 + 4\Delta^2}\right) \tag{72}$$
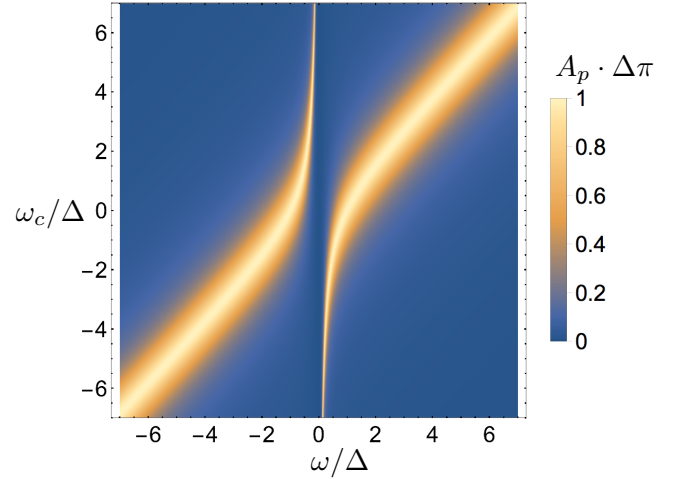
FIG. 19. $\mu = 0$: exciton-polariton spectrum as function of cavity detuning ω_c and energy ω, measured in units of the half polariton splitting Δ, with $\Gamma_c = \Delta$.

and with weights

$$W_\pm = \frac{1}{1 + \frac{4\Delta^2}{(\omega_c \pm \sqrt{4\Delta^2 + \omega_c^2})^2}}. \tag{73}$$

We note that the maximum of the polariton spectra scales as $1/\Gamma_c$ for finite Γ_c. Our spectral functions are normalized such that the total weight is unity. From Eq. (73) it is seen that the weight of the "excitonic" polaritons (corresponding to the narrow branches of Fig. 19) decays as Δ^2/ω_c^2 for large absolute values of ω_c.

B. Large Fermi energy

Let us study polariton properties in the presence of a Fermi sea. Reverting the order of presentation previously taken in the paper, we first turn to the regime of large Fermi energy, $E_B \ll \mu$. This is because for $E_B \ll \mu$ the inequality $\rho d_0^2 \ll \mu$ (62) is more easily satisfied than in the opposite limit of $E_B \gg \mu$, facilitating experimental realization. We compute the polariton properties using the electron-hole correlators as cavity photon self-energy. A similar approach was applied recently by Averkiev and Glazov [15], who computed cavity transmission coefficients semiclassically, phenomenologically absorbing the effect of the Fermi-edge singularity into the dipole matrix element. Two further recent treatments of polaritons for nonvanishing Fermi energies are found in Refs. [16,17]. In the first numerical paper [16], the Fermi-edge singularity as well as the excitonic bound state are accounted for, computing the electron-hole correlator as in Ref. [11], but an infinite mass is assumed. The second paper [17] is concerned with finite mass. However, the authors only use the ladder approximation and neglect the crossed diagrams, partially disregarding the physical ingredients responsible for the appearance of the Fermi-edge power laws. We aim here to bridge these gaps and describe the complete picture in the regime of large Fermi energy (before turning to the opposite regime of $\mu \ll E_B$).
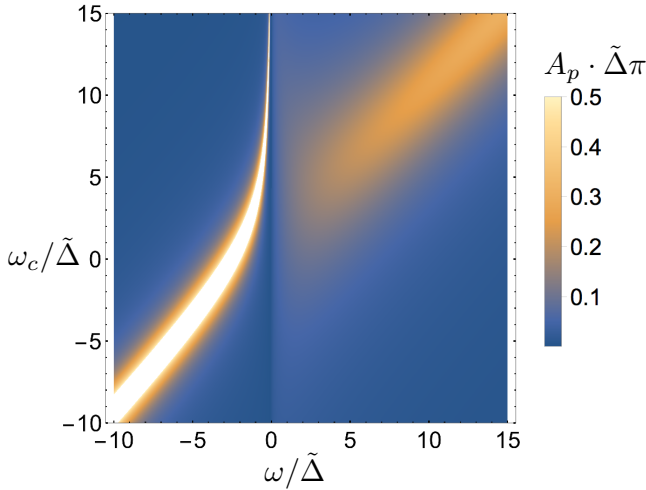
FIG. 20. $\mu \gg E_B$: infinite hole mass Fermi-edge-polariton spectrum $A_p(\omega, \omega_c)$ as a function of cavity detuning ω_c and energy ω, measured in units of the effective splitting $\tilde{\Delta}$. It was obtained by inserting Eqs. (74) and (75) into Eq. (14). Parameter values: $\mu = 30\tilde{\Delta}$, $\Gamma_c = 1\tilde{\Delta}$, and $g = 0.25$.

In the infinite mass limit we will use Eq. (64) as the photon self-energy. It is helpful to explicitly write down the real and imaginary parts of the self-energy in leading order in g:

$$\mathrm{Re}[\Pi_{\mathrm{FES}}](\omega) = \tilde{\Delta}\left(1 - \left(\frac{|\omega|}{\mu}\right)^{-2g}\right), \qquad (74)$$

$$\mathrm{Im}[\Pi_{\mathrm{FES}}](\omega) = -\tilde{\Delta}\, 2\pi g\left(\frac{\omega}{\mu}\right)^{-2g}\theta(\omega), \qquad (75)$$

$$\tilde{\Delta} \equiv \frac{d_0^2 \rho}{2g}, \qquad (76)$$

where we have introduced the parameter $\tilde{\Delta}$, which determines the splitting of the polaritons, playing a similar role to Δ in the previous case of empty CB. In the following, $\tilde{\Delta}$ will serve as the unit of energy.

For a cavity linewidth $\Gamma_c = 1\tilde{\Delta}$, a typical spectral plot of the corresponding "Fermi-edge polaritons" is shown in Fig. 20. It is qualitatively similar to the results of Ref. [15]. A quantitative comparison to the empty CB case is obviously not meaningful due to the appearance of the additional parameters μ (units of energy) and g (dimensionless). Qualitatively, one may say the following. The lower polariton is still a well-defined spectral feature. For zero cavity linewidth (see below), its lifetime is infinite. The upper polariton, however, is sensitive to the high-energy tail of the 2DEG absorption power law (75), and can decay into the continuum of CB particle-hole excitations. Its linewidth is therefore strongly broadened. Only when the 2DEG absorption is cut off by finite bandwidth effects (i.e., away from the Fermi edge), a photoniclike mode reappears in the spectrum (seen in the upper right corner of Fig. 20).

For more detailed statements, one can again consider the case of vanishing cavity linewidth Γ_c. A spectral plot with the same parameters as in Fig. 20, but with small cavity linewidth, $\Gamma_c = 0.01\tilde{\Delta}$, is shown in Fig. 21(a).
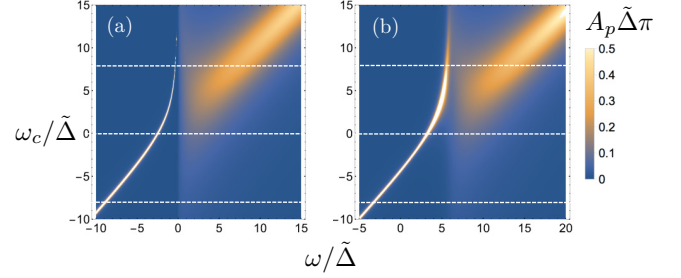

FIG. 21. $\mu \gg E_B$: (a) Fermi-edge-polariton spectrum with the same parameters as in Fig. 20, but $\Gamma_c = 0.01\tilde{\Delta}$. The white dashed lines denote the location of the spectral cuts presented in Fig. 22. (b) Spectrum with a nonzero mass-ratio $\beta = 0.2$, and otherwise the same parameters as in (a). This plot was obtained by inserting the finite mass photon self-energy of Eq. (71) into Eq. (14), with ω_c replaced by $\omega_c + \beta\mu$ to make sure that the cavity detuning is measured from the *pole* of the photon self-energy. Note that the frequency range of (b) is shifted as compared to (a).

We first examine the lower polariton (assuming zero linewidth), which is a pure delta peak. Its position is determined by the requirement

$$\omega - \omega_c - \mathrm{Re}[\Pi_{\mathrm{FES}}(\omega)] = 0. \qquad (77)$$

One may study the solution of this equation in three distinct regimes, corresponding to $\omega_c \to -\infty$, $\omega_c = 0$, and $\omega_c \to +\infty$.

For $\omega_c \to -\infty$, the solution of Eq. (77) approaches $\omega = \omega_c$, and the lower polariton acquires the full spectral weight (unity): for strong negative cavity detunings, the bare cavity mode is probed. The corresponding spectral cut is shown in Fig. 22(a) (continuous line). We will refrain from making detailed statements about the way the bare cavity mode is approached, since this would require the knowledge of the photon self-energy at frequencies far away from the threshold. As the cavity detuning is decreased, the lower polariton gets more matterlike. At zero detuning [see Fig. 22(b)], and for g not too small (with respect to $g\tilde{\Delta}/\mu$), the weight of the lower polariton is approximately given by $1/(1 + 2g)$. For large positive cavity detunings [see Fig. 22(c)], the position of the matterlike lower polariton approaches $\omega = 0$,

$$\omega \sim -\omega_c^{-1/(2g)} \quad \text{as} \quad \omega_c \to \infty. \qquad (78)$$

The lower polariton weight also scales in a power law fashion, $\sim \omega_c^{-1-1/(2g)}$, distinct from the excitonic regime, where the weight falls off quadratically [Eq. (73)].

Due to the finite imaginary part of the self-energy $\Pi_{\mathrm{FES}}(\omega)$, the upper polariton is much broader than the lower one: the photonic mode can decay into the continuum of matter excitations. At large negative detunings [see the inset to Fig. 22(a)], the upper polariton has a power law like shape (with the same exponent as the Fermi-edge singularity), and for $\omega_c \to -\infty$ its maximum approaches $\omega = 0$ from the high-energy side. As the detuning is increased (made less negative), the maximum shifts away from $\omega = 0$, approaching the free cavity mode frequency $\omega = \omega_c$ for $\omega_c \to \infty$. Since the weight and height are determined by the value of $\mathrm{Im}[\Pi_{\mathrm{FES}}]$ at the maximum, they increase correspondingly.
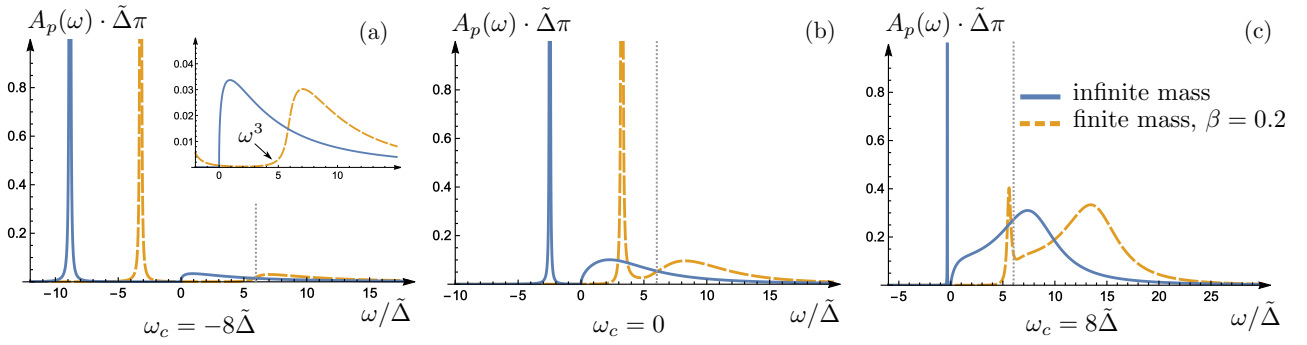
FIG. 22. $\mu \gg E_B$: spectral cuts at fixed cavity detuning through the polariton spectra of Fig. 21, for both infinite (continuous blue lines) and finite (dashed orange lines) hole mass. (a) Large negative cavity detuning. The dotted vertical line line always indicates the position of the direct threshold at $\omega = \beta\mu$. The inset is a zoom-in on the absorption onset at the indirect threshold. (b) Zero cavity detuning. (c) Large positive cavity detuning.

Let us now consider the case of finite mass. Using the finite mass photon self-energy (64) instead of (71), the Fermi-edge-polariton spectrum with a nonzero mass-ratio of $\beta = 0.2$ is plotted in Fig. 21(b). Compared to the infinite mass case of Fig. 21(a), Fig. 21(b) has the following important features. (i) The boundary line separating the lower and upper thresholds is shifted to the high-energy side from $\omega = 0$ in the infinite mass case to $\omega = \beta\mu$ in the finite mass case, reflecting the Burstein-Moss shift in the 2DEG absorption. (ii) As opposed to the infinite mass case, the lower polariton is strongly broadened at large positive detunings.

These points are borne out more clearly in Figs. 22(a)–22(c) (dashed lines), which presents cuts through Fig. 21(b) at fixed detuning. The situation at large negative detuning is shown in Fig. 22(a): compared to the infinite mass case, shown as full line, the polaritons are shifted towards higher energies. In addition, the shape of the upper polariton is slightly modified—its onset reflects the convergent phase-space power law ω^3 of Eq. (65) found for the 2DEG absorption. This is emphasized in the inset. At zero cavity detuning [Fig. 22(b)], the situation of the finite and infinite mass cases is qualitatively similar. When the cavity detuning is further increased, the position of the polelike lower polariton approaches the direct threshold at $\omega = \beta\mu$ (indicated by the vertical dotted line). When the pole is in the energy interval $[0, \beta\mu]$, the lower polariton overlaps with the 2DEG continuum absorption, and is therefore broadened. This is clearly seen in Fig. 22(c): instead of a sharp feature, there is just a small remainder of the lower polariton at $\omega = \beta\mu$. As a result, one may say that in the regime of the Fermi-edge singularity, i.e., large μ, the finite mass will cut off the excitonic features from the polariton spectrum—instead of the avoided crossing of Fig. 19, Fig. 21(b) exhibits an almost photoniclike spectrum, with a small (cavity) linewidth below the threshold at $\omega = \beta\mu$, and a larger linewidth above the threshold, reflecting the step-like 2DEG absorption spectrum of Fig. 3. The finite mass thus leads to a general decrease of the mode splitting between the two polariton branches. This trend continues when the Fermi energy is increased further.

It is instructive to compare this behavior with the experimental results reported in [13]. There, two differential reflectivity measurements were conducted, which can be qualitatively identified with the polariton spectra. The first

measurement was carried out using a low-mobility GaAs sample (which should behave similarly to the limit of large VB hole mass), and moderate Fermi energies. A clear avoided crossing was seen, with the upper polariton having a much larger linewidth than the lower one (see Fig. 2(A) of Ref. [13]). In the second measurement, the Fermi energy was increased further, and a high-mobility sample was studied, corresponding to finite mass. A substantial reduction of the mode splitting between the polaritons was observed (Fig. 2(C) of Ref. [13]). While a detailed comparison to the experiment of Ref. [13] is challenging, due to the approximations we made and the incongruence of the parameter regimes (in the experiment one has $\mu \simeq E_B$), the general trend of reduced mode splitting is correctly accounted for by our theory.

C. Small Fermi energy

We now switch to the regime of of small Fermi energy discussed in Sec. V, a regime in which the polariton spectra have not been studied analytically before. We again assume that the condition (62), required for the approximating the photon self-energy by Eq. (15), is fulfilled. This may be appropriate for systems with a large exciton-binding energies, e.g., transition metal dichalcogenide monolayers as recently studied in Ref. [14].

For infinite mass, we may use Eq. (28) as photon self-energy, multiplied by a prefactor $\Delta^2 = d_0^2 \rho E_B / g^2$ [cf. Eq. (55)], and expand the real and imaginary parts to leading order in $\alpha^2 = (\delta/\pi - 1)^2$. The energy ω is now measured from the exciton pole: $\omega = \Omega - \Omega_T^{\text{exc}}$, $\Omega_T^{\text{exc}} = E_G + \mu - E_B$. The corresponding polariton spectrum for a small cavity linewidth is shown in Fig. 23(a). Qualitatively, it strongly resembles the bare exciton case as in Fig. 19 (note that in Fig. 23 the cavity linewidth was chosen to be 100 times smaller than in Fig. 19), but with a larger linewidth of the upper polariton. This is due to the possible polariton decay into the particle hole continuum contained in the excitonic power law, Eq. (28).

The detailed discussion of polariton properties in the regime of $\mu \ll E_B$ parallels the previous discussion in the regime $E_B \ll \mu$. For small negative detuning ω_c [Fig. 24(a)], the lower polariton is found at approximately $\omega = \omega_c$. The upper polariton has a significantly smaller weight, its shape reflects
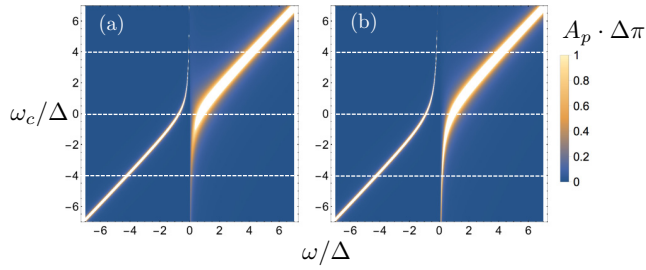
FIG. 23. $\mu \ll E_B$: exciton-polariton spectrum for small Fermi energy. The white dashed lines denote the location spectral cuts presented in Fig. 24. (a) Infinite mass. This plot was obtained by inserting the Exciton Green's function for $\mu \gtrsim 0$, given by Eq. (28) multiplied by $\Delta^2 = d_0^2 \rho E_B / g^2$, into the photon Green's function, Eq. (14). Parameters: $\mu = 10\Delta$, $\Gamma_c = 0.01\Delta$, and $\alpha^2 = (\delta/\pi - 1)^2 = 0.25$. (b) Finite mass, with mass ratio $\beta = 0.4$. In this plot, the finite mass Exciton Green's function, Eq. (57), was used, with the same parameters as in (a).

the excitonic power law of Eq. (28). However, compared to the previous spectral cuts (Fig. 22) the upper polariton peak is much more pronounced. This results from the exciton being now pole-like, as compared to the power law Fermi-edge singularity. Increasing the detuning, weight is shifted to the upper polariton. At zero detuning [Fig. 24(b)], the weight of the lower polariton is only order $\mathcal{O}(\alpha^2)$ larger than the weight of the upper polariton. At large positive detuning, the position of the lower polariton is found at approximately

$$\omega \sim -\omega_c^{-1/(1-\alpha^2)} \quad \text{as} \quad \omega_c \to \infty. \quad (79)$$

The lower polariton thus approaches the exciton line faster than in the pure exciton case, but slower than in the Fermi-edge regime [Eq. (78)]. A similar statement holds for the weight of the lower polariton, which scales as $\omega_c^{-2-\alpha^2}$.

The spectrum in the finite mass case is qualitatively similar, see Fig. 23(b). Quantitatively, a stronger peak repulsion can be seen, which may be attributed to the enhanced excitonic quasiparticle weight in the finite mass case. A comparison of spectral cuts in the finite mass case [Figs. 24(a)–24(c)] further corroborates this statement [especially in Fig. 24(c)]. Indeed, one finds that the position of the lower polariton at large cavity

detuning is approximately given by

$$\omega \sim -\beta^{\alpha^2} \cdot \omega_c^{-1} \quad \text{as} \quad \omega_c \to \infty, \quad (80)$$

i.e., the excitonic line at $\omega = 0$ is approached more slowly than in the infinite mass case, Eq. (79). The corresponding weight falls off as ω_c^{-2}. Thus the lower polariton has a slightly enhanced weight compared to the infinite mass case. In addition, in the spectral cut at large negative detuning, [inset to Fig. 24(a)], the upper polariton appears as a sharper peak compared to the infinite mass case, which again results from the enhanced quasiparticle weight of the finite mass case.

VIII. CONCLUSION

In this paper, we have studied the exciton-polariton spectra of a 2DEG in an optical cavity in the presence of finite CB electron density. In particular, we have elucidated the effects of finite VB hole mass, distinguishing between two regimes. In the first regime (small Fermi energy as compared to the exciton binding energy), we have found that excitonic features in the 2DEG absorption are enhanced by the exciton recoil and the resulting suppression of the Fermi edge singularity physics. In contrast, in the second regime of Fermi energy larger than the exciton binding energy, it is the VB hole which recoils at finite mass. This cuts off the excitonic features. These modifications also translate to polariton spectra, especially to the lower polariton at large cavity detuning, which is excitonlike. Our findings reproduce a trend seen in a recent experiment [13].

We would like to mention several possible extensions of this work. To begin with, it would be promising to study the effect of long-range interactions on the power laws, and hence on polariton spectra, from an analytical perspective. Long-range interactions are expected to be most important in the regime of small Fermi energy, leading to additional bound states and to the Sommerfeld enhancement effects [24]. Moreover, one should try to explore trionic features, for which it is necessary to incorporate the spin degree of freedom (to allow an electron to bind to an exciton despite the Pauli principle). Another interesting direction would be to tackle the limit of equal electron and hole masses, which is relevant to transition metal dichalcogenides, whose polariton spectra in the presence of a Fermi sea where measured in a recent experiment [14]. Lastly,
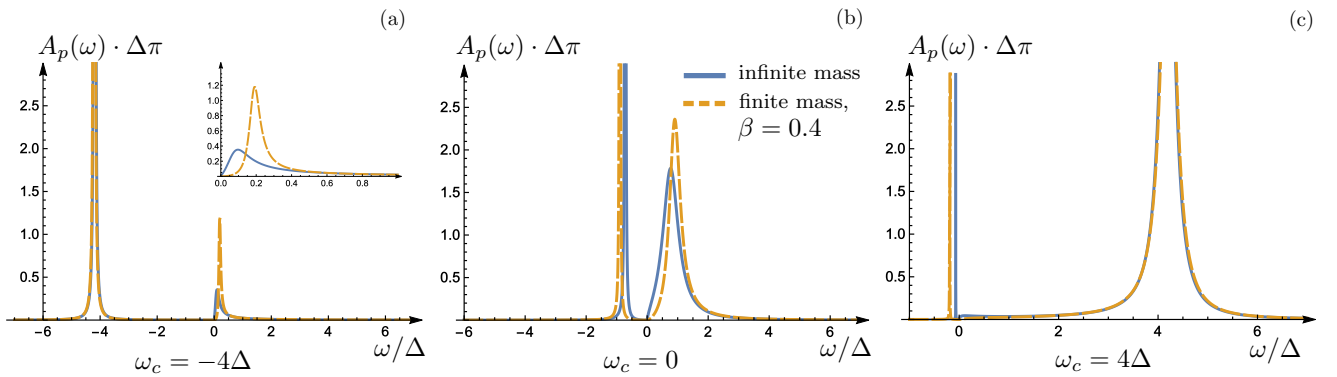

FIG. 24. $\mu \ll E_B$: spectral cuts at fixed cavity detuning through the polariton spectra of Fig. 23, for both infinite (continuous blue lines) and finite hole mass (dashed orange lines). (a) Large negative cavity detuning. The inset shows a zoom onto the upper polaritons. (b) Zero cavity detuning. (c) Large positive cavity detuning.

one should address the behavior of the polariton in the regime of small Fermi energy and strong light-matter interactions. Then, not the exciton, but rather the polariton interacts with the Fermi sea, and different classes of diagrams have to be resummed to account for this change in physics.

APPENDIX A: EVOLUTION OF ABSORPTION SPECTRA WITH INCREASING CHEMICAL POTENTIAL

In this Appendix, we present an extended overview of how the absorption spectra evolve inbetween the controlled extremal limits of $\mu \ll E_B$ and $\mu \gg E_B$. For $\mu \ll E_B$, the dominant spectral feature is the exciton. For finite mass ($\beta \neq 0$), it has a coherent delta-like part and an incoherent tail, see Eq. (5), while the infinite mass exciton ($\beta = 0$) is a purely incoherent power law, see Eq. (2). These pronounced excitonic features are well separated from the CB continuum part at $\Omega_T^{\mathrm{FES}} = E_G + \mu$ (see inset to Fig. 2).

As μ is increased, the incoherent exciton part [Eqs. (5b) and (2)] starts to overlap with the CB continuum part. Moreover, the overall relative weight of both the coherent and incoherent portions of the exciton part of the spectrum (which are both proportional to E_B) will diminish. Still, within our simplified model which neglects CB electron-CB electron interactions, and for $\beta = 0$, this exciton feature will never disappear completely, since in this model an infinite mass VB hole is simply a local attractive potential for the CB electrons, and such a potential will always have a bound state in 2D. However, for finite VB hole mass, the exciton energy (location of the coherent delta peak) will penetrate into the CB continuum when μ becomes larger than $E_B/\beta \gg E_B$ [i.e., when E_B crosses the indirect threshold, see Fig. 4(a)]. More importantly, CB electron-CB electron interactions would screen the hole potential, and will thus reduce the exciton binding energy and presumably eliminate the exciton part of the spectrum completely as soon as $\mu \gg E_B$.

To describe this situation, it has been customary in the literature [11,21] to still employ the same simplified model neglecting CB electron-CB electron interactions, but assume that the hole potential does not create a bound state for large enough μ, a practice we follow in this work as well. Then, for $\mu \gg E_B$, one should concentrate on the remaining, CB continuum part of the spectrum, which will evolve into the Fermi-edge singularity (FES), cut off by the VB hole recoil
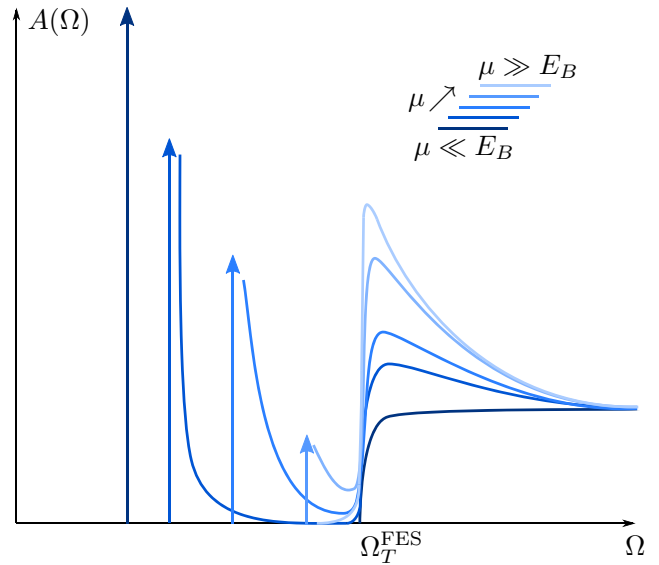

FIG. 25. Putative evolution of absorption spectra as μ is increased. The colored arrows represent delta-function peaks, their height corresponds to the relative weight of those peaks. The (hand-sketched) plots of this figure comprise the effects of a (large) finite VB hole mass ($\beta \neq 0$) and electron-electron interactions, beyond what's actually computed in this paper. For clarity, the shift of the spectra with increasing μ is disregarded. For μ even larger than shown in the sketch, the FES will reduce to a steplike feature again.

energy for $\beta \neq 0$. A putative evolution of absorption spectra with increasing μ is sketched in Fig. 25.

APPENDIX B: EVALUATION OF THE EXCITON SELF-ENERGY IN THE TIME-DOMAIN

In this Appendix, we present the time-domain evaluation of the exciton self-energy diagrams of Fig. 11. These diagrams contain one CB electron loop only, and therefore yield the leading contribution when μ/E_B is small. We will start with the direct diagrams [Fig. 11(a)] and then turn to the exchange series [Fig. 11(b)].

1. Direct diagrams

First, we note that the bare Green's functions in the time domain read

$$G_c^{(0)}(\mathbf{k},t) = -i(\theta(t) - n_{\mathbf{k}})e^{-i\epsilon_{\mathbf{k}}t}, \qquad (B1)$$

$$G_v^{(0)}(t) = i\theta(-t)e^{iE_G t}, \qquad (B2)$$

with the zero temperature Fermi function $n_{\mathbf{k}} = \theta(k_F - k)$. Using these, we will evaluate the series of direct diagrams of Fig. 11(a). The temporal structure of a generic direct diagram is illustrated via the example of Fig. 26.

To compute such a diagram, we make the following observation. Since the VB propagator has no momentum dependence, all VB phase factors simply add up to give a total factor of $e^{-iE_G t}$. Then, the step functions in the VB propagators enforce time ordering for the intermediate time integrals. In the specific case shown in Fig. 26,
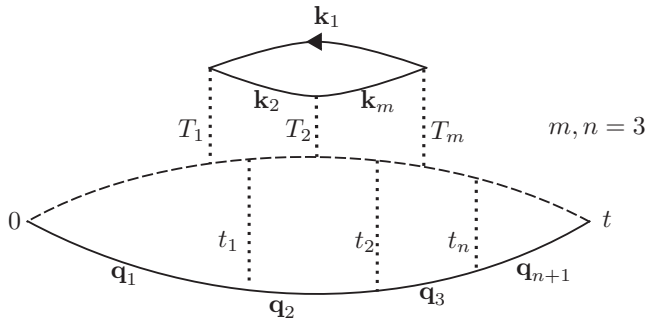
FIG. 26. A direct self-energy diagram in the time domain. The Green's function with an arrow indicates the CB electron propagating backwards in time.

$0 < T_1 < t_1 < T_2 < t_2 < T_m < t_n < t$ with $m = n = 3$ (m and n count the number of interaction lines above and

below the dashed VB line, respectively). However, there are also diagrams with $m = n = 3$, but with a different relative ordering of the interaction lines. Summing over all those diagrams for m and n fixed, one needs to integrate over the time ranges $0 < t_1 < \cdots < t_n < t \cap 0 < T_1 < \cdots < T_m < t$. This means that the time integration for the direct diagrams splits into a product of two functions, representing the propagation of a Fermi sea electron (above the VB line in Fig. 26) and a photoexcited electron (below the VB line) in the time-dependent potential.

We are now in the position to write down the full expression for the sum of direct diagrams D to all orders in the interaction, fixing the signs with Wick's theorem:

$$D(t) = - \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \, e^{-iE_G t} \tilde{B}(t) C(t), \tag{B3}$$

where

$$\tilde{B}(t) = \sum_{m=1}^{\infty} (-V_0)^m \int_{0 < T_1 < \cdots < T_m < t} dT_1 \cdots dT_m \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \cdots$$
$$\times \int_{k_m > k_F} \frac{d\mathbf{k}_m}{(2\pi)^2} \tilde{G}_c(\mathbf{k}_1, T_1 - T_m) \tilde{G}_c(\mathbf{k}_2, T_2 - T_1) \cdots \tilde{G}_c(\mathbf{k}_m, T_m - T_{m-1}), \tag{B4}$$

$$C(t) = \sum_{n=0}^{\infty} (-V_0)^n \int_{0 < t_1 < \cdots < t_n < t} dt_1 \cdots dt_n \int_{q_1 > k_F} \frac{d\mathbf{q}_1}{(2\pi)^2} \cdots \int_{q_{n+1} > k_F} \frac{d\mathbf{q}_{n+1}}{(2\pi)^2} \tilde{G}_c(\mathbf{q}_1, t_1) \tilde{G}_c(\mathbf{q}_2, t_2 - t_1) \cdots \tilde{G}_c(\mathbf{q}_{n+1}, t - t_n), \tag{B5}$$

and

$$\tilde{G}_c(\mathbf{k}_1, T_1 - T_M) = i e^{-i\epsilon_{\mathbf{k}_1}(T_1 - T_M)}, \quad \tilde{G}_c(\mathbf{p}, \tau) = -i e^{-i\epsilon_{\mathbf{p}}\tau} \quad \text{for} \quad \mathbf{p} \neq \mathbf{k}_1. \tag{B6}$$

Defining the retarded Green's function by

$$G_c^{0,R}(\mathbf{p}, \tau) = \theta(\tau) \tilde{G}(\mathbf{p}, \tau), \tag{B7}$$

we can rewrite the two factors appearing in $D(t)$ as sequences of convolutions:

$$B(t) \equiv e^{-i\epsilon_{\mathbf{k}_1} t} \tilde{B}(t) = \sum_{m=1}^{\infty} (-V_0)^m \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \cdots \int_{k_m > k_F} \frac{d\mathbf{k}_m}{(2\pi)^2} \big[G_c^{0,R}(\mathbf{k}_1,) * G_c^{0,R}(\mathbf{k}_2,) \cdots * G_c^{0,R}(\mathbf{k}_m,) * G_c^{0,R}(\mathbf{k}_1,) \big](t), \tag{B8}$$

$$C(t) = \sum_{n=0}^{\infty} (-V_0)^n \int_{q_1 > k_F} \frac{d\mathbf{q}_1}{(2\pi)^2} \cdots \int_{q_{n+1} > k_F} \frac{d\mathbf{q}_{n+1}}{(2\pi)^2} \big[G_c^{0,R}(\mathbf{q}_1,) * \cdots * G_c^{0,R}(\mathbf{q}_{n+1},) \big](t). \tag{B9}$$

Together, Eqs. (B3) and (B8)–(B9) correspond to Eq. (30) in the main text. Fourier transforming Eq. (B3) results in

$$D(\Omega) = \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \underbrace{i \int \frac{d\nu}{2\pi} B(\nu) C\big(\Omega - E_G + \epsilon_{\mathbf{k}_1} - \nu\big)}_{I(\Omega)}, \tag{B10}$$

where we defined $I(\Omega)$ for later purpose. The Fourier transform of $B(t)$ reads

$$B(\nu) = \sum_{m=1}^{\infty} (-V_0)^m \int_{k_2 > k_F} \frac{d\mathbf{k}_2}{(2\pi)^2} \cdots \int_{k_m > k_F} \frac{d\mathbf{k}_m}{(2\pi)^2} G_c^{0,R}(\mathbf{k}_1, \nu) \cdot G_c^{0,R}(\mathbf{k}_2, \nu) \cdots G_c^{0,R}(\mathbf{k}_m, \nu) \cdot G_c^{0,R}(\mathbf{k}_1, \nu), \tag{B11}$$

with retarded real frequency Green's functions:

$$G_c^{0,R}(\mathbf{k}, \nu) = \frac{1}{\nu - \epsilon_{\mathbf{k}} + i0^+}. \tag{B12}$$

Inserting (B12) into (B11), the integrations are trivially performed. The summation over interaction lines reduces to a

geometric series, yielding

$$B(\nu) = \frac{-V_0}{g} \frac{1}{(\nu - \epsilon_{\mathbf{k}_1} + i0^+)^2} \frac{1}{\ln\left(\frac{\nu - \mu + i0^+}{-E_B}\right)}, \tag{B13}$$

where we used $\ln(E_B/\xi) = -1/g$, cf. Eq. (20). For the term $C(\Omega - E_G + \epsilon_{\mathbf{k}_1} - \nu)$ appearing in (B10), we analogously

arrive at

$$C(\Omega - E_G + \epsilon_{\mathbf{k}_1} - \nu) = \frac{\rho}{g}\left(1 - \frac{1}{g\ln\left(\frac{\kappa - \nu + i0^+}{-E_B}\right)}\right),$$

$$\kappa \equiv \Omega - E_G + \epsilon_{\mathbf{k}_1} - \mu. \qquad (B14)$$

The functions $B(\nu)$ and $C(\nu)$ are difficult to integrate, because they each have both a pole and a branch cut, arising from the $1/\ln$ term. We can split these terms as follows:

$$\frac{1}{\ln\left(\frac{\nu - \mu + i0^+}{-E_B}\right)} = \frac{-E_B}{E_B + \nu - \mu + i0^+} + \left(\frac{1}{\ln\left(\frac{\nu - \mu + i0^+}{-E_B}\right)}\right.$$
$$\left. + \frac{E_B}{E_B + \nu - \mu + i0^+}\right). \qquad (B15)$$

The first term on the right-hand side of Eq. (B15) has just a simple pole, while the second one's only singularity is a branch cut. Using this representation, we can evaluate $I(\Omega)$ as defined in Eq. (B10) employing the following argument. Physically, the terms B, C represent the propagation of the two electrons in the hole potential. Comparing to the simple exciton ladder summation (see Sec. IV), we associate the poles of the

$1/\ln$-terms in these functions with the exciton contribution, while the branch cut corresponds to the continuum above the indirect threshold, $\Omega > E_G + \mu$.

Following these observations, let us split $I(\Omega)$ into a pole-pole, a pole-branch, and a branch-branch contributions. $I_{\text{branch-branch}}$ only contributes to the continuum part of the spectrum. More importantly (as explained in the main text), employing spectral representations of the retarded functions $B_{\text{branch}}, C_{\text{branch}}$, it is easily shown that $\text{Im}[I_{\text{branch-branch}}]$ (which is of potential importance for the lineshape of the exciton spectrum) vanishes for frequencies close to the exciton pole $(\omega \gtrsim 0)$. It is thus not important for our purposes.

Computing contour integrals, $I_{\text{pole-pole}}$ is easily evaluated to give

$$I_{\text{pole-pole}}(\omega) = \frac{E_B^2}{g^2}\frac{1}{(\omega + i0^+)^2}\frac{1}{E_B + \omega + \epsilon_{\mathbf{k}_1} - \mu + i0^+},$$
$$(B16)$$

where energies are measured from the exciton pole, $\omega = \Omega - (E_G + \mu) + E_B$. This contribution gives rise to trionic features in the spectrum, which are shortly discussed in Appendix C.

Last, computing contour integrals and disregarding terms which are subleading in ω/E_B, the pole-branch contribution is found to be

$$I_{\text{pole-branch}}(\omega) \simeq \frac{-E_B}{g^2}\frac{1}{(\omega + i0^+)^2}\left(\frac{1}{\ln\left(\frac{\omega + \epsilon_{\mathbf{k}_1} - \mu + i0^+}{-E_B}\right)} + \frac{E_B}{E_B + \omega + \epsilon_{\mathbf{k}_1} - \mu + i0^+}\right). \qquad (B17)$$

Inserting the Eqs. (B16) and (B17) into Eq. (B10), one finally arrives at Eq. (32) of the main text.

2. Exchange diagrams

The computation of the exchange diagrams, though technically sligthly more involved, essentially proceeds along the same lines. The general time-structure of an exchange diagram is illustrated in Fig. 27. As for the direct diagrams, the VB propagators just enforce a time ordering. In addition, there is the condition $t_n > T_1$. When this condition is violated, the diagram reduces to a ladder diagram, which must be excluded to avoid double counting. Taking this into account, the full expression for the sum of exchange diagrams reads

$$X(t) = \sum_{m,n=1}^{\infty}(-V_0)^{m+n}e^{-iE_Gt}\int_{0 < T_1 < \cdots < T_m < t}dT_1\cdots dT_m\int_{T_1}^{t}dt_n\int_0^{t_n}dt_{n-1}\cdots\int_0^{t_2}dt_1\int_{k_1 < k_F}\frac{d\mathbf{k}_1}{(2\pi)^2}\int_{k_2 > k_F}\frac{d\mathbf{k}_2}{(2\pi)^2}\cdots$$

$$\int_{k_{m+1} > k_F}\frac{d\mathbf{k}_{m+1}}{(2\pi)^2}\int_{q_1 > k_F}\frac{d\mathbf{q}_1}{(2\pi)^2}\cdots\int_{q_n > k_F}\frac{d\mathbf{q}_n}{(2\pi)^2}\tilde{G}_c(\mathbf{k}_1, T_1 - t_n)\tilde{G}_c(\mathbf{k}_2, T_2 - T_1)\cdots$$

$$\tilde{G}_c(\mathbf{k}_{m+1}, t - T_m)\tilde{G}_c(\mathbf{q}_1, t_1)\cdots\tilde{G}_c(\mathbf{q}_n, t_n - t_{n-1}). \qquad (B18)$$

To rewrite (B18) as a sum of convolutions, one can employ the following easily derived formula:

$$\mathcal{F}\left(\int_{-\infty}^{\infty}dt_1 f(t - t_1)g(t, t_1)\right)(\Omega) = \int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}f(\omega_1)g(\Omega - \omega_1, \omega_1), \qquad (B19)$$

where $\mathcal{F}$ denotes the Fourier transform, and f and g are any two well-behaved functions. Applying this result, a computation similar to the one for $D(\Omega)$ shows that the Fourier-transform of Eq. (B18) can be expressed as

$$X(\Omega) = -\int_{k_1 < k_F}\frac{d\mathbf{k}_1}{(2\pi)^2}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}(-g)\ln\left(\frac{\omega_1 - \mu + i0^+}{-\xi}\right)\frac{1}{1 + g\ln\left(\frac{\omega_1 - \mu + i0^+}{-\xi}\right)}\frac{1}{\Omega - E_G - \omega_1 + i0^+}$$

$$\times \int_{-\infty}^{\infty}\frac{d\omega_2}{2\pi}(-g)\ln\left(\frac{\omega_2 - \mu + i0^+}{-\xi}\right)\frac{1}{1 + g\ln\left(\frac{\omega_2 - \mu + i0^+}{-\xi}\right)}\frac{1}{-\omega_2 + \Omega - E_G + i0^+}\frac{1}{\omega_2 + \omega_1 - \Omega + E_G - \epsilon_{\mathbf{k}_1} - i0^+}.$$
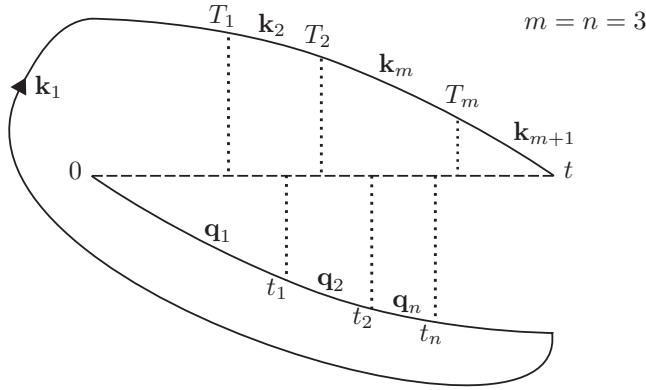
$$(B20)$$

FIG. 27. An exchange self-energy diagram in the time domain. The Green's function shown with an arrow indicates the CB electron propagating backwards in time.

This expression can be evaluated as before, splitting it into pole-pole, pole-branch, and branch-branch contributions using Eq. (B15). In complete analogy to the direct diagrams, the imaginary part of the branch-branch contribution can be shown not to contribute in the regime of interest to us, and we therefore disregard it completely. Straightforwardly evaluating the pole-pole and pole-branch contributions, one ultimately arrives at Eq. (34) in the main text.

APPENDIX C: TRION CONTRIBUTION TO THE EXCITON SELF-ENERGY DIAGRAMS

The pole-pole contribution to the direct self-energy $D(\omega)$ [Eq. (30)] physically represents two electrons tightly bound to the hole potential. Indeed, it assumes the form

$$D_{\text{pole-pole}}(\omega) = \int_{k_1 < k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} I_{\text{pole-pole}}(\omega), \qquad \text{(C1)}$$

where $I_{\text{pole-pole}}$ is given in Eq. (B16). $I_{\text{pole-pole}}$ can be identified with a bare trion Green's function, since it has a pole at $\omega = -E_B + \mu - \epsilon_{\mathbf{k}_1}$, corresponding to the binding of a second CB electron to the exciton (recall that ω is measured from the exciton threshold), where the energy $\epsilon_{\mathbf{k}_1}$ of this second electron can be from anywhere in the Fermi sea. Evaluation of (C1) close to the trion resonance $\omega \simeq -E_B$ leads to

$$D_{\text{pole-pole}}(\omega) \simeq \frac{\rho}{g^2} \ln\left(\frac{E_B + \omega + i0^+}{E_B + \omega - \mu}\right). \qquad \text{(C2)}$$

Using Eq. (33) of the main text, (C2) gives rise to a self-energy contribution to the exciton

$$\Sigma_{\text{exc}} = E_B \ln\left(\frac{E_B + \omega + i0^+}{E_B + \omega - \mu}\right). \qquad \text{(C3)}$$

This self-energy expression fully matches usual results found in works concerned with trions [14,35,37], apart from two minor differences: first, in these works the case of finite VB hole mass (of the same order as the CB mass) is considered, but reevaluation of (C3) for finite mass is straightforward and only results in some trivial factors involving mass ratios. Second, in the works cited above, the exciton is treated as an elementary entity, and the trion binding energy is therefore an adjustable parameter. By contrast, we have started from

a microscopic model, which does not contain excitons, and, accounting for exchange processes, computed excitons and trions along the way. As a result, our microscopic theory yields the same binding energy E_B for excitons and trions. However, this is clearly an artefact of disregarding electron-electron interactions (which would significantly reduce the trion binding energy), and can heuristically be accounted for by replacing E_B in Eq. (C3) by a trion binding energy $E_T \ll E_B$. Upon inserting (C3) into the exciton Green's function (27), one finds the following spectral features: First, there is a sharp resonance, red detuned with respect to the trion threshold by an order of μ, and with a weight that scales as μ/E_T. This peak is commonly called the trion, or, more appropriately, attractive polaron [14], since the trion bound state is not filled. Second, there is a small steplike feature for $0 < E_B + \omega < \mu$, arising from the imaginary part of (C3). This feature, where the trion bound state is filled and the second electron constituting the trion can come from anywhere in the Fermi sea, has smaller (but not parametrically smaller) weight than the attractive polaron, and is usually overlooked in the literature. Investigation of further trion properties is a worthwhile goal which we leave for further work.

Let us close this Appendix with a technical remark: of course, for spinless electrons, a trion cannot exist in our simple model of short range VB hole-CB electron interaction, due to the Pauli principle (two electrons cannot occupy the single bound state created by the hole). In line with that, the pole-pole contribution cancels in this case between the direct and exchange diagrams. However, in the spinful case, the direct contribution will incur a factor of two, so it does not cancel with the exchange contribution, so the trion remains.

APPENDIX D: THE SELF-ENERGY CONTRIBUTION OF THE EXCHANGE DIAGRAMS

The exchange contribution to the exciton self-energy, Eq. (35), can be understood by the following considerations. The ground-state energy of an N-particle system in the presence of an attractive delta-function potential strong enough to form a bound state is lower than the N-particle ground-state energy of the system without the potential by an amount

$$\Delta E = -E_B - (1 - \alpha)\mu, \qquad \text{(D1)}$$

which is the sum of the bound state energy E_B, and a second term which arises from the rearrangement of the Fermi sea, described by Fumi's theorem [39] [recalling that $1 - \alpha = \delta/\pi$, cf. Eq. (40)]. We find that the exchange diagrams give the contribution μ, while the term $\alpha\mu$ stems from the direct diagrams [Eq. (38)]. To create such an attractive potential, one has to lift one electron from the VB to the CB, which costs $E_G + \mu$. In our treatment, the extra cost μ appearing here is contained in the shift of the pole of the ladder diagrams, Eq. (25). Thus the minimal absorption energy predicted by our model is $E_G - E_B + \alpha\mu \approx E_G - E_B$.

At first sight, this seems to contradict the experimental results (e.g., Ref. [14]), according to which the minimal absorption energy is $E_G - E_B + \mu$ (or 2μ for equal electron-hole masses). This is attributed to "phase-space filling effects," or, in other words, the Burstein-Moss shift [26], which precisely correspond to the shift of the ladder pole, without the Fumi

contribution. The reason for this discrepancy is that our model ignores the CB electron-CB electron interaction, which would render the exciton electrically neutral and suppress the Fumi shift. Thus, as also pointed out in the literature on the x-ray edge problem, neglecting electron-electron interactions gives the right power law scalings of the spectra only, but not the correct threshold energies.

Another aspect of Eq. (35) is its lack of dependence on the frequency ω. In other words, the Anderson orthogonality power law of the exciton Green's function does not depend on $X(\omega)$. This could have been anticipated by an argument based on Hopfield's rule of thumb [41] and the results of [11]. Consider the spinful case, and study the absorption spectral function for, e.g., right-hand circularly polarized light at the exciton threshold, creating a spin down electron and a spin up hole. The spectrum should have the form

$$\frac{1}{\omega} \cdot \omega^{(1-\delta_\downarrow/\pi)^2+(1-\delta_\uparrow/\pi)^2}. \tag{D2}$$

For the spin down electrons, the exponent is $(1 - \delta_\downarrow/\pi)^2$ rather than $(\delta_\downarrow/\pi)^2$ because of the Hopfield rule: one electron is lifted from the valence band to the conduction band. For the spin up electron, no electron is lifted. However, the exciton is the secondary threshold in the spinful case (the primary one is the trion). As seen from Ref. [11], the spin-up exponent should therefore also be as in Eq. (D2). Now, in the spinful case, all direct diagrams will come with a spin factor of 2, while the exchange diagrams will not. However, we see that the exponent in (D2) is exactly 2 times the exponent the spinless case, Eq. (28), when recalling that $\delta_\uparrow = \delta_\downarrow = \delta$ for our spin-independent potential. This shows that the exchange diagrams should indeed not contribute to Anderson orthogonality, at least to leading order.

APPENDIX E: COMPUTATION OF PHASE-SPACE INTEGRALS FOR THE PARTICLE-HOLE PAIR DENSITY OF STATES

To clarify the different role of the recoil in the exciton (Sec. V B) and FES cases (Sec. VI B), let us present the computation of two important phase space integrals.

1. Exciton recoil

We start with the evaluation of the imaginary part of the exciton self-energy $\text{Im}[\Sigma](\omega)$ given in Eq. (52), focusing on zero exciton momentum. $\text{Im}[\Sigma]$ reads

$$\text{Im}[\Sigma_{\text{exc}}] \simeq -\frac{\pi V_0}{\rho g} \alpha^2 \int_{k_1<k_F} \frac{d\mathbf{k}_1}{(2\pi)^2} \int_{k_2>k_F} \frac{d\mathbf{k}_2}{(2\pi)^2}$$
$$\times \delta\big(\omega - (\mathbf{k}_2 - \mathbf{k}_1)^2/2M_{\text{exc}} - \epsilon_{\mathbf{k}_2} + \epsilon_{\mathbf{k}_1}\big). \tag{E1}$$

$\text{Im}[\Sigma_{\text{exc}}]$ can be interpreted as rate of decay of excitons into CB electron-hole pairs, or alternatively as density of state of the CB pairs. We aim to compute the leading ω-behaviour of $\text{Im}[\Sigma_{\text{exc}}]$. To put it short, the delta-function in (E1) requires $\mathbf{k}_1, \mathbf{k}_2 \simeq k_F$ and $\angle(\mathbf{k}_1, \mathbf{k}_2) \simeq 0$, and these phase space restrictions pile up to give $\text{Im}[\Sigma_{\text{exc}}] \sim \omega^{3/2}$. To perform the calculation in detail, we substitute $\mathbf{x} = \frac{\mathbf{k}_2}{\sqrt{2m}}$, $\mathbf{y} = \frac{\mathbf{k}_1}{\sqrt{2m}}$. Switching the integrals for convenience, we can rewrite (E1), to leading order in the mass ratio β, as

$$\text{Im}[\Sigma_{\text{exc}}] = -\frac{\alpha^2}{\pi} \int_{x>\sqrt{\mu}} d\mathbf{x} \int_{y<\sqrt{\mu}} d\mathbf{y} \delta(\omega - (x^2 - \mu)$$
$$+ (y^2 - \mu) - \beta(\mathbf{x} - \mathbf{y})^2). \tag{E2}$$

First, it is obvious that (E2) is proportional to $\theta(\omega)$, since all terms subtracted from ω in the delta function are positive, hence there cannot be any cancellations. Second, it is clearly seen that $x \simeq \sqrt{\mu}$, $y \simeq \sqrt{\mu}$ to yield a nonzero contribution for small ω. Thus we may linearize the dispersion relation, starting with $\mathbf{y}$:

$$\mathbf{y} = (\sqrt{\mu} + \gamma_y)\mathbf{e}_y, \tag{E3}$$
$$y^2 = \mu + 2\sqrt{\mu}\gamma_y + \mathcal{O}(\gamma_y^2). \tag{E4}$$

In doing so, we effectively disregard subleading terms of order $\mathcal{O}(\omega^2/\mu)$ in the argument of the delta function.

Introducing the notation

$$\phi = \angle(\mathbf{x}, \mathbf{y}), \quad c = \cos(\phi), \tag{E5}$$

we arrive at

$$\text{Im}[\Sigma_{\text{exc}}] = -\frac{\alpha^2 \theta(\omega)}{\pi} \int_{x>\sqrt{\mu}} d\mathbf{x} \int_{-1}^{1} \frac{2}{\sqrt{1-c^2}} \int_{-\sqrt{\mu}}^{0} d\gamma_y (\sqrt{\mu} + \gamma_y)$$

$$\times \delta(\underbrace{\omega - (x^2 - \mu) - \beta x^2 + 2\beta x\sqrt{\mu}c - \beta\mu}_{=A} + \gamma_y \underbrace{(2\beta xc - 2\beta\sqrt{\mu} + 2\sqrt{\mu})}_{=B}). \tag{E6}$$

Since the only contribution comes from γ_y close to the upper boundary, we can write $\sqrt{\mu} + \gamma_y \simeq \sqrt{\mu}$. Using $B \simeq 2\sqrt{\mu}$, the trivial integral over γ_y then results in

$$\text{Im}[\Sigma_{\text{exc}}] = -\frac{\alpha^2}{\pi} \int_{x>\sqrt{\mu}} d\mathbf{x} \int_{-1}^{1} dc \frac{1}{\sqrt{1-c^2}} \theta(A). \tag{E7}$$

To find the leading power law in ω of this expression, we assume that $\omega \ll \beta\mu$. Then, we rewrite $\theta(A)$ as

$$\theta(\overbrace{\omega - (x^2 - \mu) - \beta x^2 - \beta\mu}^{=C} + 2\beta x\sqrt{\mu}c) = \theta(c - (-C/2\beta x\sqrt{\mu})). \tag{E8}$$

We now use $x \simeq \sqrt{\mu}$. Thus we can write

$$-C/2\beta x\sqrt{\mu} \simeq 1 - \left(\frac{\omega}{2\beta\mu} - \frac{x^2 - \mu}{2\beta\mu}\right) + \mathcal{O}(\omega/\mu). \quad \text{(E9)}$$

Going back to (E7) gives

$$\text{Im}[\Sigma_{\text{exc}}] = -\frac{\alpha^2 \theta(\omega)}{\pi} \int_{x > \sqrt{\mu}} d\mathbf{x}\, \theta(\omega - (x^2 - \mu))$$

$$\times \int_{1 - (\omega - (x^2 - \mu))/2\beta\mu}^{1} dc\, \frac{1}{\sqrt{1 - c^2}}. \quad \text{(E10)}$$

Using that for $0 < t < 1$,

$$\int_{1-t}^{1} \frac{1}{\sqrt{1 - y^2}} dy = \arccos(1 - t) = \sqrt{2t} + \mathcal{O}(t^{3/2}), \quad \text{(E11)}$$

we obtain

$$\text{Im}[\Sigma_{\text{exc}}] = -2\alpha^2 \theta(\omega) \int_{\sqrt{\mu}}^{\sqrt{\mu + \omega}} x\, dx \sqrt{\frac{\omega - (x^2 - \mu)}{\beta\mu}}. \quad \text{(E12)}$$

This can be integrated exactly to give

$$\text{Im}[\Sigma_{\text{exc}}](\omega) = -\frac{2\alpha^2}{3} \frac{1}{\sqrt{\beta\mu}} \theta(\omega)\omega^{3/2}. \quad \text{(E13)}$$

The numerical prefactor should be correct, but is of no parametric relevance and is set to unity for convenience, thereby giving formula (53) of the main text.

2. FES regime: VB hole recoil

In the regime of the FES, not the exciton, but the valence-band hole recoils. Near the direct threshold at $\omega = \beta\mu$, the quantity describing the hole decay is $\text{Im}[\Sigma_{\text{VB}}(k_F, \omega)]$ as given in (66), which scales differently compared to the exciton decay because the VB hole has $\mathbf{Q} = k_F$ unlike the $\mathbf{Q} = 0$ exciton (we do not present this computation here since the power law is of
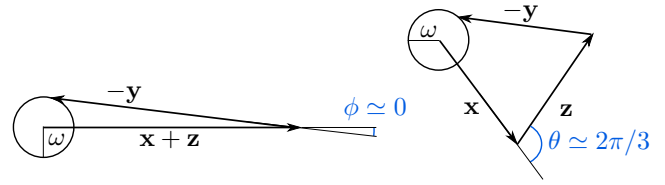

FIG. 28. Angles contributing to the indirect threshold. The ω circles indicate smallness in ω but not the exact power law or prefactor.

not much relevance for the 2DEG absorption we are interested in; see Ref. [40] for details).

Near the indirect threshold, the VB hole again has momentum $\mathbf{Q} = 0$, and the resulting 2DEG absorption $A(\omega)$ as given in (65) scales as $\sim \omega^3$. This result was already presented in [21], though without derivation. Since the computation is very similar to the previous one for the exciton decay, let us just sketch it: by performing frequency integrals in Figs. 15 and 16, and momentum substitutions as for the exciton, one arrives at

$$A(\omega) \sim \int_{x^2 > \mu} d\mathbf{x} \int_{z^2 > \mu} d\mathbf{z} \int_{y^2 < \mu} d\mathbf{y} \delta(\omega - (x^2 - \mu)$$

$$+ (y^2 - \mu) - (z^2 - \mu) - \beta(\mathbf{x} + \mathbf{z} - \mathbf{y})^2), \quad \text{(E14)}$$

which is similar to the previous expression (E2) except for an additional scattering partner, the photoexcited electron (corresponding to the $\mathbf{z}$ integral). Again, there can be no cancellations in the deltafunction, and the computation proceeds analogously to Sec. E 1. Effectively, the summands $(x^2 - \mu), (y^2 - \mu)$, and $(z^2 - \mu)$ contribute a factor of ω to $A(\omega)$. One factor is fixed by the delta function, such that in total one has ω^2. In addition, there is the hole recoil term $\beta(\mathbf{x} + \mathbf{z} - \mathbf{y})^2$. For this to be of order ω, the angles $\phi = \angle(\mathbf{x} + \mathbf{z}, \mathbf{y})$ and $\theta = \angle(\mathbf{x}, \mathbf{z})$ have to be fixed as depicted in Fig. 28. The explicit computation shows that each angle restriction give a factor of $\sqrt{\omega}$, such that in total one arrives at $A(\omega) \sim \omega^3$.

[1] I. Carusotto and C. Ciuti, Rev. Mod. Phys. **85**, 299 (2013).

[2] T. Byrnes, N. Y. Kim, and Y. Yamamoto, Nat. Phys. **10**, 803 (2014).

[3] C. Weisbuch, M. Nishioka, A. Ishikawa, and Y. Arakawa, Phys. Rev. Lett. **69**, 3314 (1992).

[4] J. Kasprzak, M. Richard, S. Kundermann, A. Baas, P. Jeambrun, J. M. J. Keeling, F. M. Marchetti, M. H. Szymanska, R. Andre, J. L. Staehli, V. Savona, P. B. Littlewood, B. Deveaud, and L. S. Dang, Nature (London) **443**, 409 (2006).

[5] F. P. Laussy, A. V. Kavokin, and I. A. Shelykh, Phys. Rev. Lett. **104**, 106402 (2010).

[6] P. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).

[7] G. D. Mahan, Phys. Rev. **163**, 612 (1967).

[8] B. Roulet, J. Gavoret, and P. Nozières, Phys. Rev. **178**, 1072 (1969).

[9] P. Nozières, J. Gavoret, and B. Roulet, Phys. Rev. **178**, 1084 (1969).

[10] P. Nozières and C. de Dominicis, Phys. Rev. **178**, 1097 (1969).

[11] M. Combescot and P. Nozières, J. Phys. **32**, 913 (1971).

[12] A. Gabbay, Y. Preezant, E. Cohen, B. M. Ashkinadze, and L. N. Pfeiffer, Phys. Rev. Lett. **99**, 157402 (2007).

[13] S. Smolka, W. Wuester, F. Haupt, S. Faelt, W. Wegscheider, and A. Imamoglu, Science **346**, 332 (2014).

[14] M. Sidler, P. Back, O. Cotlet, A. Srivastava, T. Fink, M. Kroner, E. Demler, and A. Imamoglu, Nat. Phys. **13**, 255 (2017).

[15] N. S. Averkiev and M. M. Glazov, Phys. Rev. B **76**, 045320 (2007).

[16] M. Baeten and M. Wouters, Phys. Rev. B **89**, 245301 (2014).

[17] M. Baeten and M. Wouters, Eur. Phys. J. D **69**, 1 (2015).

[18] J. Gavoret, P. Nozieres, B. Roulet, and M. Combescot, J. Phys. **30**, 987 (1969).

[19] P. Hawrylak, Phys. Rev. B **44**, 3821 (1991).

[20] T. Uenoyama and L. J. Sham, Phys. Rev. Lett. **65**, 1048 (1990).

[21] A. E. Ruckenstein and S. Schmitt-Rink, Phys. Rev. B **35**, 7551 (1987).

[22] P. Nozières, J. Phys. I **4**, 1275 (1994).

[23] A. Rosch and T. Kopp, Phys. Rev. Lett. **75**, 1988 (1995).

[24] H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (World Scientific, Singapore, 2009), Vol. 5.

[25] S. K. Adhikari, Am. J. Phys. **54**, 362 (1986).

[26] E. Burstein, Phys. Rev. **93**, 632 (1954).

[27] T. Moss, Proc. Phys. Soc. London, Sec. B **67**, 775 (1954).

[28] Y. Yamamoto and A. Imamoglu, *Mesoscopic Quantum Optics* (Wiley, New York, 1999), Vol. 1.

[29] G. D. Mahan, Phys. Rev. **153**, 882 (1967).

[30] O. Betbeder-Matibet and M. Combescot, Eur. Phys. J. B **22**, 17 (2001).

[31] M. Combescot, O. Betbeder-Matibet, and B. Roulet, Europhys. Lett. **57**, 717 (2002).

[32] M. Combescot and O. Betbeder-Matibet, Eur. Phys. J. B **31**, 305 (2003).

[33] M. Combescot, O. Betbeder-Matibet, and F. Dubin, Phys. Rep. **463**, 215 (2008).

[34] M. Combescot and O. Betbeder-Matibet, Eur. Phys. J. B **79**, 401 (2011).

[35] R. Suris, V. Kochereshko, G. Astakhov, D. Yakovlev, W. Ossau, J. Nürnberger, W. Faschinger, G. Landwehr, T. Wojtowicz, G. Karczewski, and J. Kossut, Phys. Status Solidi B **227**, 343 (2001).

[36] M. Baeten and M. Wouters, Phys. Rev. B **91**, 115313 (2015).

[37] D. K. Efimkin and A. H. MacDonald, Phys. Rev. B **95**, 035417 (2017).

[38] R. Wong, *Asymptotic Approximation of Integrals* (Academic Press, Inc., New York, 1989).

[39] G. D. Mahan, *Many Particle Physics*, 3rd ed. (Kluwer Academic/Plenum, New York and London, 2000).

[40] D. Pimenov, Master's thesis, Ludwig Maximilians University Munich, 2015.

[41] J. J. Hopfield, Comments Solid State Phys. **2**, 40 (1969).

5.3 Spectra of heavy polarons and molecules

5.3.1 Overview

Another fruitful experimental platform to study the impurity problem besides semiconductors are ultracold gases. Ever since the pioneering paper by Schirotzek et al. [SWSZ09], an important goal in this community has been to thoroughly understand the evolution from molecule to polaron, as laid out in Sec. 5.1.1. Remarkable agreement has been achieved between experiment and theory, with the latter evolving mostly along two lines:

On the analytical side, many authors employed variational ansaetze to estimate the ground state energy, in the footsteps of Chevy [Che06]. Both in 2D and 3D, two complementary ansaetze can be formulated: In the "polaronic" limit ($\mu \gg E_{\rm b}$ or $|k_F a| \gg 1$ in 3D), a fairly accurate result is already obtained from a variational wave-function that contains the impurity plus a single particle-hole excitation of the majority Fermi sea. Ansaetze with up to two particle-hole pairs have been considered [CG08]. In the opposite, molecular limit ($E_{\rm b} \gg \mu$ or $k_F a \ll 1$), the basic building block of the variational calculations is a rigid molecule, with again up to two particle-hole excitations on top [PDZ09, CGL10]. If the external momentum is set to zero, these ansaetze are orthogonal, and the obtained energies cross at a finite slope when extrapolated to intermediate values of $\mu/E_{\rm b}$ (although one has to zoom in considerably to resolve the difference in energy). This crossing is usually taken as a sign for a first order quantum phase transition. For a non-zero external momentum, the ansaetze are no longer mutually exclusive [Edw13], leaving room for a continuous transition (see also Sec. 5.4.2). Note that the expansion in electron-hole pairs is somewhat similar to our diagrammatic resummation described around Fig. 5.4; the differences are explained further below.

While a priori there is no real reason to trust the variational computations except for the extremal limits, the results are backed up by comparison to numerics, that is, mostly diagrammatic Monte-Carlo. Prokof'ev and Svistunov were the first to realize that the poles of the impurity $\propto \langle dd^\dagger \rangle$ and two-particle propagator $\propto \langle dcc^\dagger d^\dagger \rangle$ can be extracted from Monte-Carlo without the need for the ill-conditioned analytical continuation [PS08a, PS08b]. Furthermore, the sign-problem is no serious threat in the impurity context, as the alternating signs of the Feynman diagrams actually improve the convergence properties, allowing one to reach a fairly high expansion order.[4] The pole energy of the single-particle propagator is in nice agreement with the polaronic ansatz, while the pole of the two-particle propagator fits to the molecular ansatz. The agreement with the experimentally determined ground state energy is also decent, although the experimental data is not quite sufficient to distinguish between the polaronic and molecular ansaetze.
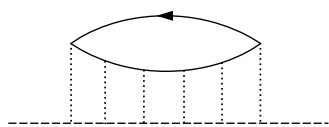

Figure 5.6 Ladder diagram for $\langle dd^\dagger \rangle$.

Thus, the question of the ground state energies is mostly settled. What is more difficult to predict are the spectral lineshapes measured in experiment, since this also requires the knowledge of the excited states. Deriving spectra from Monte-Carlo is challenging due to the necessary analytical continuation [GMPS16], and the variational ansaetze only yield ground state energies by design. One exception exists: the simple Chevy (polaronic) ansatz is known to be equivalent to the T-matrix approach/ladder diagram summation (see publication below for references), from which the impurity spectrum can easily be extracted.

A representative diagram of the ladder series for the single-particle propagator, with a single majority hole indicated by the arrow, is shown in Fig. 5.6. The spectrum obtained by

---

4 Although the series are still divergent and need to be resummed with Borel-type methods.

using these diagrams as impurity self-energy features a prominent delta-peak at the lowest energies, the "attractive polaron".

While the validity of the T-matrix picture is undisputed in the polaronic limit $\mu/E_b \gg 1$ (due to the equivalence to the polaronic ansatz), one may question it in the opposite limit $E_b \gg \mu$. In particular, the attractive polaron should aquire a finite line-width due to the decay into the low-lying molecular state. So, how to find the spectrum in this limit? As in the previous paper, our approach in the publication below is to start from the limit of infinitely heavy impurities, and then study the effects of finite mass. This is of particular interest for ultracold experiments with ^{40}K and ^6Li mixtures [KZJ$^+$12, CJL$^+$16] that do feature heavy impurities, although we believe that qualitatively our results extend to the equal mass case as well, and also apply to the (equal mass) TMD experiment [SBC$^+$17]. Let us summarize our analysis:

○ As discussed before, in the infinite mass limit the single-impurity spectrum can be found exactly [CN71, SKI$^+$18]; for $E_b \gg \mu$, the lowest energy feature is a weakly divergent power law previously shown in Fig. 5.3(a). This "molecule-hole continuum" reflects the formation of a bound state from the impurity and a majority electron, with a majority hole left behind. Thus, the molecule leaves a marked trace in the single-particle, and not just in the two-particle spectrum. The naive expectation would be that the attractive polaron splits off as a delta peak from the molecule-hole continuum for finite masses.

○ However, this turns out not to be the case in accordance with our discussion in Sec. 5.1.2. In fact, the finite mass cuts off the singularity completely, resulting in a broad incoherent spectral feature. This is somewhat unexpected given the T-matrix- and Monte-Carlo results [GMPS16], but in fact qualitatively agrees with the ultra cold gas experiments, where the attractive feature in the measured single-particle spectrum gradually broadens for increasing E_b/μ.

○ From the infinite mass point of view, one can schematically characterize the T-matrix approach for $E_b \gg \mu$ as follows: by using the single-hole ladder series of Fig. 5.6 as an impurity self-energy, one effectively resums a single logarithmic channel. This is similar to, say, trying to reduce the parquet series of Fig. 5.2 to the parallel bubbles only, by solving a single-channel Bethe-Salpeter equation. In this case, as the partial cancellation from the competing logarithmic channel is lost, the weakly divergent Fermi-edge singularity is replaced by a spurious delta-peak. In the same fashion, in the infinite mass limit the attractive polaron results from an uncontrolled approximation. Upon a proper summation of the leading logarithms, achieved by expanding the single-particle propagator in the number of holes without the appeal to Dyson's equation, the attractive polaron disappears. For finite mass (nonzero mass ratios β), this statement should remain true as long as $|\log(\beta)|$ is still large.

○ In addition to the attractive feature, our analysis also reveals a repulsive polaron at positive energies in accordance with prior literature. This excited state is of the strong singularity type for $E_b \gg \mu$ (Fig. 5.3(b)), i.e. it becomes even sharper when finite masses are considered; It also develops a delta-peak, which may additionally be broadened by decay into the molecule-hole continuum.

Spectra of heavy polarons and molecules coupled to a Fermi sea

by

D. Pimenov,[1] and M. Goldstein[2]

[1] Physics Department, Arnold Sommerfeld Center for Theoretical Physics, Center for NanoScience, and Munich Center for Quantum Science and Technology (MCQST), Ludwig-Maximilians-Universität München, 80333 München, Germany

[2]Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel

Spectra of heavy polarons and molecules coupled to a Fermi sea

Dimitri Pimenov[1,*] and Moshe Goldstein[2]

[1]*Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-University Munich, 80333 Munich, Germany*
[2]*Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel*

We study the spectrum of an impurity coupled to a Fermi sea (e.g., minority atom in an ultracold gas, exciton in a solid) by attraction strong enough to form a molecule/trion. We introduce a diagrammatic scheme which allows treating a finite mass impurity while reproducing the Fermi-edge singularity in the immobile limit. For large binding energies the spectrum is characterized by a semicoherent repulsive polaron and an incoherent molecule-hole continuum, which is the lowest-energy feature in the single-particle spectrum. The previously predicted attractive polaron seems not to exist for strong binding.

Introduction. The interaction of a single impurity with a surrounding fermionic bath is a problem at the very heart of quantum many-body physics, which is easily formulated, and yet difficult to solve. It is characterized by a rich interplay of kinetic and interaction effects, which can strongly modify the quasiparticle (polaronic) nature of the impurity. Controlled experimental realization and analysis of impurity physics has recently been achieved in ultracold gas setups [1–5], where the impurity is usually an excited hyperfine state of an atom, and the interaction strength is tunable via Feshbach resonances [6]. An alternative are semiconductor or transition metal dichalcogenide (TMD) experiments [7,8], where the impurity is a valence band hole or exciton, in the presence of a finite conduction band population controlled by gate voltage.

On the theory side, a major part of the literature is devoted to the computation of ground-state energies following Chevy's [9] pioneering work, which proposed an ansatz for the ground-state wave function consisting of the impurity dressed by a single electron-hole pair. This ansatz works well in the polaronic regime where the impurity-bath interaction is weak, but breaks down if the formation of a molecule, or trion in semiconductor language, becomes favorable. This regime can be described by a complementary ansatz [10–12] involving a dressed molecule. In two dimensions (2D), a similar picture applies [13–15].

The variational energy has recently been verified using diagrammatic quantum Monte Carlo [16–21]. The situation is quite different for the *impurity spectrum*, which is the actual quantity measured in experiments: In Monte Carlo, extracting the spectrum is difficult due to the infamous analytical continuation problem, and only few definite statements can be made [22]. Analytically, it has been realized that the Chevy ansatz is equivalent to the non-self-consistent T-matrix approach [23], from which spectra can be easily extracted [1,5,8,24–26]. However, this ansatz is *a priori* reliable for weak coupling only. In the molecule limit, extracting the spectrum from a variational ansatz is difficult since the coefficients are not

analytically known. As for the functional renormalization group [27], its accuracy is hard to assess [28].

Besides the interaction strength and Fermi energy, a third control parameter in the impurity problem is the impurity mass M. Infinitely heavy impurities are subject to Anderson orthogonality [29], and the universal properties of the impurity spectrum in the presence of a bound state can be computed exactly from a functional determinant [30–33]. The goal of this Rapid Communication is to characterize the spectrum for arbitrary impurity mass, while maintaining consistency with all known limits. Building on the framework developed in our recent work [34], we find that a rigorous expansion in the number of fermion-hole pairs reproduces the infinite mass spectrum, and obtain controlled estimates of the impurity spectrum deep in the molecular limit; in particular, we present a controlled computation of the incoherent molecular feature in the single-particle spectrum. We mostly focus on 2D for clarity, listing the modifications in 3D along the way.

Model. Consider a single impurity (annihilation operator d) immersed in a bath of fermions (c). In a cold atom system, the impurity can be a spin-up fermion in a bath of spin-down particles; in semiconducting systems, the impurity is usually an exciton containing a conduction electron with a given spin, together with a bath of the opposite spin conduction electrons [35]. The usual model Hamiltonian reads

$$H = \sum_{\mathbf{k}}(\epsilon_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}} + E_{\mathbf{k}} d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}}) - \frac{V_0}{\mathcal{S}} \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}-\mathbf{q}} d_{\mathbf{p}}^{\dagger} d_{\mathbf{p}+\mathbf{q}}, \quad (1)$$

with $\epsilon_{\mathbf{k}} = k^2/2m$, $E_{\mathbf{k}} = k^2/2M$. $V_0 > 0$ is the attractive contact interaction [36], $\mathcal{S}$ the system area, and $\hbar = 1$. Our goal is to find the single-particle spectrum $\mathcal{A}(\omega)$ at zero momentum, which is proportional to the Fourier transform of the imaginary part of the retarded impurity Green's function, $D(t) = -i\theta(t)\langle 0|d_0(t)d_0^{\dagger}(0)|0\rangle$, where $|0\rangle$ is the Fermi sea without impurity. We work in the real frequency formalism at zero temperature.

Chevy's ansatz versus the Fermi-edge singularity. Chevy's ansatz corresponds to a summation of all impurity self-energy diagrams Σ_1 with a single hole (the T-matrix series), shown
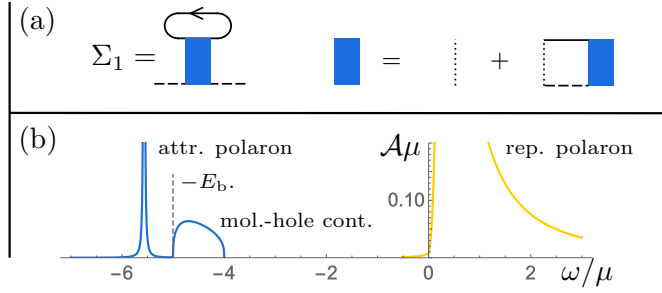
*D.Pimenov@physik.lmu.de

FIG. 1. (a) Self-energy diagrams with one hole, indicated by the arrow. Solid (dashed) lines denote electron (impurity) propagators. The blue box indicates the T matrix. (b) ($M = \infty$) spectrum for $E_b = 5\mu$ from the Chevy ansatz. Repulsive and attractive features in the spectrum are shown in different colors for clarity. For the attractive polaron, a finite width is used.
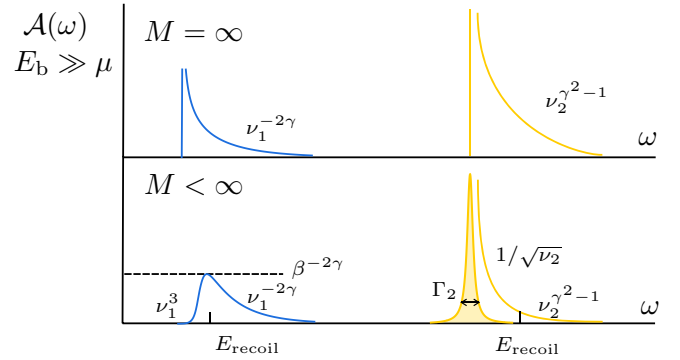

FIG. 2. Sketch of the spectrum for $E_b \gg \mu$. Power laws are measured from the respective thresholds; $\Gamma_2 \simeq \mu^4 / E_b^4 \beta$ is the width of the repulsive polaron (see main text). Colors are chosen as in Fig. 1.

in Fig. 1(a). For infinite mass, one finds, in 2D,

$$\Sigma_1(\omega) = -\int_0^\mu d\epsilon_{\mathbf{k}} \frac{1}{\ln\left(\frac{\omega + \epsilon_{\mathbf{k}} - \mu + i0^+}{-E_b}\right)}. \quad (2)$$

Here, ω is the energy measured from the impurity level, and $\mu = k_F^2 / 2m$ is the Fermi energy. We define the complex logarithm with a branch cut on the negative half axis. $-E_b$ is the energy of the bound state of the attractive contact potential, which always exists in 2D. It is determined from the pole of the T matrix. Due to this bound state, $\mathrm{Im}[\Sigma_1](\omega)$ has a molecule continuum $\propto \theta(\omega + E_b)$. Its width is μ, representing the different energies of the hole in the Fermi sea created when the impurity binds an electron. For $E_b \gg \mu$, inserting Σ_1 into the bare impurity Green's function $D_0(\omega) = 1/(\omega + i0^+)$ leads to three prominent features: First, the bare pole of the impurity is shifted ("repulsive polaron"). Second, the aforementioned molecule-hole continuum is created. Third, $\mathrm{Re}[\Sigma_1]$ gives rise to another pole below the molecule-hole continuum, the "attractive polaron." Between the latter two there is a spectral gap of $\simeq -0.582\mu$ as $E_b/\mu \to \infty$ [13]. A typical plot is shown in Fig. 1(b). The 3D result is similar (see Supplemental Material [37]). For finite mass, expression (2) is more complicated, but the qualitative form of the spectrum is unchanged [25].

The "Chevy" spectrum for $M = \infty$ is to be contrasted with the exact result of Combescot and Nozières [30], who showed that the spectrum is dominated by two divergent power laws [38] $\mathcal{A}(\omega) \propto \sum_{i=1,2} (\omega - \omega_{\mathrm{th},i})^{\alpha_i} \theta(\omega - \omega_{\mathrm{th},i})$. Here, $\omega_{\mathrm{th},i}$ are the threshold energies determined from Fumi's theorem [39], and the exponents α_i are characterized by δ, the phase shift of the bath fermions at the Fermi energy due to their scattering by the immobile impurity, $\alpha_1 = (\delta/\pi)^2 - 1$, $\alpha_2 = (1 - \delta/\pi)^2 - 1$. For infinite mass, the dimensionality of the problem only affects the value of δ. For $E_b \gg \mu$ one can then approximate [40,41]

$$1 \gg 1 - \delta/\pi \simeq \gamma \equiv \begin{cases} 1/\ln(E_b/\mu) & \text{for } d = 2, \\ k_F a/\pi & \text{for } d = 3, \end{cases} \quad (3)$$

with the 3D scattering length a. In this limit, with exponents to leading order in γ, the spectrum looks like

$$\mathcal{A}(\omega) \simeq \theta(\nu_1) \nu_1^{-2\gamma} + \theta(\nu_2) \nu_2^{\gamma^2 - 1}, \quad \nu_i \equiv \omega - \omega_{\mathrm{th},i}. \quad (4)$$

A sketch is shown in Fig. 2 (upper panel). The lower (blue) feature, which starts close to $\omega = -E_b$ and corresponds to the molecule-hole continuum, has a weak power law (close to a step). The upper feature, which can be identified with the repulsive polaron, has a strong power-law spectrum (close to a delta function). Note that there is no well-defined "attractive polaron" in the spectrum. We claim that, for $E_b \gg \mu$, this will persist for finite masses M, and thus the Chevy spectrum of Fig. 1(b) is incorrect *for large binding energies*.

Method. Our approach is to reproduce Eq. (4) in a diagrammatic expansion in γ generalizable to finite mass. However, γ does not directly appear in the Hamiltonian; instead, one must resort to an expansion in the *number of holes*: A diagram involving n holes contains n integrations over filled states $\propto \mu^n$, and μ is small in units of E_b. In effect, as shown below, this leads to an expansion in γ [12,14,15,23,42].

The one-hole diagrams are already considered as the impurity self-energy within the Chevy approach [Fig. 1(a)], and resummed with Dyson's equation. For heavy impurities, this resummation is uncontrolled. Instead, one must add up the most important (log-divergent) diagrams order by order in γ, which ultimately removes the attractive polaron from the spectrum. Thus, we reattach the impurity lines to Σ_1, defining $H_1(\omega) = D_0(\omega)^2 \Sigma_1(\omega)$. Of course, H_1 only represents the first-order process: The impurity can interact with an arbitrary number of electrons, creating electron-hole excitations in the Fermi sea. The processes involving two holes are represented in Fig. 3(a). Here, the interaction lines can be drawn arbitrarily often in any order, as long as the structure of the diagrams is preserved, e.g., in diagram H_2^a the first and last interaction lines should connect to the lower part of the "horseshoe," and to the upper loop in diagram H_2^c. These diagrams can also be redrawn with T-matrix blocks, as exemplarily shown in Fig. 3(b); we never expand in the number of T matrices, but always resum diagrams with an infinite number of T matrices at the two-hole level. We note that the contribution of the two-hole diagrams to the ground-state energy is much less significant [12,43].

Results: The molecule/attractive polaron spectrum. For dispersionless infinite mass impurities, the evaluation of all two-hole diagrams is possible. Following Ref. [34], one can either work in the time or frequency domain, employing
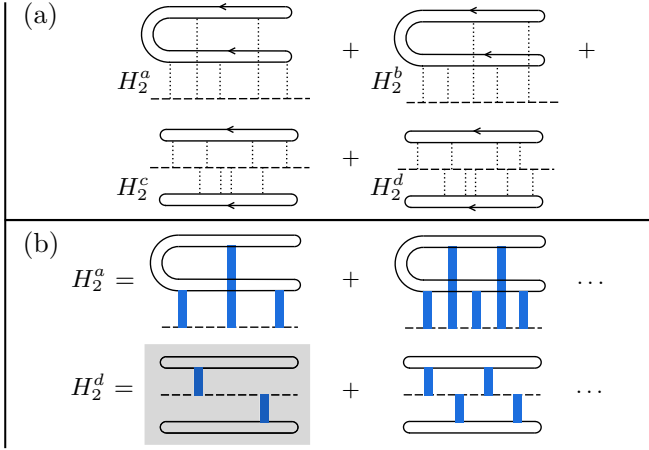
FIG. 3. (a) All relevant two-hole diagrams. (b) T-matrix representation of diagram series H_2^a, H_2^d. The gray-shaded diagram is contained in the Chevy approach.

different approximations [37]. In particular, for small $\omega + E_B$, close to the molecular threshold, we find [44]

$$
H_1(\omega) + H_2(\omega)
$$
$$
\simeq \frac{1}{E_b}\left[\ln\left(\frac{\omega + E_b + i0^+}{-\mu}\right) - \gamma \ln^2\left(\frac{\omega + E_b + i0^+}{-\mu}\right)\right],
\tag{5}
$$

where $H_2 = \sum_i H_2^i$. The term $\propto \ln^2$ in Eq. (5) arises solely from diagram H_2^a. Curiously, the contribution of H_2^d is subleading, while the contribution from diagrams H_2^b, H_2^c effectively shifts the bound-state energy as $E_b \to E_b + \mu(1 - \gamma)$ in 2D, or $E_b + \mu(1 - 2\gamma/3)$ in 3D, in agreement with Fumi's theorem [39] to leading order in γ. Redefining ν_1 to include these shifts, we find a contribution to the spectrum,

$$
\mathcal{A}_1(\nu_1) \simeq \frac{\theta(\nu_1)}{E_b}(1 - 2\gamma \ln[\nu_1/\mu]),
\tag{6}
$$

in agreement with Eq. (4) when expanded in γ. This expansion has the same form as the perturbative expansion of the polarization in the standard Fermi-edge singularity case [45–47]. This was to be expected, as in the limit $E_b \to \infty$ we can formally regard the diagrams $H_{1,2}$ as polarization diagrams containing a molecule and a bath fermion, with an effective molecule-bath interaction γ. We expect higher-order leading logarithmic (parquet) contributions to arise in a similar fashion from diagrams containing a larger number of holes.

Let us now address the modification of the molecule-hole feature for a large but finite impurity mass M. The general strategy is to reevaluate the frequency-domain diagrams of Fig. 3(b) for finite mass [37], and trace the modification of the logarithmic singularities [34,48–52]. Our results hold to leading order in the mass ratio $\beta = m/M$ only, but we expect them to be qualitatively correct all the way up to $\beta \simeq 1$. First, introducing a finite mass shifts the binding energy, $E_b \to \tilde{E}_b$, but we will not compute those shifts in detail, limiting ourselves to the form of the spectrum. In terms of $\nu_1 = \omega + \tilde{E}_b$, the real part of the logarithmic singularities is modified as $\ln(\max[\nu_1 - \beta\mu, \gamma^2\beta\mu]/\mu)$, again reminiscent of the

Fermi-edge singularity case [48]. In contrast to $M = \infty$, the logarithmic singularities for finite mass are peaked at $\nu_1 = \beta\mu$ ("direct threshold" [48]). This is simply understood: When an incoming zero momentum impurity binds an electron and leaves behind a low-energy hole, the resulting molecule must have a momentum $\simeq k_F$ by momentum conservation. Since the molecule is now mobile, with mass $M_+ = M + m$, one must pay its recoil energy $E_{\text{recoil}} \simeq \beta\mu$, which shifts the maximum of the logarithms to $\nu_1 = \beta\mu$. Subsequently, the so created molecule can decay into a zero momentum state, by exciting an electron-hole pair. The rate of this indirect process is $\Gamma_1 = \gamma^2\beta\mu$, leading to a cutoff of the logarithmic singularities. Mathematically, this cutoff arises from the diagram H_2^c, which can be interpreted as a molecule self-energy diagram with an imaginary part Γ_1. For large frequencies, $\nu_1 \gg E_{\text{recoil}}$, one recovers the infinite mass behavior $\propto \nu_1^{-2\gamma}$.

Apart from cutting off the singularity, the decay of the molecule leads to a shift of the threshold from the direct to the "indirect" one at $\nu_1 = 0$, which corresponds to the creation of zero momentum molecules. Near the indirect threshold, the spectrum starts continuously, with a power law $\propto \nu_1^3$ in 2D and $\propto \nu_1^{7/2}$ in 3D. This behavior is obtained by computing the imaginary parts of diagrams $H_2^{a,c}$, which yield the leading contributions in γ via standard phase space estimation [34,37]. For a spinless Fermi sea, the two contributions cancel; however, even in this case we expect that the power-law behavior is robust, since it is (a) determined from a generic phase space estimate and (b) there may well be processes involving three holes that yield the same behavior. Exponentiating the logarithms [34], one finds the spectrum near both thresholds to be

$$
\mathcal{A}_1(\nu_1) \simeq \frac{1}{E_b}\left(\frac{\sqrt{(\nu_1 - \beta\mu)^2 + (\gamma^2\beta\mu)^2}}{\mu}\right)^{-2\gamma}\theta(\nu_1)f_1(\nu_1),
\tag{7}
$$

where $f_1(\nu_1)$ smoothly interpolates between $f_1(\nu_1) \simeq \gamma^2(\nu_1/\beta\mu)^3$ in 2D and $f_1(\nu_1) \simeq \gamma^2(\nu_1/\beta\mu)^{7/2}$ in 3D, for $\nu_1 \ll \beta\mu$, and $f_1(\nu_1) \simeq \pi$ for $\nu_1 \gtrsim \beta\mu$. A typical plot of the resulting spectrum is shown in Fig. 2 (blue feature in the lower panel). Let us reiterate: The ground-state signal in the spectrum is purely incoherent, with maximum $\propto (\beta)^{-2\gamma}$ [53]; there is no polaronic delta peak.

Results: The repulsive polaron spectrum. We now discuss the repulsive polaron already predicted by the Chevy ansatz [5,14,27,54–56]. For $E_b \gg \mu$, the repulsive polaron contains most of the spectral weight, $\sim 1 - \mu/E_b$, as seen in Fig. 1(b): As $E_b/\mu \to \infty$, the repulsive polaron is essentially a spectral probe of the impurity without the Fermi sea, with unit weight. For infinite mass, the asymptotic form of the repulsive polaron is given by the second term in Eq. (4), with $\omega_{\text{th},2} \simeq \gamma\mu$ in 2D and $\omega_{\text{th},2} \simeq \frac{2}{3}\gamma\mu$ in 3D. To leading order in γ, $\mathcal{A}_2(\nu_2) \simeq \gamma^2\theta(\nu_2)/\nu_2$, which reduces to a delta function as $\gamma \to 0$. This leading-order term can already be obtained from the first-order diagram H_1 for small positive frequencies. One can also reproduce the full power-law singularity in a linked cluster approach, formally exponentiating H_1. Extending the latter approach to finite mass, one finds a delta peak with weight β^{γ^2}, on top of an incoherent background $\propto 1/\sqrt{\nu_2}$
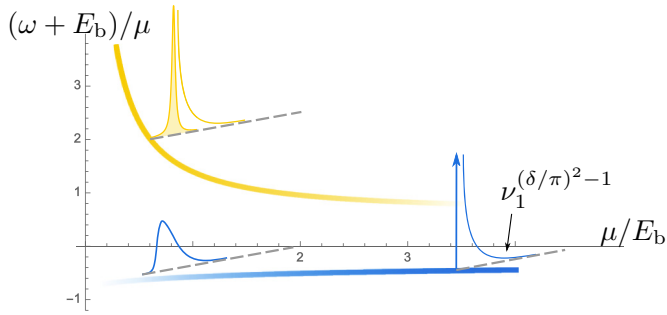
FIG. 4. Sketch of the full 2D spectrum for general values of μ/E_b. Thick lines indicate the threshold position determined from Fumi's theorem.

for $\nu_2 \ll \beta\mu$, similar to the results of Ref. [51]. In 3D, the incoherent part is approximately constant. For much larger frequencies $\nu_2 \gtrsim \beta\mu$, one recovers the infinite mass behavior $\propto \nu_2^{\gamma^2-1}$ [37].

Thus, in a first approximation, the repulsive polaron is a delta peak plus incoherent background. However, for finite mass, the delta peak may be broadened due to decay into the low-lying molecule-hole continuum, resulting in a finite width Γ_2. This width can be estimated by computing the self-energy part of the diagrams H_2 (called Σ_2) at the repulsive polaron threshold $\nu_2 = 0$. Note that, for infinite mass, the problem becomes single particle [33], forbidding such a transition; this behavior is reproduced by our calculations. Unfortunately, for finite mass a complete evaluation of Im$[\Sigma_2(\nu = 0)]$ is out of reach. A simple estimate can be obtained from a golden-rule-type expansion of Σ_2 in T matrices [37], similar to Ref. [57]; we find, in 2D, $\Gamma_2 \sim \gamma^2 \beta \frac{\mu^4}{E_b^3}$; in 3D, Γ_2 should still be small in μ/E_b, but the scaling could be different. Putting everything together, an approximate expression for the repulsive polaron spectrum reads

$$\mathcal{A}_2(\nu_2) \simeq (\beta)^{\gamma^2} \frac{\Gamma_2}{\nu_2^2 + \left(\frac{1}{2}\Gamma_2\right)^2} + f_2(\nu_2), \qquad (8)$$

where $f_2(\nu_2)$ interpolates between the limits $f_2(\nu_2) \simeq \gamma^2/\sqrt{\beta\mu\nu_2}$ in 2D and $f_2(\nu_2) \simeq \gamma^2/(\beta\mu)$ in 3D, for $\nu_2 \ll \beta\mu$, and $f_2 \simeq 1/\mu(\nu_2/\mu)^{\gamma^2-1}$ for $\nu_2 \gtrsim \beta\mu$. A sketch is shown in Fig. 2 (yellow feature in the lower panel).

Discussion. So far, we have only discussed the spectrum in the molecular limit $E_b \gg \mu$. In the opposite limit, the influence of the bound state should be negligible. The spectrum of a heavy impurity without a bound state was computed in Ref. [51], and we expect the same result here: a single feature of a form similar to the repulsive polaron described above, but with a delta peak that is not broadened, and singularity

exponents controlled by $\delta \ll 1$ for $\mu \gg E_b$. Both known limits (in 2D) are sketched in Fig. 4, along with the thresholds as determined from Fumi's theorem, which should be approximately correct for large masses. Note that if we follow the lower spectral feature, we see a "molecule-to-polaron transition," since, for $\mu \ll E_b$, the single-particle spectrum is fully incoherent, but fully coherent in the opposite limit. The details of this transition/crossover [58] remain to be explored. In particular, it would be interesting to analyze this in 3D, where a vacuum bound state only forms at $a > 0$.

Let us also comment on the connection to quantum Monte Carlo and experiments. A major difference is that the Monte Carlo works extract the molecule solely from a pole in the two-particle propagator. The latter was obtained in our recent work [34], and we found essentially opposite behavior to the one presented here, e.g., for $E_b \gg \mu$, there is a sharp feature related to the molecule, and a broad continuum at larger energies. However, here we have argued that the molecule emerges as an incoherent ground-state feature in the single-particle propagator as well. This seems to be in agreement with the ultracold gas experiments in both 3D [1,2,4,5] and 2D [3], while the results of the 2D TMD experiment are somewhat less clear [8]. The incoherent molecule feature was not seen in the "polaron spectra" of the recent Monte Carlo work [22], possibly due to problems with analytical continuation. Finally, let us note that most Monte Carlo works, in 3D [16–18,21,22] and 2D [19,20], deal with the (almost) equal mass case, while in the experiment also heavily mass-imbalanced ^{6}Li-^{40}K mixtures are used. Anyway, we do not expect significant changes in the spectra for equal masses, except for the disappearance of the orthogonality power laws beyond E_{recoil}.

Conclusion. We presented a controlled computation of polaron spectra, providing the connection to the infinite mass limit. We found that, for large binding, the attractive polaron and molecule-hole continuum merge into a single incoherent feature, and also gave a detailed description of the repulsive polaron spectrum. Our work paves the way towards the study of many impurity physics, including the effective interaction between impurities, molecular condensate versus polaron Fermi gas, etc. [59,60].

[1] A. Schirotzek, C.-H. Wu, A. Sommer, and M. W. Zwierlein, Phys. Rev. Lett. **102**, 230402 (2009).

[2] C. Kohstall, M. Zaccanti, M. Jag, A. Trenkwalder, P. Massignan, G. M. Bruun, F. Schreck, and R. Grimm, Nature (London) **485**, 615 (2012).

[3] M. Koschorreck, D. Pertot, E. Vogt, B. Fröhlich, M. Feld, and M. Köhl, Nature (London) **485**, 619 (2012).

[4] M. Cetina, M. Jag, R. S. Lous, I. Fritsche, J. T. Walraven, R. Grimm, J. Levinsen, M. M. Parish, R. Schmidt, M. Knap *et al.*, Science **354**, 96 (2016).

[5] F. Scazza, G. Valtolina, P. Massignan, A. Recati, A. Amico, A. Burchianti, C. Fort, M. Inguscio, M. Zaccanti, and G. Roati, Phys. Rev. Lett. **118**, 083602 (2017).

[6] W. Ketterle and M. W. Zwierlein, in *Ultracold Fermi Gases*, Proceedings of the International School of Physics "Enrico Fermi," Course CLXIV, Varenna, 2006, edited by M. Inguscio, W. Ketterle, and C. Salomon (IOS Press, Amsterdam, 2008).

[7] S. Smolka, W.Wuester, F. Haupt, S. Faelt, W. Wegscheider, and A. Imamoglu, Science **346**, 332 (2014).

[8] M. Sidler, P. Back, O. Cotlet, A. Srivastava, T. Fink, M. Kroner, E. Demler, and A. Imamoglu, Nat. Phys. **13**, 255 (2017).

[9] F. Chevy, Phys. Rev. A **74**, 063628 (2006).

[10] M. Punk, P. T. Dumitrescu, and W. Zwerger, Phys. Rev. A **80**, 053605 (2009).

[11] C. Mora and F. Chevy, Phys. Rev. A **80**, 033607 (2009).

[12] R. Combescot, S. Giraud, and X. Leyronas, Europhys. Lett. **88**, 60007 (2010).

[13] M. M. Parish, Phys. Rev. A **83**, 051603 (2011).

[14] M. M. Parish and J. Levinsen, Phys. Rev. A **87**, 033616 (2013).

[15] J. Levinsen and M. M. Parish, Annu. Rev. Cold At. Mol. **3**, 1 (2015).

[16] N. V. Prokof'ev and B. V. Svistunov, Phys. Rev. B **77**, 020408 (2008).

[17] N. V. Prokof'ev and B. V. Svistunov, Phys. Rev. B **77**, 125101 (2008).

[18] J. Vlietinck, J. Ryckebusch, and K. Van Houcke, Phys. Rev. B **87**, 115133 (2013).

[19] J. Vlietinck, J. Ryckebusch, and K. Van Houcke, Phys. Rev. B **89**, 085119 (2014).

[20] P. Kroiss and L. Pollet, Phys. Rev. B **90**, 104510 (2014).

[21] P. Kroiss and L. Pollet, Phys. Rev. B **91**, 144507 (2015).

[22] O. Goulko, A. S. Mishchenko, N. Prokof'ev, and B. Svistunov, Phys. Rev. A **94**, 051605 (2016).

[23] R. Combescot, A. Recati, C. Lobo, and F. Chevy, Phys. Rev. Lett. **98**, 180402 (2007).

[24] M. Punk and W. Zwerger, Phys. Rev. Lett. **99**, 170404 (2007).

[25] R. Schmidt, T. Enss, V. Pietilä, and E. Demler, Phys. Rev. A **85**, 021602 (2012).

[26] D. K. Efimkin and A. H. MacDonald, Phys. Rev. B **95**, 035417 (2017).

[27] R. Schmidt and T. Enss, Phys. Rev. A **83**, 063620 (2011).

[28] F. B. Kugler and J. von Delft, J. Phys.: Condens. Matter **30**, 195501 (2018).

[29] P. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).

[30] M. Combescot and P. Nozières, J. Phys. **32**, 913 (1971).

[31] M. Baeten and M. Wouters, Phys. Rev. B **91**, 115313 (2015).

[32] R. Schmidt, M. Knap, D. A. Ivanov, J.-S. You, M. Cetina, and E. Demler, Rep. Prog. Phys. **81**, 024401 (2018).

[33] P. Nozières and C. T. De Dominicis, Phys. Rev. **178**, 1097 (1969).

[34] D. Pimenov, J. von Delft, L. Glazman, and M. Goldstein, Phys. Rev. B **96**, 155310 (2017); D. Pimenov, Fermi-edge polaritons

[35] For an exciton to bind two same-spin electrons, the "trion" angular momentum must be odd by Pauli exclusion, and this is energetically unfavorable for heavy excitons [14,61].

[36] This also applies to electronic systems due to screening [34] .

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.98.220302 which includes Refs. [62–64], for further technical details.

[38] We neglect band bottom features discussed in Ref. [32] .

[39] G. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer Academic/Plenum, New York/London, 2000).

[40] S. K. Adhikari, Am. J. Phys. **54**, 362 (1986).

[41] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Phys. Rev. B **41**, 327 (1990).

[42] P. Bloom, Phys. Rev. B **12**, 125 (1975).

[43] R. Combescot and S. Giraud, Phys. Rev. Lett. **101**, 050404 (2008).

[44] The computation is restricted to a parametrically large window of not too small frequencies, $\mu^2/E_b \ll \omega + E_b \ll \mu$ [37] .

[45] G. D. Mahan, Phys. Rev. **153**, 882 (1967).

[46] B. Roulet, J. Gavoret, and P. Nozieres, Phys. Rev. **178**, 1072 (1969).

[47] P. Nozieres, J. Gavoret, and B. Roulet, Phys. Rev. **178**, 1084 (1969).

[48] J. Gavoret, P. Nozieres, B. Roulet, and M. Combescot, J. Phys. (Paris) **30**, 987 (1969).

[49] A. E. Ruckenstein and S. Schmitt-Rink, Phys. Rev. B **35**, 7551 (1987).

[50] P. Nozières, J. Phys. I France **4**, 1275 (1994).

[51] A. Rosch and T. Kopp, Phys. Rev. Lett. **75**, 1988 (1995).

[52] A. Rosch, Adv. Phys. **48**, 295 (1999).

[53] The coefficient in front of β cannot be determined exactly with logarithmic accuracy, but is between γ^2 and 1 [34,48].

[54] X. Cui and H. Zhai, Phys. Rev. A **81**, 041602 (2010).

[55] P. Massignan and G. Bruun, Eur. Phys. J. D **65**, 83 (2011).

[56] V. Ngampruetikorn, J. Levinsen, and M. M. Parish, Europhys. Lett. **98**, 30005 (2012).

[57] G. M. Bruun and P. Massignan, Phys. Rev. Lett. **105**, 020403 (2010).

[58] D. Edwards, J. Phys.: Condens. Matter **25**, 425602 (2013).

[59] L. Radzihovsky and D. E. Sheehy, Rep. Prog. Phys. **73**, 076501 (2010).

[60] H. Hu, B. C. Mulkerin, J. Wang, and X.-J. Liu, Phys. Rev. A **98**, 013626 (2018).

[61] C. J. M. Mathy, M. M. Parish, and D. A. Huse, Phys. Rev. Lett. **106**, 166404 (2011).

[62] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 2014).

[63] M. Punk, Many-particle physics with ultracold gases, Ph.D. thesis, Technical University of Munich, 2010.

[64] T. Kopp, A. E. Ruckenstein, and S. Schmitt-Rink, Phys. Rev. B **42**, 6850 (1990).

with finite hole-mass, Master's thesis, Ludwig Maximilians University of Munich, 2015.

Supplementary Material for "Heavy polarons and molecules coupled to a Fermi sea"

Dimitri Pimenov[1, *] and Moshe Goldstein[2]

[1] *Arnold Sommerfeld Center for Theoretical Physics,*
Ludwig-Maximilians-University Munich, 80333 Munich, Germany
[2] *Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel*

In this supplement, we present the detailed derivation of our results. In Sec. S.A we recapitulate Fumi's theorem and the determination of the phase shift δ. The evaluation of the one-hole diagram H_1 in the infinite mass limit is presented in Sec. S.B for 2D, and in Sec. S.C for 3D. Focusing on 2D, the two-hole diagrams H_2 are computed in the time domain in Sec. S.D and in the frequency domain in Sec. S.E. Then, in Sec. S.F, the molecule-hole continuum is determined for finite mass. Finally, the linked cluster approach is used to compute the repulsive polaron for infinite mass in Sec. S.G, and for finite mass in Sec. S.H.
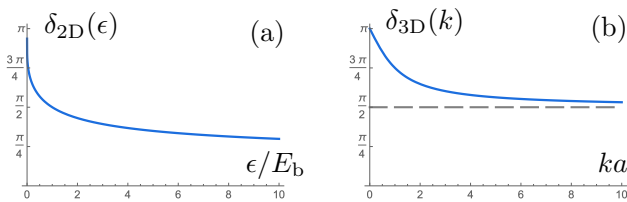

Figure S1. Scattering phase shifts (a) 2D phase shift (c) 3D phase shift.

S.A. DETERMINATION OF THE PHASE SHIFTS AND FUMI'S THEOREM

The universal properties of immobile impurities are characterized by the energy dependent scattering phase shift of the two-particle problem, $\delta(\epsilon)$. In particular, the power law exponents discussed in the main text are determined by the phase shift at the Fermi energy, $\delta(\mu)$. For a zero-range interaction potential in 2D, δ is given by [1, 2]

$$\delta_{2D}(\epsilon) = \cot^{-1}\left(\ln\left[\frac{\epsilon}{E_b}\right]/\pi\right), \quad \text{(S1)}$$

where E_b is the 2D binding energy (chosen as positive, i.e., the bound state occurs at $-E_b$). In 3D, provided the potential is strong enough to form a bound state, the phase shift is usually presented as [2]

$$\delta_{3D}(k) = \pi + \arctan(-ka), \quad k = \sqrt{2m\epsilon}, \quad \text{(S2)}$$

with the 3D scattering length a. Note that the reduced mass m_r equals m for immobile impurities. δ_{2D} and δ_{3D} are plotted in Fig. S1. In terms of δ, the lower threshold of the spectrum $\omega_{\text{th},1}$ is obtained from Fumi's theorem (see, e.g., [3]):

$$\omega_{\text{th},1} = -E_b - \int_0^\mu \frac{d\epsilon}{\pi}\, \delta(\epsilon), \quad \text{(S3)}$$

———
* D.Pimenov@physik.lmu.de

with $E_b = 1/(2ma^2)$ for $M = \infty$ in $d = 3$. The second threshold is reached by removing the Fermi sea electron from the bound state and putting it on top of the Fermi sea, thus $\omega_{\text{th},2} = \omega_{\text{th},1} + E_b + \mu$. The two thresholds are plotted in Fig. 4 of the main text.

S.B. $M = \infty$: EVALUATION OF H_1 IN 2D

The T-matrix corresponds to the sum of all "ladder"-diagrams for the two-particle vertex and is pictorially defined in Fig. 1(a) of the main text. It is solely a function of total energy-momentum $(\omega + \epsilon, \mathbf{k})$. In the evaluation of the T-matrix and all further diagrams, the summation over internal frequencies is trivial: the bare propagator of the single impurity is purely retarded (since there is no "impurity Fermi sea"), which effectively sets all internal frequencies on-shell and restricts the momenta of electrons propagating forward (backward) in time to be above (below) k_F. With finite impurity masses, in 2D the T-matrix is therefore given by:

$$T(\omega + \epsilon, \mathbf{k}) =$$
$$\left(-1/V_0 - \int d\mathbf{p} \frac{1}{\omega + \epsilon - \epsilon_{\mathbf{p}} - E_{\mathbf{k}-\mathbf{p}} + i0^+}\right)^{-1}. \quad \text{(S4)}$$

Here and henceforth, we use the convention

$$\int d\mathbf{k} = \int_{k < k_F} \frac{d^2 p}{(2\pi)^2}, \quad \int d\mathbf{p}\, d\mathbf{q} = \int_{k_F < p, q < p_\xi} \frac{d^2 p}{(2\pi)^2} \frac{d^2 q}{(2\pi)^2}, \quad \text{(S5)}$$

where $p_\xi = \sqrt{2m\xi}$, and ξ is a UV cutoff. The pole of Eq. (S4) for $\mathbf{k}, k_F \to 0$ defines the vacuum binding energy $-E_b$ [4]. One can also define a 2D scattering length $a_{2D} = \sqrt{2m_r E_b}$ with reduced mass m_r, but we will not use this quantity further.

For $M = \infty$, the impurity is dispersionless, $E = 0$, and Eq. (S4) reduces to

$$T(\omega + \epsilon) = -\frac{1}{\rho} \frac{1}{\ln\left(\frac{\omega + \epsilon - \mu + i0^+}{-E_b}\right)}, \quad \text{(S6)}$$

with the density of states $\rho = m/2\pi$. E_{b} is given by $\xi e^{-1/(\rho V_0)}$. From Eq. (S6), the one-hole self-energy is obtained by closing the electron loop. This yields Eq. (2) of the main text:

$$\Sigma_1(\omega) = -\int_0^\mu d\epsilon_{\mathbf{k}} \frac{1}{\ln\left(\frac{\omega + \epsilon_{\mathbf{k}} - \mu + i0^+}{-E_{\mathrm{b}}}\right)}. \tag{S7}$$

The one-hole diagrams H_1 are obtained by reattaching the impurity lines, $H_1(\omega) = \Sigma_1(\omega) D_0(\omega)^2$.

Molecule-hole feature

To find the contribution of H_1 to the molecule feature, we expand Eq. (S7) around $\omega = -E_{\mathrm{b}}$. Thus

$$H_1(\omega) \simeq \frac{1}{E_{\mathrm{b}}} \ln\left(\frac{\omega + E_{\mathrm{b}} + i0^+}{\omega + E_{\mathrm{b}} - \mu}\right) \tag{S8}$$

$$\simeq \frac{1}{E_{\mathrm{b}}} \ln\left(\frac{\omega + E_{\mathrm{b}} + i0^+}{-\mu}\right) \text{ for } |\omega + E_{\mathrm{b}}| \ll \mu,$$

yielding the first logarithm in Eq. (5) of the main text.

Repulsive polaron

For the repulsive polaron, we need to evaluate Eq. (S7) for $\omega \gtrsim 0$. A useful formula is

$$\int dx \frac{x^n}{\ln(x)} = \begin{cases} \frac{x^n}{n+1}\frac{1}{\ln x} + \mathcal{O}\left(\frac{x^n}{\ln^2(x)}\right), & n \neq -1 \\ \ln(\ln(x)) & n = -1 \end{cases}. \tag{S9}$$

Schematically, this formula implies that the $1/\log$ terms can be pulled out from integrals with "logarithmic accuracy" (l.a.). As a result, we find

$$\Sigma_1(\omega) \simeq -\frac{\mu}{\ln\left(\frac{\mu}{E_{\mathrm{b}}}\right)} + \frac{\omega \cdot \ln\left(\frac{|\omega|}{\mu}\right)}{\ln\left(\frac{\mu}{E_{\mathrm{b}}}\right)\ln\left(\frac{|\omega|}{E_{\mathrm{b}}}\right)} - i\frac{\pi\omega\theta(\omega)}{\ln^2\left(\frac{|\omega|}{E_{\mathrm{b}}}\right)}. \tag{S10}$$

Upon resummation, the first term in Eq. (S10) shifts the repulsive polaron threshold to $\omega = \gamma\mu \simeq \omega_{\mathrm{th},2}$, with γ as defined in Eq. (3) of the main text. To interpret the other terms, we restrict ourselves to a parametrically large window of frequencies $\mu^2/E_{\mathrm{b}} \ll \omega \ll \mu$, which allows the simplification $\ln(\omega/E_{\mathrm{b}}) = \ln(\omega/\mu \cdot \mu/E_{\mathrm{b}}) \simeq \ln(\mu/E_{\mathrm{b}})$ with l.a. This restriction is specific to our diagrammatic approach, and is not required in the infinite mass treatment [5]. Thus, we believe that our results hold down all the way to $\omega \to 0$. Taking into account the threshold shift, i.e. shifting to $\nu_2 = \omega - \omega_{\mathrm{th},2}$, and reattaching the impurity lines yields:

$$H_1(\nu_2) \simeq \frac{\gamma^2}{\nu_2} \ln\left(\frac{\nu_2 + i0^+}{-\mu}\right). \tag{S11}$$

Taking the imaginary part leads to a spectrum $\mathcal{A}_2(\nu_2) \simeq \gamma^2\theta(\nu_2)/\nu_2$ given in the main text.

S.C. $M = \infty$: EVALUATION OF H_1 IN 3D

To substantiate our claim that our results apply to 3D in analogous fashion, here we present the evaluation of H_1 in 3D. We start from the infinite mass T-matrix analogous to Eq. (S4):

$$T(\Omega) = \left(-1/V_0 - \int d^3\mathbf{p} \frac{1}{\Omega - \epsilon_{\mathbf{p}} + i0^+}\right)^{-1}, \quad \Omega = \omega + \epsilon. \tag{S12}$$

The 3D integrals follow the convention of Eq. (S5) adapted to 3D (in this section only). To regularize the T-matrix, we apply the Lippmann-Schwinger equation [6, 7]

$$\frac{1}{-V_0} = \frac{m}{2\pi a} - \int_{>0} d^3\mathbf{p} \frac{1}{\epsilon_{\mathbf{p}}}, \tag{S13}$$

where the integral ranges over all momenta $0 < p < p_\xi$. As a result,

$$T(\Omega) = \tag{S14}$$

$$\left(\frac{m}{2\pi a} - \left[\int d^3\mathbf{p} \frac{1}{\Omega - \epsilon_{\mathbf{p}} + i0^+} - \int_> d^3\mathbf{p} \frac{1}{-\epsilon_{\mathbf{p}}}\right]\right)^{-1}.$$

After some straightforward algebra (see, e.g., [8, 9]), T can be rewritten as

$$T(\Omega) = \left(\frac{m}{2\pi a} - c_1 R(\Omega)\right)^{-1}, \quad c_1 = \frac{m^{3/2}}{\sqrt{2}\pi^2}, \tag{S15}$$

$$R(\Omega) = \theta(\Omega)\left(2\sqrt{\mu} + \sqrt{\Omega}\ln\left|\frac{\sqrt{\mu} - \sqrt{\Omega}}{\sqrt{\mu} + \sqrt{\Omega}}\right|\right) + \tag{S16}$$

$$\theta(-\Omega)\left(\pi\sqrt{-\Omega} + 2\sqrt{\mu} - 2\sqrt{-\Omega}\arctan\left(\frac{\mu}{\sqrt{-\Omega}}\right)\right)$$

$$- i\pi\sqrt{\Omega} \cdot \theta(\Omega - \mu).$$

In the vacuum limit $\mu \to 0$, T has a pole at $\Omega = -E_{\mathrm{b}} = -1/(2ma^2)$ as in 2D. From T, the one-hole self-energy is obtained by closing the loop

$$\Sigma_1(\omega) = \int d^3\mathbf{k}\, T(\omega + \epsilon_{\mathbf{k}} + i0^+). \tag{S17}$$

Molecule-hole feature

We focus on large binding energies, $E_{\mathrm{b}} \gg \mu \Leftrightarrow k_F a \ll 1$, and expand around $\omega = -E_{\mathrm{b}}$. Keeping terms up to order $\mathcal{O}(\mu^{3/2}/E_{\mathrm{b}})$, we find (compare also [7]):

$$\Sigma_1(\omega) = \frac{2\sqrt{E_{\mathrm{b}}}}{\pi}\int_0^\mu d\epsilon \frac{\sqrt{\epsilon}}{\omega + E_{\mathrm{b}} + \epsilon - \frac{2}{3}\gamma\mu + i0^+}, \tag{S18}$$

with $\gamma = k_F a/\pi$ as defined in Eq. (3) of the main text. From Eq. (S18) one can deduce, reattaching the hole lines

$$H_1(\omega) \simeq \tag{S19}$$

$$\frac{2\sqrt{\mu}}{\pi E_{\mathrm{b}}^{3/2}} \ln\left(\frac{\omega + E_{\mathrm{b}} - \omega_0 + i0^+}{-\mu}\right), \quad \omega_0 = -\mu + \frac{2}{3}\gamma\mu,$$

which holds for $|\omega + E_{\rm b} - \omega_0| \ll \mu$. This result is very similar to 2D, Eq. (S8), apart from a different non-universal prefactor, and a shift of the molecule feature by ω_0. Thereby, already at one-hole level the molecule is placed at the right energy $\omega \simeq -E_{\rm b} + \omega_0 \simeq \omega_{\rm th,1}$ up to order $\mathcal{O}(a)$, as can be checked by inserting Eq. (S2) into (S3). The energy shift might cause some technical modifications at the two-hole level, that is, for diagrams H_2 in Fig. 3 of the main text (whose 3D evaluation is beyond the scope of this work), but the overall 2D strategy should remain valid. Let us note that resumming the self-energy of Eq. (S18) with Dyson's equation (which is incorrect for large masses as explained in the main text) yields a spurious attractive polaron, determined from

$$\omega - \Sigma_1(\omega) = 0. \qquad (S20)$$

For $E_{\rm b} \gg \mu$ this equation is readily solved, and yields $\omega = \omega_{\rm th,1} - \mathcal{O}\left(\mu \exp(-1/\gamma)\right)$, i.e. the gap between the continuum and the polaron is only exponentially small. Correct evolution of H_2 and higher diagrams should eliminate the polaron as in 2D.

Repulsive polaron

For $\omega \gtrsim 0$, $\Sigma_1(\omega)$ reads

$$\Sigma_1(\omega) = \int_0^\mu d\epsilon \frac{\sqrt{\epsilon}}{\sqrt{\mu}/\gamma - \left(2\sqrt{\mu} + \sqrt{\omega + \epsilon}\ln\left|\frac{\sqrt{\mu}-\sqrt{\omega+\epsilon}}{\sqrt{\mu}+\sqrt{\omega+\epsilon}}\right|\right)}. \qquad (S21)$$

We restrict ourselves to frequencies $\mu \exp(-1/\gamma) \ll \omega \ll \mu$ similar to 2D (except that the small parameter γ is not logarithmic anymore). Then we find, with l.a.:

$$\Sigma_1(\omega) = \frac{2}{3}\gamma\mu + \gamma^2\omega \ln\left(\frac{\omega}{\mu}\right) - i\pi\gamma^2\omega\theta(\omega), \qquad (S22)$$

which is in full agreement with the 2D result of Eq. (S10) except for the factor $2/3$ dictated by Fumi's theorem.

S.D. $M = \infty$: EVALUATION OF H_2 IN THE TIME DOMAIN IN 2D

The evaluation of the diagrams H_2 drawn in Fig. 3 of the main text is similar to Ref. [10]. Let us first focus on the diagrammatic series H_2^a, and evaluate the corresponding contribution to the self-energy part Σ_2^a, i.e., amputate the external impurity lines first. The relevant diagram is redrawn in Fig. S2. We specialize on energies $\omega \simeq -E_{\rm b}$. When the energy is measured from the impurity level, the time-domain impurity Green function for infinite mass reduces to a step-function: $D_0(t) = -i\theta(t)$. Thus, the impurity lines impose the time-ordering of the interactions only. We parenthetically note that for finite mass, the impurity propagator aquires a non-trivial momentum-dependence, which obstructs the time-domain evaluation, and is the reason for going into the more complicated frequency domain calculation in the next Section. The general expression for all diagrams which preserve the structure of Σ_2^a, with the interaction lines at initial and final times connecting to the lower part of the "horseshoe", reads

$$\Sigma_2^a(t) = -iV_0^2 \sum_{n=1}^\infty (-V_0)^n \sum_{m=0}^\infty (-V_0)^m \theta(t) \int d\mathbf{k}_x d\mathbf{k}_y \qquad (S23)$$

$$\int_{0<T_1\ldots<T_n<t} dT_1\ldots dT_n \int d\mathbf{q}_1\ldots d\mathbf{q}_{n-1} G(\mathbf{k}_x, T_1 - t)G(\mathbf{q}_2, T_2 - T_1)\ldots G(\mathbf{q}_{n-1}, T_n - T_{n-1})G(\mathbf{k}_y, -T_n)$$

$$\int_{0<t_1\ldots<t_m<t} dt_1\ldots dt_m \int d\mathbf{p}_1\ldots \int d\mathbf{p}_{m+1} G(\mathbf{p}_1, t_1)\ldots G(\mathbf{p}_m, t_m - t_{m-1})G(\mathbf{p}_{m+1}, t - t_m),$$

where $G(\mathbf{k}, t) = -i(\theta(t) - n_\mathbf{k})\exp(-i\epsilon_\mathbf{k} t)$, and $n_\mathbf{k} = \theta(k_F - k)$ is the zero temperature Fermi function. Introducing retarded Green functions as $G^R(t) = G(t)\theta(t)$, Eq. (S23) can be rewritten as:

$$\Sigma_2^a(t) = -iV_0^2 \int_{<k_F} d\mathbf{k}_x d\mathbf{k}_y \exp(i(\epsilon_{\mathbf{k}_x} + \epsilon_{\mathbf{k}_y})t) \cdot A(t)B(t) \qquad (S24)$$

$$A(t) = \sum_{n=1}^\infty (-V_0)^n \int d\mathbf{q}_1\ldots d\mathbf{q}_{n-1} \left[G^R(\mathbf{k}_x, \cdot) * G^R(\mathbf{q}_1, \cdot) * \ldots * G^R(\mathbf{q}_{n-1}, \cdot) * G^R(\mathbf{k}_y, \cdot)\right](t) \qquad (S25)$$

$$B(t) = \sum_{m=0}^\infty (-V_0)^m \int d\mathbf{p}_1\ldots d\mathbf{p}_{m+1} \left[G^R(\mathbf{p}_1, \cdot) * \ldots * G^R(\mathbf{p}_{m+1}, \cdot)\right](t), \qquad (S26)$$

where $*$ denotes convolutions, $[f * g](t) = \int d\tilde{t} f(t - \tilde{t})g(t)$.

Fourier-transformation using the convolution theorem
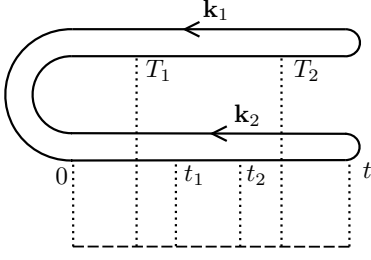
turns the convolutions into a geometric series, which are

Figure S2. A representative of the series Σ_2^a, see also Fig. 3 of the main text.

resummed in the same way as the T-matrix, Eq. (S4). Changing to energy integrations for the remaining two momentum integrals over $\mathbf{k}_x$ and $\mathbf{k}_y$ results in

$$\Sigma_2^a(\omega) = -i \int_0^\mu dx dy \int_{-\infty}^\infty \frac{d\omega_1}{2\pi} \frac{1}{\omega_1 - x + i0^+} \frac{1}{\omega_1 - y + i0^+}$$
$$\frac{1}{\ln\left(\frac{\omega_1 - \mu + i0^+}{-E_\mathrm{b}}\right)} \frac{1}{\ln\left(\frac{\Omega_1 - \mu + i0^+}{-E_\mathrm{b}}\right)}, \quad \Omega_1 = \omega + x + y - \omega_1. \tag{S27}$$

For the remaining ω_1-integration we use the following approach: we split the $1/\ln$-terms into a part containing a pole and a part containing a branch cut:

$$\frac{1}{\ln\left(\frac{\omega_1 - \mu + i0^+}{-E_\mathrm{b}}\right)} = \frac{-E_\mathrm{b}}{\omega_1 + E_\mathrm{b} - \mu + i0^+} \tag{S28}$$
$$+ \left(\frac{1}{\ln\left(\frac{\omega_1 - \mu + i0^+}{-E_\mathrm{b}}\right)} + \frac{E_\mathrm{b}}{\omega_1 + E_\mathrm{b} - \mu + i0^+}\right).$$

The first part, containing the pole, can be interpreted as bound state propagator, while the second part, containing the branch cut, corresponds to the continuum contribution; this is also in agreement with the evaluation of H_1 in Sec. S.B. Employing the spectral representation, it is easily shown that the combination of the branch cut contributions for both $1/\ln$-functions yields a result with vanishing imaginary part for $\omega \simeq -E_\mathrm{b}$. Since the spectrum is determined by the latter, we omit this part. The remainder is evaluated using Cauchy's theorem:

$$\Sigma_2^a(\omega) = \int_0^\mu dx dy \frac{E_\mathrm{b}}{\omega + E_\mathrm{b} + x - \mu + i0^+} \tag{S29}$$
$$\frac{1}{\omega + E_\mathrm{b} + y - \mu + i0^+} \frac{1}{\ln\left(\frac{\omega + E_\mathrm{b} + x + y - 2\mu + i0^+}{-E_\mathrm{b}}\right)}.$$

Restricting to $\mu^2/E_\mathrm{b} \ll \omega + E_\mathrm{b} \ll \mu$ as explained below Eq. (S10), we find, in agreement with Eq. (5) of the main text.

$$H_2^a(\omega) \simeq -\frac{\gamma}{E_\mathrm{b}} \ln^2\left(\frac{\omega + E_\mathrm{b} + i0^+}{-\mu}\right). \tag{S30}$$
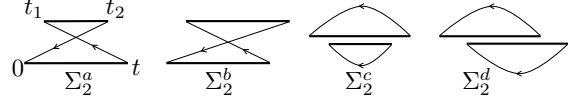

Figure S3. Two-hole diagrams Σ_2 (with amputated impurity lines) in bold-line representation. Initial, final and intermediated times are indicated for Σ_2^a only.

The diagrams $H_2^{c,d}$ can be evaluated along the same lines; the evaluation of H_2^b requires a "generalized convolution theorem":

$$\mathcal{F}\left(\int_{-\infty}^\infty dt_1 f(t - t_1) g(t, t_1)\right)(\Omega) = \tag{S31}$$
$$\int_{-\infty}^\infty \frac{d\omega_1}{2\pi} f(\omega_1) g(\Omega - \omega_1, \omega_1),$$

where $\mathcal{F}$ denotes the Fourier transform, $\mathcal{F}(f)(\Omega) = \int dt f(t) \exp(i\Omega t)$, and f and g are any two well-behaved functions; for more details, see Appendix B.2 of Ref. [10]. There is, however, a quicker way to arrive at the results: one can combine the impurity interacting with a forward propagating electron into a "bold" propagator of the form

$$D_\mathrm{b}(t) \propto \int_{-\infty}^\infty \frac{d\omega_1}{2\pi} \exp(-it\omega_1) \frac{\theta(t)}{\ln\left(\frac{\omega_1 - \mu + i0^+}{-E_\mathrm{b}}\right)}, \tag{S32}$$

The self-energy parts of the series H_2 of Fig. 3 in the main text can be redrawn in this bold-line representation as shown in Fig. S3. The thin lines correspond to the two holes, with Green function $G(x,t) = i\theta(-t)n_F(x)\exp(-ixt)$, where x is an energy variable as in Eq. (S27). With these ingredients, $\Sigma_2^a(t)$ can immediately be written down in a closed form up to an overall phase factor:

$$\Sigma_2^a(t) \propto \int_0^\mu dx dy \int_0^t dt_2 \int_0^{t_2} dt_1$$
$$D_\mathrm{b}(t_2 - t_1) D_\mathrm{b}(t) G(x, -t_2) G(y, t_1 - t) \tag{S33}$$

By Fourier transformation one easily reproduces Eq. (S27); the overall phase is fixed referring to the ordinary diagrams. The time domain representations of the remaining diagram read

$$\Sigma_2^b(t) \propto \int_0^\mu dx dy \int_0^t dt_2 \int_0^{t_2} dt_1 \tag{S34}$$
$$D_\mathrm{b}(t - t_1) D_\mathrm{b}(t_2) G(x, -t) G(y, t_1 - t_2),$$

$$\Sigma_2^c(t) \propto \int_0^\mu dx dy \int_0^t dt_2 \int_0^{t_2} dt_1 \tag{S35}$$
$$D_\mathrm{b}(t) G(x, -t) D_\mathrm{b}(t_2 - t_1) G(y, t_1 - t_2),$$

$$\Sigma_2^d(t) \propto \int_0^\mu dx dy \int_0^t dt_1 \int_0^t dt_2 \tag{S36}$$
$$D_\mathrm{b}(t_1) G(x, -t_1) D_\mathrm{b}(t - t_2) G(y, t_2 - t).$$

Fourier transforming with help of Eq. (S31), and splitting into pole and branch cut contributions as described in Eq. (S28), we find for $H_2^{b,c}$

$$H_2^b(\omega) + H_2^c(\omega) \simeq \frac{1}{E_{\mathrm{b}}} \int_0^\mu dx \frac{1}{(\omega + E_{\mathrm{b}} + x - \mu + i0^+)^2}$$
$$\cdot \left(-\mu - \int_0^\mu dy \frac{1}{\ln\left(\frac{\omega + E_{\mathrm{b}} + x + y - 2\mu + i0^+}{-E_{\mathrm{b}}} \right)} \right). \qquad (S37)$$

Comparison with Eqs. (S17), (S10) shows that these contributions shift the molecule threshold, $E_{\mathrm{b}} \to E_{\mathrm{b}} + \mu(1 - \gamma)$, as claimed in the main text. This shift is the only relevant self-energy effect: the non-trivial frequency-dependence of the molecule self-energy involves a factor $\gamma^2 \ln[(\omega + E_{\mathrm{b}})/(-\mu)]$, which is subleading compared to Eq. (S30).

In addition, we find that H_2^d is subleading with an extra factor μ/E_{b} compared to the other diagrams, and we may therefore savely neglect it.

S.E. $M = \infty$: EVALUATION OF H_2 IN THE FREQUENCY DOMAIN IN 2D

The evaluation of H_2 in the time-domain is instructive, but crucially depends on the fact that the impurity is dispersionless. Thus, it does not simply generalize to finite mass impurities. To circumvent this problem, we first recover the results of the previous section in the frequency domain, which allows for extension to the finite mass case. In this approach, the diagrams are organized by the number of T-matrices. For H_2^a, the first two diagrams are shown in Fig. 3(b) of the main text. The lowest order diagram (3 T-matrices) reads

$$H_{2,1}^a(\omega) \simeq \qquad (S38)$$
$$\frac{1}{E_{\mathrm{b}}^2} \int_0^\mu dx dy \int_\mu^\xi d\epsilon_p \frac{1}{\ln\left(\frac{\omega + x - \mu + i0^+}{-E_{\mathrm{b}}} \right)} \frac{1}{\ln\left(\frac{\omega + y - \mu + i0^+}{-E_{\mathrm{b}}} \right)}$$
$$\frac{1}{\ln\left(\frac{\omega + x + y - \epsilon_p - \mu + i0^+}{-E_{\mathrm{b}}} \right)} \frac{1}{\omega - \epsilon_p + x + i0^+} \frac{1}{\omega - \epsilon_p + y + i0^+}.$$

For $\omega \simeq -E_{\mathrm{b}}$, the ϵ_p-integral is dominated by the pole of the third logarithm, around which we can expand. The remaining Green functions effectively cut off the integration at $\epsilon_p \simeq E_{\mathrm{b}}$. Since this scale only appears in the argument of a logarithm (see below), an exact determination

is not required within l.a.. Thus, we can approximate

$$H_{2,1}^a(\omega) \simeq -\frac{1}{E_{\mathrm{b}}} \int_0^\mu dx dy \qquad (S39)$$
$$\frac{1}{\omega + E_{\mathrm{b}} + x - \mu + i0^+} \frac{1}{\omega + E_{\mathrm{b}} + y - \mu + i0^+} \cdot I$$
$$I = \int_\mu^{E_{\mathrm{b}}} d\epsilon_p \frac{1}{\omega + E_{\mathrm{b}} + x + y - \epsilon_p - \mu + i0^+} \simeq \quad (S40)$$
$$\ln\left(\frac{\omega + E_{\mathrm{b}} + x + y - 2\mu + i0^+}{-E_{\mathrm{b}}} \right), \ \omega \simeq -E_{\mathrm{b}}.$$

A similar evaluation of the second diagram of the series H_2^a reproduces Eq. (S39) with I replaced by I^3. Extrapolating this behavior, the higher order diagrams yield a series $I + I^3 + I^5 + \ldots = I/(1 - I^2) \simeq 1/(-I)$. In total:

$$H_2^a(\omega) = \frac{1}{E_{\mathrm{b}}} \int_0^\mu dx dy \frac{1}{\omega + E_{\mathrm{b}} + x - \mu + i0^+} \qquad (S41)$$
$$\frac{1}{\omega + E_{\mathrm{b}} + y - \mu + i0^+} \frac{1}{\ln\left(\frac{\omega + E_{\mathrm{b}} + x + y - 2\mu + i0^+}{-E_{\mathrm{b}}} \right)},$$

in agreement with Eq. (S29). The remaining expressions $H_2^{b,c,d}$ can be evaluated along the same lines. For $H_2^{b,c}$, the result is in perfect agreement with Eq. (S37). For H_2^d, one arrives at

$$H_2^d(\omega) = \frac{1}{E_{\mathrm{b}}} \int_0^\mu dx dy \qquad (S42)$$
$$\frac{1}{\omega + E_{\mathrm{b}} + x - \mu + i0^+} \frac{1}{\omega + E_{\mathrm{b}} + y - \mu + i0^+} \cdot \frac{1}{I^2}.$$

Since I is a large logarithm of order $1/\gamma$, this result is subleading, although only by a factor γ and not μ/E_{b} as in the time domain; this slight discrepancy can be attributed to the evaluation with l.a.

S.F. THE MOLECULE CONTINUUM FOR FINITE MASS

With the infinite mass results for H_1, H_2 at hand, we move to the finite mass case, focusing on $\omega \simeq -E_{\mathrm{b}}$. We start from re-evaluation of Eq. (S4) to leading order in $\beta = m/M$:

$$T(\omega + \epsilon, \mathbf{k}) = -\frac{1 + \beta}{\rho} \frac{1}{\ln\left(\frac{\omega + \epsilon - \mu\beta - \mu - k^2/2M_+ + i0^+}{-E_{\mathrm{b}}} \right)}, \qquad (S43)$$

where $M_+ = m + M$, and the vacuum binding energy E_{b} now reads $(1 + \beta)\xi e^{-(1+\beta)/(\rho V_0)}$. In the following, we will only resolve the factors $(1 + \beta)$ in the numerator of the logarithms, since the other factors just rescale the energies. Closing the contour to obtain $H_1(\omega)$, we find

$$H_1(\omega) \simeq \frac{1}{E_{\mathrm{b}}} \int_0^\mu dx \frac{1}{\omega + E_{\mathrm{b}} - \beta\mu + x - \mu - \beta x + i0^+}$$
$$\simeq \ln\left(\frac{\omega + E_{\mathrm{b}} - 2\beta\mu + i0^+}{-\mu} \right). \qquad (S44)$$

The resulting contribution to the spectrum is given by

$$-\,\mathrm{Im}\,[H_1(\omega)] \tag{S45}$$
$$= \frac{1}{E_{\mathrm{b}}} \int_0^\mu dx\, \delta\left(\omega - (-E_{\mathrm{b}} + \beta\mu + \beta x + (\mu - x))\right).$$

The energy-conservation imposed by the delta-function describes the following process: an impurity with energy ω decays into a bound state, with "potential energy" $-E_{\mathrm{b}} + \beta\mu$ and kinetic energy βx, and a hole with energy $(\mu - x)$. At the treshold, the hole peels off right at the Fermi surface, and the bound state has a kinetic energy

$\beta\mu$. The modification of the binding energy in the presence of a Fermi sea $-E_{\mathrm{b}} \to -E_{\mathrm{b}} + \beta\mu$ is only of secondary importance, since there are further molecule self-energy diagrams with renormalize the binding energy anyway (see Eq. (S37)). We will not evaluate these in detail for finite mass, and just write $\nu_1 = \omega + E_{\mathrm{b}} - \beta\mu$ henceforth. Thus, we see that the logarithm in Eq. (S44) is peaked at the "direct threshold" $\nu_1 = \beta\mu$, involving the creation of a molecule with momentum k_F.

Next, we evaluate the lowest order contribution (3 T-matrices) to diagram H_2^a. Resumming the logarithms, the finite mass generalization of Eq. (S38) reads

$$H_{2,1}^a(\omega) = \frac{1}{E_{\mathrm{b}}^2 \rho^3} \int d\mathbf{k}_x d\mathbf{k}_y \int d\mathbf{p} \frac{1}{\omega - \epsilon_p + x - E_{\mathbf{k}_x - \mathbf{p}} + i0^+} \frac{1}{\omega - \epsilon_p + y - E_{\mathbf{k}_y - \mathbf{p}} + i0^+} \tag{S46}$$
$$\frac{1}{\ln\left(\frac{\omega + x - \mu(1+\beta) - E_{\mathbf{k}_x} + i0^+}{-E_{\mathrm{b}}}\right)} \frac{1}{\ln\left(\frac{\omega + y - \mu(1+\beta) - E_{\mathbf{k}_y} + i0^+}{-E_{\mathrm{b}}}\right)} \frac{1}{\ln\left(\frac{\omega + x + y - \epsilon_p - \mu(1+\beta) - E_{\mathbf{k}_x + \mathbf{k}_y - \mathbf{p}} + i0^+}{-E_{\mathrm{b}}}\right)}.$$

To compute the $\mathbf{p}$-integral, we expand the last $1/\ln$-function around its pole. The resulting non-trivial logarithmic integral reads

$$\tilde{I} = \frac{1}{\rho} \int_{k_F < \mathbf{p} < \sqrt{2mE_{\mathrm{b}}}} d\mathbf{p} \tag{S47}$$
$$\frac{1}{\nu_1 + x + y - \mu - \epsilon_p - E_{\mathbf{k}_x + \mathbf{k}_y - \mathbf{p}} + i0^+},$$

c.f. Eq. (S40). Integration with logarithmic accuracy (which only gives access to $\mathrm{Re}[\tilde{I}]$) yields

$$\mathrm{Re}[\tilde{I}] \simeq \ln\left(\frac{\max(\nu_1 + x + y - 2\mu, \beta\mu)}{-E_{\mathrm{b}}}\right). \tag{S48}$$

Postponing evaluation of $\mathrm{Im}[\tilde{I}]$, one can extrapolate to the full series $H_{2,1}^a$ as in Sec. S.E. Repeating this procedure for $H_2^{b,c,d}$, we find

$$H_2^a(\nu_1) \simeq \frac{1}{E_{\mathrm{b}}\rho^2} \int d\mathbf{k}_x d\mathbf{k}_y \frac{1}{\nu_1 + x - \mu - E_{\mathbf{k}_x} + i0^+}$$
$$\frac{1}{\nu_1 + y - \mu - E_{\mathbf{k}_y} + i0^+} \cdot \frac{1}{\tilde{I}} \tag{S49}$$

$$H_2^b(\nu_1) \simeq -\frac{1}{E_{\mathrm{b}}\rho^2} \int d\mathbf{k}_x d\mathbf{k}_y \frac{1}{(\nu_1 + x - \mu - E_{\mathbf{k}_x} + i0^+)^2}$$
$$\cdot \left(1 + \mathcal{O}\left(1/\tilde{I}^2\right)\right) \tag{S50}$$

$$H_2^c(\nu_1) \simeq -\frac{1}{E_{\mathrm{b}}\rho^2} \int d\mathbf{k}_x d\mathbf{k}_y \frac{1}{(\nu_1 + x - \mu - E_{\mathbf{k}_x} + i0^+)^2} \cdot \frac{1}{\tilde{I}} \tag{S51}$$

$$H_2^d(\nu_1) \simeq +\frac{1}{E_{\mathrm{b}}\rho^2} \int d\mathbf{k}_x d\mathbf{k}_y \frac{1}{(\nu_1 + x - \mu - E_{\mathbf{k}_x} + i0^+)}$$
$$\frac{1}{(\nu_1 + y - \mu - E_{\mathbf{k}_y} + i0^+)} \cdot \frac{1}{\tilde{I}^2}. \tag{S52}$$

Since $\tilde{I}$ is still a large logarithm of order $1/\gamma$, H_2^d is subleading as for infinite mass. The other contributions behave as follows: with l.a., H_2^a reads

$$H_2^a(\nu_1) \simeq -\frac{\gamma}{E_{\mathrm{b}}} \ln^2\left[(\nu_1 - \beta\mu + i0^+)/(-\mu)\right], \tag{S53}$$

i.e., essentially the same result as in the infinite mass case, Eq. (S30), except that the peak of the logarithm is at the direct threshold, as discussed below Eq. (S45). $H_2^{b,c}$ again act as molecular self-energy terms. First, their real parts lead to a shift of E_{b}, which we do not compute. More importantly, the imaginary part of H_2^c cuts off the logarithmic singularity at the direct threshold. This can be seen extracting the molecule self-energy part from Eq. (S51):

$$\Sigma_{\mathrm{mol}}(\nu_1, \mathbf{k}_x) = -\frac{1}{\rho} \int d\mathbf{k}_y \frac{1}{\tilde{I}}. \tag{S54}$$

Using Eq. (S47), we find

$$\mathrm{Im}\left[\Sigma_{\mathrm{mol}}\right](\nu_1, \mathbf{k}_x) = \int d\mathbf{k}_y \frac{\mathrm{Im}[\tilde{I}]}{\mathrm{Re}[\tilde{I}]^2 + \mathrm{Im}[\tilde{I}]^2} \simeq \tag{S55}$$
$$-\gamma^2 \pi \int d\mathbf{k}_y d\mathbf{p}\, \delta(\nu_1 + x + y - \mu - \epsilon_p - E_{\mathbf{k}_x + \mathbf{k}_y - \mathbf{p}}).$$

Evaluation of this standard phase space integral (see e.g. Appendix E of Ref. [10] for examples) for $k_x = k_F$ yields

$$\mathrm{Im}[\Sigma_{\mathrm{mol}}(\nu_1, k_F)] \propto -\gamma^2 \frac{\nu_1^2}{\beta\mu}. \tag{S56}$$

Near the direct threshold, $\nu_1 \simeq \beta\mu$, this leads to a molecule decay rate $\Gamma \propto \gamma^2 \beta\mu$. Appropriately resummed,

this rate cuts all logarithms; e.g., the one-hole result of Eq. (S44) is modified as

$$H_1(\nu_1) \simeq \frac{1}{E_{\mathrm{b}}} \ln \left(\frac{\nu_1 - \beta\mu + i\Gamma}{-\mu} \right)$$

$$\overset{l.a.}{\simeq} \frac{1}{E_{\mathrm{b}}} \ln \left(\max[\nu_1 - \beta\mu, \gamma^2 \beta\mu]/\mu \right), \qquad (S57)$$

and likewise for the term H_2^a in Eq. (S53). The physical reason for this cut-off is the decay of the molecule at the direct threshold with $k = k_F$ into a zero momentum molecule, two holes and an electron. This process shifts the threshold to the "indirect" one at $\nu_1 = 0$. For $0 < \nu_1 \ll \beta\mu$, the spectrum is perturbative (i.e., no large logarithms need to be resummed), and can be obtained from $\mathrm{Im}[H_2^a, H_2^c]$, Eqs. (S49), (S51). For spinless electrons, these contributions cancel to leading order. For spinful electrons, H_2^c incurs an extra factor of two, and the perturbative spectrum reads

$$\mathcal{A}_{\mathrm{pert}}(\nu_1) \simeq \gamma^2 \frac{1}{E_{\mathrm{b}}(\beta\mu)^2} \frac{\pi}{\rho^2} \int d\mathbf{k}_x d\mathbf{k}_y d\mathbf{p} \qquad (S58)$$

$$\delta(\nu_1 + x + y - \epsilon_p - \mu - E_{\mathbf{k}_x + \mathbf{k}_y - \mathbf{p}}) \propto \frac{\gamma^2}{E_{\mathrm{b}}} \left(\frac{\nu_1}{\beta\mu} \right)^3 \theta(\nu_1)$$

in 2D, while in 3D the extra phase space restriction should lead to $\mathcal{A}_{\mathrm{pert}} \propto \nu_1^{7/2}$ [11]. Exponentiating the cut-off logarithms (S57) with a correct imaginary part to capture the perturbative spectrum yields Eq. (7) of the main text; the square root is yet another, continuous reformulation of the logarithm cutoff.

S.G. $M = \infty$: REPULSIVE POLARON FROM THE LINKED-CLUSTER APPROACH

The leading contribution to the repulsive polaron for $M = \infty$ was already obtained in Eq. (S11). The full power law singularity can be reproduced in a linked cluster approach (see also Refs. [3], [12]). One starts from the following set of identities for the impurity propagator

$$D(t) = -i\theta(t) \langle 0|S(t)|0\rangle \qquad (S59)$$

$$\langle 0|S(t)|0\rangle = \exp \left(\sum_n F_n(t) \right) \qquad (S60)$$

$$F_n(t) = \frac{(-i)^n}{n} \int_0^t dt_1 \ldots \int_0^t dt_n \langle 0|\hat{T}\left\{ \hat{V}(t_1) \ldots \hat{V}(t_n) \right\}|0\rangle \qquad (S61)$$

$$\hat{V}(t_i) = -V_0 \sum_{\mathbf{k},\mathbf{p}} c_{\mathbf{k}}^\dagger c_{\mathbf{p}} \theta(t_i)\theta(t - t_i), \qquad (S62)$$

where $S(t)$ is the S-matrix, and $\hat{T}$ the time-ordering operator. Note that the impurity has effectively been eliminated from the problem, which results in Feynman diagrams such as those shown in Fig. S4. The expressions above imply an expansion in the bare interaction V_0. Our
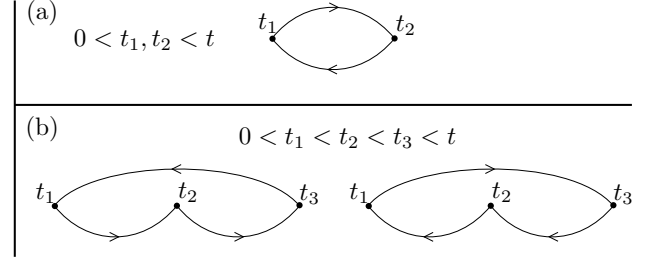

Figure S4. (a) $n = 2$ cluster diagram. Full lines indicate electron Green functions, and dots V_0-insertions. Internal times t_1, t_2 independently range from 0 to t. (b) *Time-ordered $n = 3$ cluster diagrams.* The left diagram shows an ordering with one hole, which is invariant under cyclic permutation of times. The right diagram contains two holes.

goal is to substitute this by an expansion in number of holes, getting rid of the V_0-dependence. Let us collect all one-hole diagrams: one is drawn in Fig. S4(a). It is more convenient to reexpress it imposing a time-ordering $0 < t_1 < t_2 < t$. This results in a factor of two which cancels the factor $1/2$ in Eq. (S61) for $n = 2$. Generalizing this approach, the one-hole diagrams can be extracted by drawing "loop diagrams" containing n interaction insertions at times $t_1, \ldots t_n$, and reordering them in such a fashion that only one electron propagates backwards in time; this cancels the factor $1/n$ in Eq. (S61). An example for $n = 3$ is shown in Fig. S4(b). Performing this reorganization, at one-hole level we can write

$$D(t) \simeq \exp(C(t)), \qquad (S63)$$

$$C(t) = -\int \frac{d\omega}{\pi} \left(\frac{-it}{\omega} - \frac{1}{\omega^2} \left(\exp(-it\omega) - 1 \right) \right) N(\omega) \qquad (S64)$$

$$N(\omega) = \mathrm{Im} \left\{ \int_0^\mu dx \frac{1}{\ln \left(\frac{\omega + x - \mu + i0^+}{-E_{\mathrm{b}}} \right)} \right\} \qquad (S65)$$

where we resummed the one-hole diagrams similar to Sec. S.B and employed the spectral representation of the retarded $1/\ln$-function. In Eq. (S64), the part linear in t just shifts the polaron threshold, and we may omit it. $N(\omega)$ measures the phase space for scattering of polarons with Fermi electrons. Eqs. (S17) and (S10) show that in the most important spectral window $\mu^2/E_{\mathrm{b}} \ll \omega \ll \mu$ we can approximate $N(\omega) \simeq \gamma^2 \pi\omega$. Therefore, evaluation of Eqs. (S63) and (S64) similar to Sec. 8.3.C of Ref. [3] and Fourier transformation directly results in a repulsive polaron spectrum

$$\mathcal{A}_2(\nu_2) \propto \frac{\theta(\nu_2)}{\mu} \left(\frac{\nu_2}{\mu} \right)^{\gamma^2 - 1} \qquad (S66)$$

with frequencies measured from the polaron threshold.

S.H.   THE REPULSIVE POLARON FOR FINITE MASS

Modified linked-cluster approach

The procedure above can also be adapted for finite mass (see, e.g., Sec. 3.6.B of [3] or [13]). In effect, we need to reevalute the phase-space factor $N(\omega)$ at one-hole level, and find, in 2D

$$N(\omega) \simeq \gamma^2 \pi \int d\mathbf{k} d\mathbf{p} \; \delta\left(\omega + \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{p}} - E_{\mathbf{p}-\mathbf{k}}\right)$$

$$\simeq \begin{cases} \gamma^2 \omega^{3/2}/\sqrt{\beta\mu}\theta(\omega) & \omega \ll \beta\mu \\ \gamma^2 \pi \omega & \beta\mu \ll \omega \ll \mu, \end{cases} \quad \text{(S67)}$$

and $N(\omega) \propto \omega^2$ for $\omega \ll \beta\mu$ in 3D. Thus, for energies beyond the recoil energy $\beta\mu$, the phase space factor assumes the infinite mass form. For smaller energies, the scattering phase space is suppressed, since processes where the polaron is scattered to large momenta of order k_F involve a minimal energy cost of order $\beta\mu$ [14].

We insert Eq. (S67) into (S64) and first study the limit $t \to \infty$. Again ignoring the term linear in t, in the infinite mass case one can show that $C(t)$ diverges as $-\gamma^2 \ln(|t|\mu)$. In contrast, for finite mass we find $\lim_{t\to\infty} C(t) \simeq \gamma^2 \ln(\beta)$. Inserted into the Green function of Eq. (S63), this limit gives rise to a finite quasiparticle-weight of the polaron, $Z \propto \beta^{\gamma^2}$. Again, the emergence of this quasi-particle weight is a consequence of the restricted low-energy scattering phase space, which partially reduces the repulsive polaron to its non-interacting form. Moreover, we can extract the incoherent polaron spectrum for small detuning from the threshold $0 < \nu_2 < \beta\mu$ simply by expanding the exponential in Eq. (S64), since there is no large logarithmic quantity to prevent it. This yields, in 2D

$$\mathcal{A}(\nu_2) \propto \gamma^2 \frac{1}{\sqrt{\beta\mu\nu_2}}, \quad \text{(S68)}$$

while in 3D the incoherent part is approximately constant $\propto \gamma^2/(\beta\mu)$. For $\nu_2 \gg \beta\mu$ one recovers the infinite mass behavior. Interpolating between these two limits yields formula (8) of the main text, apart from the finite width of the repulsive polaron quasiparticle to be discussed below.

Width of the repulsive polaron

In the previous section, we only considered the decay of the single impurity into particle-hole excitations, but neglected the decay into the molecular state. To incorporate this process, we need to go to two-hole level. The decay rate vanishes for infinite mass, since the problem becomes single-particle, hence the molecule and repulsive polaron sectors decouple. Indeed, starting from the exact expressions for the impurity self-energy Σ_2 in the time
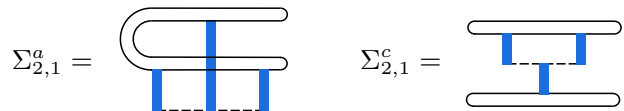

Figure S5. Diagrams $\Sigma_{2,1}^{a;c}$ contributing to the decay of the repulsive polaron

domain, of a form similar to Eq. (S27), one can show that $\sum_i \text{Im}[\Sigma_2^i](\omega = 0^-) = 0$, where the frequency argument 0^- is chosen to exclude the UV-tail arising from electron-hole excitations. This cancellation implies that the repulsive polaron does not acquire a Lorentzian IR-tail for infinite mass. This is not necessarily true for finite mass. An evaluation of all two-hole diagrams for $\omega = 0^-$ similar to Sec. S.E appears too involved. A simpler estimate can be given by restriction to the "first-order" diagrams $\Sigma_{2,1}$ with the minimal number of T-matrices. The diagrams with a minimal number of 3 T-matrices and nonvanishing imaginary parts for $\omega = 0^-$ are $\Sigma_{2,1}^{a,c}$, shown in Fig. S5.

The rate resulting from these diagrams has been evaluated effectively for $E_\text{b} \to 0$ in Ref. [15] for 3D, but here we focus on $E_\text{b} \gg \mu$. We approximate the central T-matrix by a pole to incorporate the molecule, and the remaining two T-matrices by γ. Taking the imaginary part, we find

$$\tilde{\Gamma}_2 \simeq \gamma^2 \frac{E_\text{b}}{\rho^3} \int d\mathbf{k}_x d\mathbf{k}_y d\mathbf{p} \; \delta(x + y - \epsilon_p - E_{\mathbf{k}_x + \mathbf{k}_y - \mathbf{p}} + E_\text{b})$$

$$\frac{1}{\epsilon_p - x + E_{\mathbf{k}_x - \mathbf{p}}} \left(\frac{1}{\epsilon_p - x + E_{\mathbf{k}_x - \mathbf{p}}} - \frac{1}{\epsilon_p - y + E_{\mathbf{k}_y - \mathbf{p}}} \right)$$

$$\propto \gamma^2 \frac{\mu^4}{E_\text{b}^3} (1 - 3\beta), \quad \text{(S69)}$$

where the last estimate holds to leading order in $\mu/E_\text{b}, \beta$. In (S69) we may omit the β-independent part, since the infinite mass cancels upon complete evaluation as discussed above. The remainder can be used to estimate $\Gamma_2 \sim \beta\gamma^2 \frac{\mu^4}{E_\text{b}^3}$. For 3D, we expect a similar behavior, although details of the scaling could be different.

[1] S. K. Adhikari, Am. J. Phys. **54**, 362 (1986).

[2] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Phys. Rev.

B **41**, 327 (1990).

[3] G. Mahan, *Many-particle-physics*, 3rd ed. (Kluwer Academic/Plenum Publishers, New York and London, 2000).

[4] M. M. Parish, Phys. Rev. A **83**, 051603 (2011).

[5] M. Combescot and P. Nozières, J. Phys. **32**, 913 (1971).

[6] W. Ketterle and M. W. Zwierlein, arXiv preprint arXiv:0801.2500 (2008).

[7] R. Combescot, A. Recati, C. Lobo, and F. Chevy, Phys. Rev. Lett. **98**, 180402 (2007).

[8] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products* (Academic press, 2014).

[9] M. Punk, *Many-particle-physics with ultracold gases*,

Ph.D. thesis (2010).

[10] D. Pimenov, J. von Delft, L. Glazman, and M. Goldstein, Phys. Rev. B **96**, 155310 (2017).

[11] A. E. Ruckenstein and S. Schmitt-Rink, Phys. Rev. B **35**, 7551 (1987).

[12] T. Kopp, A. E. Ruckenstein, and S. Schmitt-Rink, Phys. Rev. B **42**, 6850 (1990).

[13] D. Pimenov, *Fermi-Edge polaritons with finite hole-mass*, Master's thesis, Ludwig Maximilians University Munich (2015).

[14] A. Rosch and T. Kopp, Phys. Rev. Lett. **75**, 1988 (1995).

[15] G. M. Bruun and P. Massignan, Phys. Rev. Lett. **105**, 020403 (2010)

5.4 Molecule-to-polaron transition

5.4.1 Phenomenology of the molecule-to-polaron transition

To keep the common thread of this thesis in sight, in this extended outlook I will present some preliminary thoughts on the molecule-polaron transition/crossover. First, let us recapitulate the conventional wisdom: it states that there are two well-defined separate phases, a "molecular" and a "polaronic" one. In the molecular phase, connected to the limit $E_b \gg \mu$, the ground state contains a bound state between the impurity and a majority particle, while the bound state is part of an excited state (or does not exist) in the polaronic phase ($\mu \gg E_b$). The quantity that distinguishes between these two phases is the quasi-particle weight Z of the lowest-in-energy feature of the *single-particle spectrum* $\mathcal{A}(\omega) \propto \mathrm{Im}\,\langle dd^\dagger \rangle$ [PDZ09]: Indeed, deep in the molecular limit, bound state formation results in a single hole in the majority Fermi sea. This hole can have an arbitrary energy up to μ, determining the spectral width of the ground state feature. There is no delta peak ($Z = 0$). By contrast, in the polaronic limit the particle is essentially free, with a large quasiparticle weight, possibly on top of a small incoherent background of width $\simeq \mu$. In the infinite mass case, the polaronic delta peak is replaced by a sharp power law, but the delta peak reemerges for finite masses as discussed in the previous sections. Sketches of these line-shapes are shown in Fig. 5.7(a),(b). Note that the overall weight of the molecule-hole feature of Fig. (a) is small in the indicated limit, since most of the weight is concentrated in the "repulsive polaron", see previous publication [PG18].
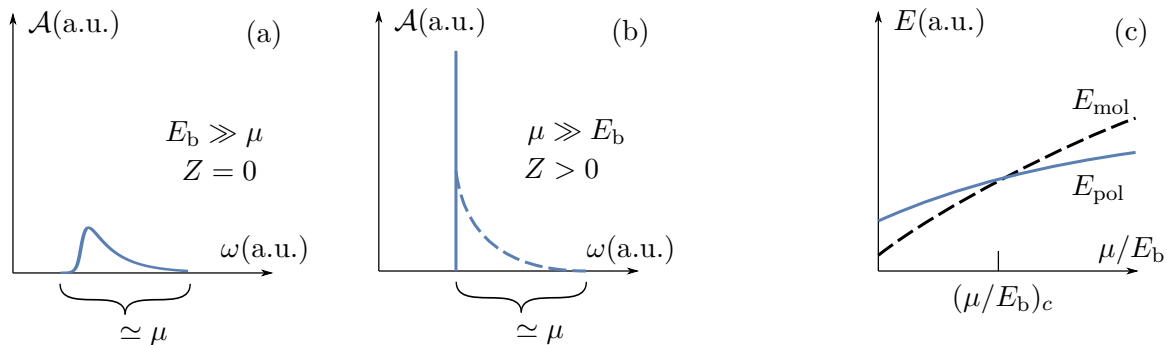

Figure 5.7 (a),(b): Spectral line shapes deep in the molecular and polaronic limit. The incoherent background shown in Fig. (b) (dashed) may or may not be divergent. (c) Typical crossing of the polaronic and molecular ground state energies as obtained from the variational ansaetze or Monte-Carlo. All dimensionful quantities are shown in arbitrary units (a.u.).

How does the transition between these two limits occur? The standard answer is that the transition is first order, i.e. Z jumps discontinuously at a critical value $(\mu/E_b)_c$. This assertion is supported by the following arguments: First of all, the ground state energies obtained from the minimization of polaron and molecule ansaetze or Monte-Carlo cross at a finite slope (Fig. 5.7(c)), at least for zero-momentum molecules (see below). As a result, the coefficients of the ground state wave-function in the variational approximation change discontinuosly at $(\mu/E_b)_c$, and so does Z – it is non-zero when determined from the polaronic ansatz, but vanishes for the molecular ansatz [PDZ09].

The quasi-particle weight extracted from Monte-Carlo does not jump but vanishes smoothly [VRVH13], but it does so as $\mu/E_b \to 0$, and not at a finite value μ/E_b. This can be taken as indication that the polaron still exists as a quasi-stable excited state for $E_b \gtrsim \mu$; this excited state should acquire a finite (but small) linewidth due to the decay into the molecule, which however cannot be properly resolved in Monte-Carlo. Again, the quasi-stability of

the polaron is typical for a first-order transition. An explanation for it can be found in a phase-space argument [PS08b, BM10]: If one views the molecule and polaron as rigid particles, the leading decay channel for a zero-momentum polaron into a zero-momentum molecule necessary involves two majority holes and a majority electron, see Fig. 5.8(a). This is required to place all scattered majority particles close to the Fermi surface, and to ensure energy and momentum conservation. The resulting condition for the momenta of the scattered majority particles, $|\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{k}_1| \simeq 0$ as shown in Fig. 5.8(a) severely restricts the phase space for this process. In effect, the estimated decay rate for the polaron is parametrically smaller than the energy separation to the molecule pole [BM10]. This means that the polaron remains sharp at the transition, and does not evolve into the molecule smoothly.
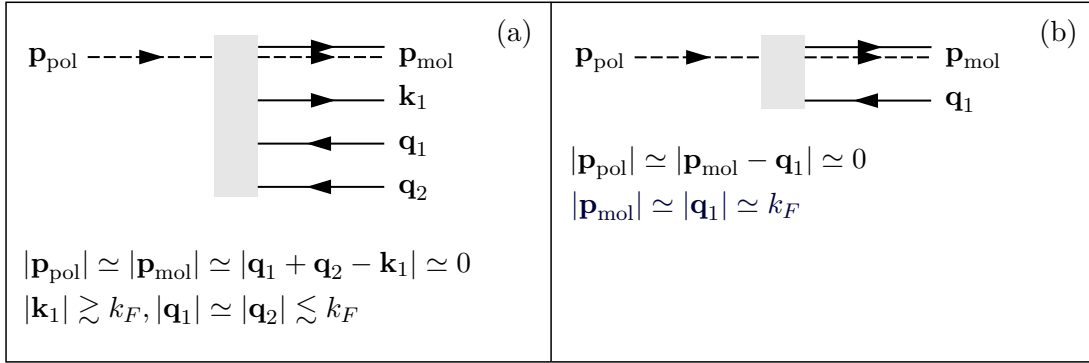
Figure 5.8 Leading order decay channels for rigid molecules and polarons. Full lines represent majority particles, the dashed line the polaron (dressed impurity), and the double line the molecule. The decay involves a zero-momentum molecule, Fig. (a), and a molecule with momentum $\simeq k_F$, Fig. (b).

Last, the first order nature of the transition is also supported by an RG computation which effectively incorporates similar scattering processes, see Sec. 5.4.3.

Although the argumentation for the first order nature is fairly sound, there is room for criticism: To begin with, we know that in the solvable infinite mass limit there certainly is no discontinuous transition, but the ground-state feature in the single-particle spectrum just keeps getting sharper as μ/E_b is increased [CN71, SKI$^+$18], and, in fact, all properties of the system are smooth as a function of the interaction strength [LL14, KM65]. Our findings in the paper above [PG18] do leave room for a similar continuous evolution for finite mass (see Fig. 2 in the paper), but we certainly cannot say for sure since our computation is controlled for $E_b \ll \mu$ or $\mu \ll E_b$ only. What is more, the arguments for the first-order transition implicitly assume that both polaron and molecule can be approximated as quasi-elementary particles, composed from the bare impurity/bare (vacuum) bound state plus a low number of particle-hole excitations. For $\mu \simeq E_b$ there is, however, no small parameter that guarantees this. In particular, the phase-space computation of Ref. [BM10] assumes that the polaron decays into a molecule corresponding to a well-defined pole in the *two-particle propagator*, with weight Z_M, and a quadratic dispersion $p^2/2m_{\mathrm{Mol}}$. But, first, the weight Z_M may vanish at the transition, and second, the dispersion may be minimal at a non-zero momentum.

The latter case is akin to the FFLO order discussed in Sec. 3.2, and in particular molecules with preferential momentum $|\mathbf{p}_{\mathrm{mol}}| \simeq k_F$ are of interest [MPH11, PL13]: As sketched in Fig. 5.8(b), the zero-momentum polaron can decay into such a k_F-molecule without the additional creation of an electron-hole pair, since the emitted majority hole is automatically close to the Fermi surface. Similar reasoning applies to the reversed process. As a result, the phase space for the decay of polaron/molecule is significantly increased as compared

to Fig. 5.8(a), leading to a continuous (un-)binding transition without quasi-stable excited states as in the first-order scenario.

To motivate the appearance of such an FFLO molecule, on can appeal to the variational ansaetze, for now casting aside the doubts about their range of validity. In fact, as shortly discussed below, inclusion of a finite molecule momentum has lead to a postulate by Edwards [Edw13] that the molecule-polaron transition is always just a smooth crossover.

5.4.2 Transition vs. crossover from the variational ansaetze

To set the stage for Edwards' argument, let us shortly recapitulate the variational treatment of Refs. [MPH11] (3D) and [PL13] (2D). These authors considered polaronic and molecular ansaetze with a finite momentum $\mathbf{p}$, and with up to one electron-hole pair. In their notation, these ansaetze read:

$$
\begin{aligned}
|P_1(\mathbf{p})\rangle &= d_{\mathbf{p}}^\dagger |N\rangle & (5.7) \\
|M_2(\mathbf{p})\rangle &= \sum_{\mathbf{k}} \phi_{\mathbf{k}} d_{\mathbf{p-k}}^\dagger c_{\mathbf{k}}^\dagger |N-1\rangle \\
|P_3(\mathbf{p})\rangle &= \alpha_0 d_{\mathbf{p}}^\dagger |N\rangle + \sum_{\mathbf{k,q}} \alpha_{\mathbf{kq}} d_{\mathbf{p+q-k}}^\dagger c_{\mathbf{k}}^\dagger c_{\mathbf{q}} |N\rangle \\
|M_4(\mathbf{p})\rangle &= \sum_{\mathbf{k}} \phi_{\mathbf{k}} d_{\mathbf{p-k}}^\dagger c_{\mathbf{k}}^\dagger |N-1\rangle + \sum_{\mathbf{kk'q}} \phi_{\mathbf{kk'q}} d_{\mathbf{p+q-k-k'}}^\dagger c_{\mathbf{k}}^\dagger c_{\mathbf{k'}}^\dagger c_{\mathbf{q}} |N-1\rangle \quad \dots \, .
\end{aligned}
$$

Here, $|P_i\rangle\,(|M_i\rangle)$ is a polaronic (molecular) ansatz with i-particles on top of the Fermi sea $|N\rangle\,(|N-1\rangle)$ consisting of $N(N-1)$ majority fermions. Sums over $\mathbf{k(q)}$ run over $|\mathbf{k}| > k_F$ ($|\mathbf{q}| < k_F$). To find the ground state energy E, one minimizes the expectation value of the Hamiltonian with respect to the coefficients α, ϕ and the modulus of momentum $|\mathbf{p}|$, subject to a normalization constraint. Focusing on the 2D case, one can show that the polaronic (molecular) ansaetze correctly reproduce the ground-state energy in the perturbatively controlled limits $\mu \gg E_{\mathrm{b}}$ ($E_{\mathrm{b}} \gg \mu$). Deep in these limits, the obtained variational momentum is zero. However, for $\mu \simeq E_{\mathrm{b}}$ and a fairly light impurity, a region with an FFLO phase is obtained, while always $\mathbf{p} = 0$ for the optimal polaronic wave-functions. The phase diagram of Ref. [PL13] is reprinted in Fig. 5.9.
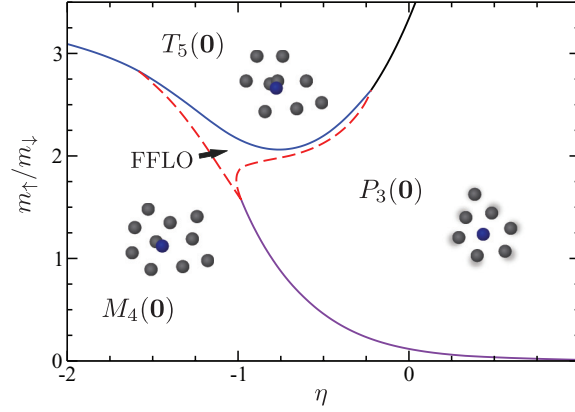
Figure 5.9 Phase diagram from the one particle-hole ansaetze, reprinted with permission from Ref. [PL13]. Copyright (2013) by the American Physical Society. Used notation is $\eta = \ln(k_F a_{2D}) \simeq \ln(\mu/E_b)$, with a_{2D} the $2D$ scattering length, and $m_\downarrow \hat{=} M, m_\uparrow \hat{=} m$. $T_5(\mathbf{0})$ corresponds to a "trimer ansatz" (two majority particles bound to the impurity) not discussed here. In the region indicated as "FFLO", a molecule with non-zero momentum $\mathbf{p}$ is obtained, with $|\mathbf{p}| = 0$ at the left and $|\mathbf{p}| = k_F$ at the right dashed region boundary.

To summarize, the conclusion of Ref. [PL13] is that a continuous transition is possible for small impurity masses, but for impurities with $M \gtrsim 0.7m$ the transition is still first order.

This conclusion was subsequently questioned by Edwards [Edw13], by virtue of the following argument: Let's e.g. take the wave-function $|M_4(\mathbf{p})\rangle$ and set $|\mathbf{p}| = k_F$ as well as

$$\phi_{\mathbf{k}} = \delta_{\mathbf{k},\mathbf{p}} \cdot \alpha_0, \quad \phi_{\mathbf{k}\mathbf{k'}\mathbf{q}} = \frac{1}{2}\left(\alpha_{\mathbf{k'}\mathbf{q}}\delta_{\mathbf{k},\mathbf{p}} - \alpha_{\mathbf{k}\mathbf{q}} \cdot \delta_{\mathbf{k'},\mathbf{p}}\right). \tag{5.8}$$

This definition accounts for the required $\mathbf{k} \leftrightarrow \mathbf{k'}$ antisymmetry of the coefficient function. Then $|M_4\rangle$ reduces to

$$|M_4(\mathbf{p})\rangle \to |\tilde{P}_3\rangle = \alpha_0 d_0^\dagger c_{\mathbf{p}}^\dagger |N-1\rangle + \sum_{\mathbf{k},\mathbf{q}} \alpha_{\mathbf{k}\mathbf{q}} d_{\mathbf{q}-\mathbf{k}}^\dagger c_{\mathbf{k}}^\dagger c_{\mathbf{q}} c_{\mathbf{p}}^\dagger |N-1\rangle . \tag{5.9}$$

If we identify $c_{\mathbf{p}}^\dagger |N-1\rangle \hat{=} |N\rangle$, we can conclude $|\tilde{P}_3\rangle = |P_3(0)\rangle$, and therefore $E(M_4) \leq E(P_3)$ by the variational principle. This means that there is no need to introduce a polaronic ansatz, since the polaronic limit is correctly reproduced by the molecular wave-functions with the sharply peaked "polaronic" coefficients of Eq. (5.8). As a result, Edwards claims that the evolution from molecule to polaron is simply a smooth crossover, for any value of the impurity mass, and the phase diagram of Fig. 5.9 is incorrect.

While the observation of Edwards that a single ansatz is in fact sufficient is certainly interesting, I believe his conclusion is too drastic: One the one hand, the identification $c_{\mathbf{p}}^\dagger |N-1\rangle \hat{=} |N\rangle$ may be questioned, although it should be passable in the thermodynamic limit as far as ground state energies are concerned. But more importantly, usage of a single variational ansatz still allows for both a continuous *and* a first order transition: the question is simply how the minimizing coefficients of the ansatz behave, i.e. whether they evolve continuously or jump as μ/E_b is varied. To illustrate this, let us look at the simplest ansaetze without electron-hole pair, $|P_1(\mathbf{p})\rangle, |M_2(\mathbf{p})\rangle$: If there is no energy offset for the impurity, $E_0 = 0$ (recall the impurity Hamiltonian (5.1)), the ground state energy obtained from $|P_1(\mathbf{p})\rangle$ is simply constant, $E(P_1) = 0$. By contrast, the energy $E(M_2)$ can be determined

from the equation [PL13]

$$\frac{d\langle M_2(\mathbf{p})|H - E|M_2(\mathbf{p})\rangle}{d\bar{\phi}_{\mathbf{k}}} = \phi_{\mathbf{k}}\left(\epsilon_{\mathbf{k}} - \mu + E_{\mathbf{k}-\mathbf{p}} - E\right) + \frac{V_0}{\mathcal{S}}\sum_{\mathbf{k}'}\phi_{\mathbf{k}'} \overset{!}{=} 0. \qquad (5.10)$$

Here and in the following, we take the limits $V_0 \to 0$, $\Lambda(\text{UV cutoff}) \to \infty$ in such a way that the identity

$$-\frac{1}{V_0} = \frac{1}{\mathcal{S}}\sum_{0<|\mathbf{k}|<\Lambda}\frac{1}{E_{\mathrm{b}} + \epsilon_{\mathbf{k}} + E_{\mathbf{k}}} \qquad (5.11)$$

yields a finite vacuum binding energy E_{b}.

The usual way to treat Eq. (5.10) [Pun10] is to solve for $\phi_{\mathbf{k}}$, sum over $\mathbf{k}$, and divide by $\sum_{\mathbf{k}}\phi_{\mathbf{k}}$. In this way one finds

$$\frac{1}{\mathcal{S}}\sum_{\mathbf{k}}\frac{1}{\epsilon_{\mathbf{k}} - \mu + E_{\mathbf{p}-\mathbf{k}} - E} = \frac{1}{\mathcal{S}}\sum_{\mathbf{k}>0}\frac{1}{E_{\mathrm{b}} + \epsilon_{\mathbf{k}} + E_{\mathbf{k}}}. \qquad (5.12)$$

In the continuum limit, the sums are transformed into integrals, and one can solve for $E = E(M_2)$, which is also minimized w.r.t. $\mathbf{p}$. For equal masses $m = M$, $E(M_2)$ is shown in Fig. 5.10 (blue curve), and the corresponding optimal momentum in Fig. 5.10(b).

To derive (5.12), one implicitly excludes $\epsilon_{\mathbf{k}} - \mu + E_{\mathbf{p}-\mathbf{k}} - E = 0$. However, this combination can lead to another solution, the "polaronic one" ,

$$\phi_{\mathbf{k}} = \delta_{\mathbf{k},\mathbf{p}}, \quad |\mathbf{p}| = k_F, \quad E = 0, \qquad (5.13)$$

which solves Eq. (5.10) in the limit $V_0 \to 0$. To put this differently: the standard minimization of Eq. (5.12) is restriced to the domain of continuous functions $\phi_{\mathbf{k}}$, and thus the discontinous solution (5.13) is missed. Therefore, the energy E determined from (5.12) is no longer the true minimum as soon as $E > 0$. The situation is illustrated in Fig. 5.10(a).
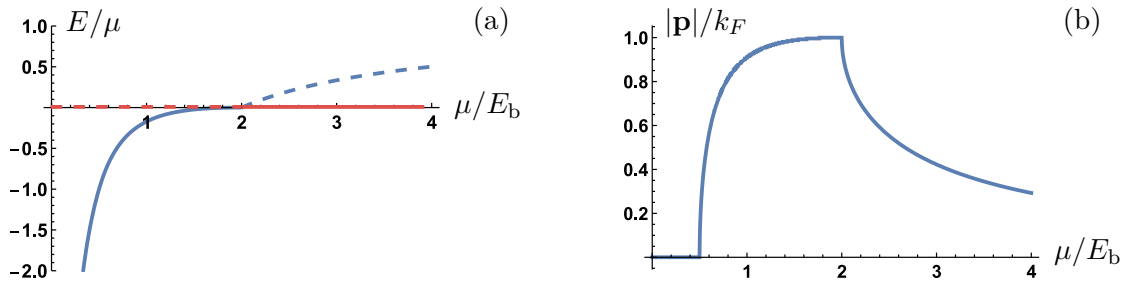

Figure 5.10 (a) Two different solutions to Eq. (5.10). The blue line is determined from Eq. (5.12), and the red line corresponds to Eq. (5.13). The full line indicates the correct ground state energy $E(M_2)$. (b) Optimal momentum corresponding to the blue line in Fig. (a).

Thus, we see that the correct minimization of the molecular wave-function $|M_2(\mathbf{p})\rangle$ captures both polaronic (sharply peaked coefficients as in Eq. (5.13)) and conventional molecular behaviour. Now, to decribe the transition, one must trace how the coefficients of the wavefunction evolve as μ/E_{b} is increased. By simply solving (5.10) for $\phi_{\mathbf{k}}$ and enforcing normalization of the wavefunction, one can easily show

$$\text{As } E \nearrow 0, |\mathbf{p}| \nearrow k_F: \quad \phi_{\mathbf{p}} \to \frac{\mu^{3/4}}{|E|^{3/4}\cdot\sqrt{N}}, \qquad \phi_{\mathbf{k}} \to \frac{|E|^{1/4}}{\mu^{1/4}\sqrt{N}}, \quad \mathbf{k} \neq \mathbf{p} . \qquad (5.14)$$

I.e., the coefficients evolve continuously into $\delta_{\mathbf{k},\mathbf{p}}$ as $N \to \infty$ and $E \nearrow 0$ (or $\mu/E_{\mathrm{b}} \nearrow 2$) in Fig. 5.10(a), indicating a continuous transition (within this crude approximation). Note, however, that this is still distinct from a smooth crossover, which would correspond to an evolution of $\phi_{\mathbf{k}}$ into $\delta_{\mathbf{k},\mathbf{p}}$ in the limit $\mu/E_{\mathrm{b}} \to \infty$ – furthermore, the ground-state curve in Fig. 5.10 is not smooth at $\mu/E_{\mathrm{b}} = 2$ (the second derivative of the blue curve does not vanish).

It should be noted that, while the coefficients evolve continuously, strictly speaking the quasiparticle weight is zero for $E < 0$: E.g., if we follow the weight estimation of [PDZ09] which associates the quasiparticle weight with the jump of the Fermi distribution function, we find

$$Z \le \langle M_2(\mathbf{p})|d_0^\dagger d_0|M_2(\mathbf{p})\rangle \overset{(5.7)}{=} |\phi_{\mathbf{p}}|^2 \; . \tag{5.15}$$

As seen in Eq. (5.14), $\phi_{\mathbf{p}}$ generically scales as $1/\sqrt{N}$, but with a divergent prefactor as $E \nearrow 0$. Thus $Z = 0$ for $E < 0$, $Z = 1$ for $E > 0$. Still, I believe the continuous evolution of the coefficients makes it questionable whether such a jump in the quasiparticle weight can be observed in reality.

Let us return to the more accurate ansaetze with one electron-hole pair, $|P_3(\mathbf{p})\rangle, |M_4(\mathbf{p})\rangle$. While the determination of the energies $E(P_3)$ and $E(M_4)$ can be inferred from Ref. [PL13], a correct evaluation of the minimizing coeffients analogous to Eq. (5.14) is not available yet. Our expectation, which would reconcile the basic observation of Edwards and the phase diagram of Fig. 5.9, is as follows: As μ/E_{b} increased, the optimal coefficients $\phi_{\mathbf{k}}, \phi_{\mathbf{kk}'}$ of $|M_4(\mathbf{p})\rangle$ always evolve into the sharp, "polaronic" limit given in Eq. (5.8). But his evolution is continuous similar to Eq. (5.14) only for light impurities, when the FFLO-region of Fig. 5.9 is crossed. For heavy impurities, a correct optimization of the coefficients should show a discontinuous jump towards Eq. (5.8) at the transition point, with the optimal momentum jumping from $\mathbf{p} = 0$ to $|\mathbf{p}| = k_F$.

To summarize the take-away message of this section: If one believes the truncated variational ansaetze, retaining a single momentum-dependent ansatz (say, $|M_4(\mathbf{p})\rangle$) is indeed sufficient to exhaust the full phase diagram as found by Edwards [Edw13]; however, this potentially allows for both continouus and first-order transitions, since the optimal coefficients of that ansatz may be discontinuous as a function of μ/E_{b}, and thus the phase diagram of Fig. 5.9 should be correct within the variational approximation.

5.4.3 RG approach to the molecule-to-polaron transition

Can we describe the molecule-to-polaron transition within Landau-Ginzburg theory? At present, the answer is no: First of all, in the thermodynamic limit for the majority fermions only the impurity degrees of freedom can become critical at the transition, and there is no symmetry breaking of the full system that can be captured by an order-parameter theory. In the above, we relied on the single-particle weight Z to distinguish the phases, but it is not clear how to utilize Z in a LGW functional; Z is an order-parameter only in a colloquial sense.

Alternatively, we can attempt a renormalization group treatment. Of course, if the transition is indeed first order, this may seem as a somewhat futile endeavour, since there is no room for universal critical behaviour. But, as described above, a continuous transition is also imaginable, and ultimately the RG is the best tool to make the judgement. The starting point for the RG will be the impurity Hamiltonian (5.13). This has the downside that the single-impurity limit is not incorporated directly. As an alternative, one could rephrase the impurity fields as impurity position and momentum operators, followed by a Lee-Low-Pines rotation into a moving frame [RK95, KL17], which effectively allows to eliminate the impurity.

However, the resulting theory for the majority fermions has a fairly complicated interaction term, so we will stick to Eq. (5.1) for the time being.

The short RG analysis presented here will be a slightly reformulated version of the FRG approach by Schmidt and Enss (SE) [SE11], which is the only detailed RG treatment of the molecule-to-polaron transition to date (some related results are also found in [GS08]). SE focused on the computation of spectra; here, we will look for critical phenomena, indicating possible extensions of [SE11] along the way.

SE work with the action corresponding to (5.13) with majority (minority) fields $c(k), \bar{c}(k)$ $(d(k), \bar{d}(k))$. As the first step, they perform a Hubbard-Stratonovich decoupling in the pairing channel $\phi(k) \propto \langle \sum_p c(k+p)d(-p) \rangle$ that makes the molecule explicit. The result is the "two-channel model"

$$
\mathcal{S} = \int_k \bar{c}(k)(-i\omega + \mathbf{k}^2/2m - \mu)c(k) + \int_k \bar{d}(k)(-i\omega + \mathbf{k}^2/2M)d(k) \tag{5.16}
$$
$$
+ \int_k \bar{\phi}(k)G_\phi^{-1}(\omega, \mathbf{k})\phi(k) + \lambda \int_{k,p} (\bar{\phi}(k)c(k+p)d(-p) + \phi(k)\bar{d}(-p)\bar{c}(k+p)),
$$

where we left the molecule propagator unspecified (it will be constant in the UV, but a frequency and momentum dependence will arise at small energies) and introduced an independent Yukawa coupling λ, check our discussion in Sec. 3.1.3. It is tempting to go one step further and integrate out all fermions in favour of a purely bosonic theory, interpreting $\langle \phi \rangle$ as order parameter for the molecular phase. However, with our previous discussion of the Hertz-Millis approach to metallic criticality in mind (Sec. 3.1.4), this step of integrating out gapless degrees of freedom seems dangerous, and we will retain all three particle species.

Now, when running the RG, the following basic problem arises: low-energy majority fermions c have momenta $\simeq k_F$, and thus we must scale towards the Fermi surface. By contrast, for the minority particles d (and possibly the molecule ϕ, see below) we must scale towards $\mathbf{k} = 0$. In the previously discussed RG approaches to Fermions of Shankar- (Sec. 3.1.2) and hot-spot type (Sec. 3.1.3), we could get rid of the Fermi momentum by expanding the dispersion around it, the assumption being that k_F is the largest scale in the problem. But now we want to describe a transition with a competing scale $E_b \simeq \mu$, so clearly such an expansion is no longer legitimate!

The FRG approach (see Sec. 4.1.3) of SE handles this difficulty very gracefully by choosing different regulators for majority, impurity, and molecule. The regularized flowing Green's functions at scale Λ are proportional to

$$
G_{\phi,\Lambda}(\omega, \mathbf{k}) \propto \theta(|\mathbf{k}| - \Lambda), \quad G_{d,\Lambda}(\omega, \mathbf{k}) \propto \theta(|\mathbf{k}| - \Lambda), \quad G_{c,\Lambda}(\omega, \mathbf{k}) \propto \theta(|\mathbf{k}^2 - k_F^2| - \Lambda^2) . \tag{5.17}
$$

Note the first proportionality: it implies that the low-energy molecule (as $\Lambda \to 0$) has its dispersion minimum at zero momentum. As we discussed in the previous sections, this typically results in a first order transition, and this is indeed found by SE. By contrast, to obtain a continuous transitions, we should endow the molecule with a regulator $\theta(|\mathbf{k}^2 - k_F^2| - \Lambda^2)$ as well.

To perform the FRG, we must prescribe a truncation. SE neglect higher-order vertices beyond the ones appearing in (5.16), but allow for the most general form of impurity/molecule propagators G_d, G_ϕ, which is e.g. necessary to reproduce the correct vacuum limit for the molecule. As seen from the diagrammatic representation of the FRG flow equations [SE11], this truncation is akin to a *self-consistent* T-matrix approach (see also Sec. 5.3.1). Note that the majority particles can only pick up non-trivial renormalizations of order $\mathcal{O}(1/N)$ from a single impurity, to be neglected here.

To keep things transparent, here we will limit ourselves to a gradient expansion also discussed by SE; this cannot reproduce correct spectra, but the qualitative features of the phase transition agree with the full treatment of SE. For simplicity, let us also focus on equal masses $m = M \equiv 1/2$, and thus choose

$$G_{d,\Lambda}(\omega, \mathbf{k}) = \theta(|\mathbf{k}| - \Lambda) \left(Z_d(-i\omega + \mathbf{k}^2) + r_d \right)^{-1}, \tag{5.18}$$

$$G_{\phi,\Lambda}(\omega, \mathbf{k}) = \theta(|\mathbf{k}| - \Lambda) \left(Z_\phi(-i\omega + \mathbf{k}^2/2) + r_\phi \right)^{-1}, \tag{5.19}$$

where the the molecule has twice the impurity mass. The flowing parameters are the wave-function renormalization functions $Z_{\phi,d}$ and the "gaps" $r_{\phi,d}$. Schematically, the polaronic (molecular) phase will be identified as $r_d < r_\phi (r_\phi < r_d)$, with more details specified below. In this truncation, the relevant flow equations are readily derived from the Wetterich equation [SE11], since the required integrals collapse due to the sharp cutoff functions. Here, we generalize the 3D flow-equations of SE to d spatial dimensions, and introduce standard dimensionless quantities (see, e.g., [KBS10]) as

$$\tilde{r}_\alpha = \frac{r_\alpha}{Z_\alpha \Lambda^2}, \quad \eta_\alpha = \frac{d\log(Z_\alpha)}{dl} = \frac{1}{Z_\alpha}\frac{dZ_\alpha}{dl}, \quad \tilde{\lambda} = \frac{\lambda}{\Lambda^{\frac{4-d}{2}}\sqrt{Z_d Z_\phi}}, \quad \tilde{\mu} = \frac{\mu}{\Lambda^2}, \tag{5.20}$$

where $\alpha = (d, \phi)$, $\Lambda = \Lambda_0 \exp(-l)$, and η_α are the anomalous dimensions. In these terms, the flow equations read

$$\frac{d\tilde{r}_d}{dl} = (2 - \eta_d)\tilde{r}_d - 2c_d\tilde{\lambda}^2\theta(\tilde{\mu} - 2)\left(\frac{(\sqrt{\tilde{\mu}-1})^{d-2}}{1 + \tilde{\mu} + 2\tilde{r}_\phi} + \frac{1}{-1 + 2\tilde{\mu} + 2\tilde{r}_\phi}\right) \tag{5.21}$$

$$\frac{d\tilde{r}_\phi}{dl} = (2 - \eta_\phi)\tilde{r}_\phi - c_d\tilde{\lambda}^2\frac{(\sqrt{\tilde{\mu}+1})^{d-2}}{2 + \tilde{\mu} + \tilde{r}_d},$$

$$\eta_d = 4c_d\tilde{\lambda}^2\theta(\tilde{\mu} - 2)\left(\frac{(\sqrt{\tilde{\mu}-1})^{d-2}}{(1 + \tilde{\mu} + 2\tilde{r}_\phi)^2} + \frac{1}{(-1 + 2\tilde{\mu} + 2\tilde{r}_\phi)^2}\right), \quad \eta_\phi = c_d\tilde{\lambda}^2\frac{(\sqrt{\tilde{\mu}+1})^{d-2}}{(2 + \tilde{\mu} + \tilde{r}_d)^2}$$

$$\frac{d\tilde{\lambda}}{dl} = \tilde{\lambda}\left(\frac{4-d}{2} - \frac{\eta_d}{2} - \frac{\eta_\phi}{2}\right), \quad \frac{d\tilde{\mu}}{dl} = 2\tilde{\mu},$$

with c_d a positive numerical coefficient. Let us first have a look at betafunctions for $\tilde{\lambda}$ and $\tilde{\mu}$ in the last line. In the given truncation, this flow arises solely for the dimensionless quantities due to the change of the scale Λ. The simple flow equation for $\tilde{\mu}$ has a fixed point $\tilde{\mu}^\star = 0$ (vacuum limit), which also implies $\eta_d = 0$ and $d_l\tilde{r}_d = 2\tilde{r}_d$. If we choose $\tilde{r}_d = 0$ (which seems natural in the UV), the flow equation for $\tilde{\lambda}$ transforms to $d_l\tilde{\lambda} = \tilde{\lambda}((4-d)/2 - \tilde{\lambda}^2)$ modulo trivial rescaling, which has a *stable* fixed point $\lambda^\star \propto \sqrt{4-d}$. This is nothing but the vacuum fixed point of the two-component Fermi gas obtained by Nikolić and Sachdev [NS07a] via expansion in $\epsilon = (4 - d)$. Of course, it is not surprising that the FRG reproduces the vacuum limit, since the only relevant diagrams for $\mu = 0$ (ladder diagrams) are correctly incorporated.

So what about the finite-density system, $\tilde{\mu} > 0$? Obviously, $\tilde{\mu}$ will flow to infinity for any non-zero starting value. Since in (5.21) the equations for the anomalous dimensions η_α scale as $(\tilde{\mu})^{-2}$, it follows that in the IR $\eta_\alpha \to 0$ unless there is some extreme finetuning of both the couplings $\tilde{r}_d, \tilde{r}_\phi$. In consequence, the betafunction $d_l\tilde{\lambda}^2$ simply becomes $d_l\tilde{\lambda} = \tilde{\lambda}(4-d)/2$, which wipes out the $\tilde{\mu} = 0$ fixed point, and $\tilde{\lambda}$ flows to strong coupling for all $d < 4$. Furthermore, one can simply show that the interaction-dependent contributions in the flow of $\tilde{r}_\alpha$ scale as $\mathcal{O}(1)$ in the IR, which is subleading as compared to the exponentially diverging engineering parts, and thus the equations for $\tilde{r}_{\phi,d}$ essentially decouple. We therefore see that there are always

strong runaway flows unless we fine-tune at least two relevant parameters, which is indicative of a first-order transition. Of course, this is not suprising: The truncated RG effectively incorporates scattering processes similar to the ones arising in the ($\mathbf{p} = 0$) variational ansaetze (albeit in self-consistent fashion), and the basic prediction of both methods agrees.

For a quantitative understanding of the transition, we must discuss two major points glossed over so far: How to guarantee the single-impurity limit and precisely distinguish the two phases (polaronic and molecular) in the FRG formalism? SE give the following prescription: by choosing appropriate boundary conditions, that is, finetuning $\tilde{r}_d(0)$, one has to make sure that in the end of the RG flow, $l \to \infty$, only one mode (polaron or molecule) is occupied, while the other is empty. The filled mode has $\tilde{r}_\alpha(l \to \infty) \to \mathcal{O}(1)$, corresponding to $r_\alpha \to 0$. This is right inbetween the macroscopic occupation $r_\alpha < 0$ and the empty state $r_\alpha > 0$. By contrast, the other mode should fulfill $r_{\bar{\alpha}} > 0$. All other boundary conditions are specified by the dimensionless coupling constant ak_F (appropriately generalized to d dimensions) and the fact that the two-channel model (5.16) should map onto the single-channel model (5.1) in the UV. In this way, SE find a molecule-to-polaron transition at $ak_F \simeq 1.04$, in reasonable agreement with their full numerical treatment and with diagrammatic Monte-Carlo.

When comparing to the diagrammatic computations of the previous sections, I am not certain that such a fine-tuning of $\tilde{r}_d(0)$ is actually necessary: recall from Sec. 5.1.2 that a sufficient condition for ensuring the single-particle limit was to use strictly retarded impurity propagators, as well as retarded molecule propagators in the two-channel formulation. This retardation condition was implicitly used in deriving Eq. (5.21) – there would be additional contributions from advanced impurity particles, arising from the other way of closing contours in frequency integrals. If we exclude these "advanced contributions", the single-impurity limit is still implemented in Eq. (5.21) also for $\tilde{r}_d, \tilde{r}_\phi < 0$. From a different perspective, we could supplement the impurity and molecule dispersions by large energy offset, $\mathbf{k}^2 \to \mathbf{k}^2 + E_0$ as done in Sec. 5.1.1. In the limit $E_0 \to \infty$ we can then derive the flow for r_α as $d_l r_\alpha = d_l G_\alpha(\omega + i0^+ = E_0 + i0^+, \mathbf{k} = 0)$ after analytical continuation. In this way, the single-particle limit is manifest without additional finetuning, since contributions with n impurities/molecules involve far-detuned energies $\simeq nE_0$.

If we believe that the single-particle limit is automatically implemented, we are free to specify any value of $\tilde{r}_d(0)$, and may e.g. choose $\tilde{r}_d(0) = 0$ to conform with the vacuum limit. Of course, this does not change the basic conclusion of a first-order transition in any way. To illustrate the transition, some plots of the flowing gap terms $\tilde{r}_\alpha$ for $d = 3$ are shown in Fig. 5.11, renormalized by the exponentially diverging chemical potential $\tilde{\mu}$. In these plots, we tune through the transition by increasing $\tilde{r}_\phi(0)$. By comparing the converged molecule propagator to the exactly known static molecule propagator in vacuum, we can also extract a value for the dimensionless interaction parameter ak_F [SE11] as denoted in the figure. The critical value $ak_F \simeq 2$ obtained this way compares less favourable to the Monte-Carlo data than the finetuning applied by SE, but on the other hand we can hardly expect our rough truncation to be numerically precise.

To summarize, at the present stage the FRG does seem to predict a first order transition, signaled by strong runaway flows. At the same time, the divergence of the Yukawa coupling $\tilde{\lambda}$ renders the RG uncontrolled.

Can we improve on this? The engineering part of the beta-function $d_l\tilde{\lambda}$ in Eq. (5.21) does suggest an ϵ-expansion; alas, for $\tilde{\mu} > 0$ there is no meaningful fixed point $\tilde{\lambda}^\star$ which can be tuned to small values.[5] To find such a fixed point, and thus get a perturbative hold of the molecule-polaron transition, we can attempt the following (this is currently work in progress):

---

[5] ϵ-expansion results for the related problem of an imbalanced Fermi gas have been obtained by Nishida and Son [NS06, NS07b], but they work deep in the superfluid phase which is not permittable here.
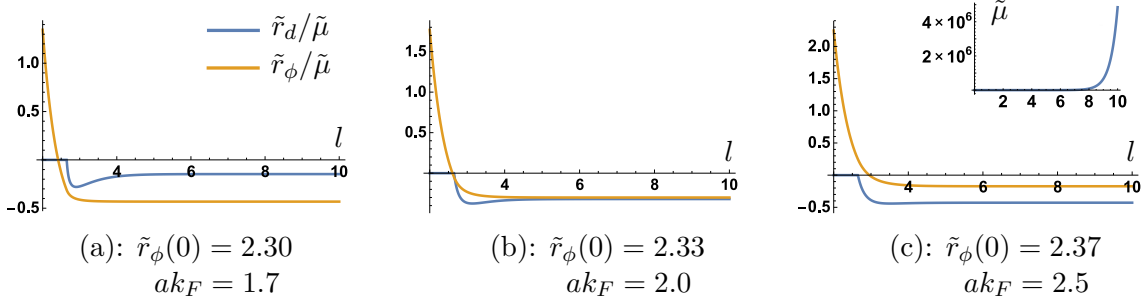
(a): $\tilde{r}_\phi(0) = 2.30$
$ak_F = 1.7$

(b): $\tilde{r}_\phi(0) = 2.33$
$ak_F = 2.0$

(c): $\tilde{r}_\phi(0) = 2.37$
$ak_F = 2.5$

Figure 5.11 Flow of the gap terms for the molecule $\tilde{r}_\phi$ and polaron $\tilde{r}_d$ as a function of the logarithmic RG time l. The gaps are regularized by the exponentiallly diverging chemical potential, shown in the inset to Fig. (c). (a) Molecular phase (b) Transition (c) Polaronic phase. Used parameters are $\tilde{r}_d(0) = 0, \tilde{\mu}(0) = 0.01, \tilde{\lambda}(0) = 10$; the large starting value for the Yukawa coupling is chosen in order to correctly reproduce the original single-channel model in the UV [Pun10].

as we have seen in the previous sections, a continuous (un-)binding transition should involve a molecule with a finite momentum $|\mathbf{p}| = k_F$. Thus, we can modify the molecule propagator of Eq. (5.19) as

$$G_{\phi,\Lambda}(\omega, \mathbf{k}) = \theta(|\mathbf{k}^2 - p_0^2| - \Lambda^2) \left(Z_\phi(-\omega + (|\mathbf{k}| - p_0)^2/(m + M) + r_\phi \right)^{-1}. \qquad (5.22)$$

I.e., the minimum of the molecule dispersion is located at a momentum p_0, which should be allowed to flow. In addition, we have reinstalled the masses m, M with the phase diagram of Fig. 5.9 in mind.

Indeed, first results for a fixed value of $p_0 = k_F$ do indicate the reemergence of nontrivial fixed point $\tilde{\lambda}^\star$, but more work needs to be done. In further pursuing these ideas, at least for certain mass ratios m/M we hope to arrive at a universal description of the molecule-polaron transition in the future.

6 Conclusion and Outlook

In this thesis, we have covered three fairly diverse fields of physics, held together by a somewhat loose common thread: the violation of the Landau-Ginzburg-Wilson paradigm of critical phenomena. In all cases, a simple order parameter description was insufficient to describe the phase transitions, since it missed crucial low-energy degrees of freedom. For the quantum critical metals, these are the fermionic excitations. To incorporate them, we performed an RG analysis of a Yukawa-like theory, which we got under control by working close to the upper critical dimension. Interesting non-Fermi liquid physics ensued. For the insulating spin system, we saw that the phase transition is effected by fractionalized degrees of freedom, which are topological in nature, and become important (gapless) at the critical point only. Finally, for the polaron problem, again there is an abundance of low-energy electron-hole pairs, leading to interesting orthogonality power laws – in fact, similar to the non-Fermi liquids. A fully conclusive picture of the molecule-to-polaron transition has not yet been reached, but it is clear that an order parameter description does not suffice.

Where to go next? In the problem of $SU(3)$ deconfined criticality, there are many open questions, such as the precise role of the constraint softening or the detailed classification of the vortex defects, but let us rather turn to the remaining two subjects which we are currently working on: Concerning metallic criticality, it is worthwhile to consider the $2k_F$ problem in a commensurate configuration, extending the existing RPA analysis [BCP$^+$12]. This problem also seems suitable for an exact non-perturbative solution, similar to the exciting exact solution of the spin-density-wave problem [SLL17]: the key idea is not to start from a Gaussian scaling ansatz, but rather from an "interaction-driven" one, where the Yukawa term is marginal from the outset. In this scaling, certain terms in the bare action are irrelevant by power counting and flow to the zero in the IR. But as it turns out, they can be utilized as small parameters which suppress most Feynman diagrams, and in this sense the interaction-driven scaling becomes exact in the IR.

Furthermore, the precise role of competing instabilities at the metallic critical points is still largely unexplored. In the usual approach, which we also followed in Sec. 3.2, one defines a "primary" instability with an associated critical point, and perturbatively computes anomalous dimensions of other instabilities at this point. To improve on this, one could consider hot-spot theories with multiple equal-right instability channels. A suitable technical tool for such an analysis seems to be the recently developed multiloop FRG, which resums parquet-like contributions [KvD18].

Concerning polarons, the analysis of the molecule-to-polaron transition described in Sec. 5.4 needs to be completed. The ultimate goal should be to understand whether the first order transition for heavy impurities and second order transition for light impurities is indeed robust, or simply an artefact of the low-order variational ansaetze.

Finally, an interesting extension of the standard polaron problem are impurities coupled to topological insulators, as e.g. recently studied in Ref. [CGGMB19]: if the interaction with the majority particles is large enough compared to the insulator gap, the impurity inherits the topological properties of the majority particles. In the spirit of chapter Sec. 5, one could review this problem for infinite mass impurities. This could raise the exciting prospect of obtaining orthogonality power laws determined from topological quantum numbers.

Bibliography

[ACS03]   A. Abanov, A. V. Chubukov, and J. Schmalian, *Quantum-critical theory of the spin-fermion model and its application to cuprates: Normal state analysis*, Advances in Physics **52** (2003), no. 3, 119–218. See pages: 10, 11, and 14

[AIM95]   B. L. Altshuler, L. B. Ioffe, and A. J. Millis, *Critical behavior of the t=0 $2k_f$ density-wave phase transition in a two-dimensional fermi liquid*, Phys. Rev. B **52** (1995), 5563–5572. See pages: 11 and 43

[And67]   P. W. Anderson, *Infrared catastrophe in fermi gases with local scattering potentials*, Phys. Rev. Lett. **18** (1967), 1049–1051. See page: 86

[Asl69]   L. G. Aslamazov, *Influence of impurities on the existence of an inhomogeneous state in a ferromagnetic superconductor*, Sov. Phys. JETP **28** (1969), no. 4, 773–775. See page: 22

[Aue94]   A. Auerbach, *Interacting electrons and quantum magnetism*, 1 ed., Springer, New York, 1994. See pages: 59 and 60

[Bar13]   L. Bartosch, *Corrections to scaling in the critical theory of deconfined criticality*, Phys. Rev. B **88** (2013), no. 19, 195140. See pages: 59, 65, and 66

[BCL+12]   B. Bauer, P. Corboz, A. M. Läuchli, L. Messio, K. Penc, M. Troyer, and F. Mila, *Three-sublattice order in the su (3) heisenberg model on the square and triangular lattice*, Phys. Rev. B **85** (2012), no. 12, 125116. See page: 67

[BCP+12]   D. Bergeron, D. Chowdhury, M. Punk, S. Sachdev, and A.-M. S. Tremblay, *Breakdown of fermi liquid behavior at the $(\pi, \pi) = 2k_F$ spin-density wave quantum-critical point: The case of electron-doped cuprates*, Phys. Rev. B **86** (2012), 155123. See page: 143

[BM10]   G. Bruun and P. Massignan, *Decay of polarons and molecules in a strongly polarized fermi gas*, Phys. Rev. Lett. **105** (2010), no. 2, 020403. See page: 134

[CG08]   R. Combescot and S. Giraud, *Normal state of highly polarized fermi gases: full many-body treatment*, Phys. Rev. Lett. **101** (2008), no. 5, 050404. See page: 116

[CGGMB19]   A. Camacho-Guardian, N. Goldman, P. Massignan, and G. M. Bruun, *Dropping an impurity into a chern insulator: A polaron view on topological matter*, Phys. Rev. B **99** (2019), 081105. See page: 143

[CGJT10]   C. Chin, R. Grimm, P. Julienne, and E. Tiesinga, *Feshbach resonances in ultracold gases*, Rev. Mod. Phys. **82** (2010), 1225–1286. See page: 84

[CGL10]   R. Combescot, S. Giraud, and X. Leyronas, *Analytical theory of the dressed bound state in highly polarized fermi gases*, Europhys. Lett. **88** (2010), no. 6, 60007. See page: 116

[Che06]   F. Chevy, *Universal phase diagram of a strongly interacting fermi gas with unbalanced spin populations*, Phys. Rev. A **74** (2006), no. 6, 063628. See page: 116

[CJL+16]   M. Cetina, M. Jag, R. S. Lous, I. Fritsche, J. T. M. Walraven, R. Grimm, J. Levinsen, M. M. Parish, M. Schmidt, R. Schmidt, et al., *Ultrafast many-body interferometry of impurities coupled to a fermi sea*, Science **354** (2016), no. 6308, 96–99. See pages: 84 and 117

[CN71]   M. Combescot and P. Nozières, *Infrared catastrophy and excitons in the x-ray spectra of metals*, J. Phys. **32** (1971), no. 11-12, 913–929. See pages: 87, 88, 117, and 134

[CPC+18]   P. Chen, W. W. Pai, Y.-H. Chan, V. Madhavan, M. Y. Chou, S.-K. Mo, A.-V. Fedorov, and T.-C. Chiang, *Unique gap structure and symmetry of the charge density wave in single-layer* vse$_2$, Phys. Rev. Lett. **121** (2018), 196402. See page: 43

[CPR04]   A. V. Chubukov, C. Pépin, and J. Rech, *Instability of the quantum-critical point of itinerant ferromagnets*, Phys. Rev. Lett. **92** (2004), 147003. See page: 12

[DH81]   C. Dasgupta and B. I. Halperin, *Phase transition in a lattice model of superconductivity*, Phys. Rev. Lett. **47** (1981), 1556–1560. See page: 62

[DL13]   D. Dalidovich and S.-S. Lee, *Perturbative non-fermi liquids from dimensional regularization*, Phys. Rev. B **88** (2013), no. 24, 245106. See pages: 15 and 19

[DM06]   L. Dell'Anna and W. Metzner, *Fermi surface fluctuations and single electron excitations near pomeranchuk instability in two dimensions*, Phys. Rev. B **73** (2006), 045127. See pages: 7 and 12

[DORB+19]   N. Darkwah Oppong, L. Riegger, O. Bettermann, M. Höfer, J. Levinsen, M. M. Parish, I. Bloch, and S. Fölling, *Observation of coherent multiorbital polarons in a two-dimensional fermi gas*, Phys. Rev. Lett. **122** (2019), 193604. See page: 84

[Edw13]   D. M. Edwards, *A smooth polaron–molecule crossover in a fermi system*, J. Phys. Cond. Mat. **25** (2013), no. 42, 425602. See pages: 116, 135, 136, and 138

[EMS16]   A. Eberlein, I. Mandal, and S. Sachdev, *Hyperscaling violation at the ising-nematic quantum critical point in two-dimensional metals*, Phys. Rev. B **94** (2016), 045133. See page: 20

[FF64]   P. Fulde and R. A. Ferrell, *Superconductivity in a strong spin-exchange field*, Phys. Rev. **135** (1964), no. 3A, A550. See page: 21

[FvW15]   F. Flicker and J. van Wezel, *Charge order from orbital-dependent coupling evidenced by nbse2*, Nature Communications **6** (2015), 7034 EP –. See page: 43

[FvWW+15]   Y. Feng, J. van Wezel, J. Wang, F. Flicker, D. M. Silevitch, P. B. Littlewood, and T. F. Rosenbaum, *Itinerant density wave instabilities at classical and quantum critical points*, Nature Physics **11** (2015), 865 EP –. See page: 43

[FWJ+12]   Y. Feng, J. Wang, R. Jaramillo, J. van Wezel, S. Haravifard, G. Srajer, Y. Liu, Z.-A. Xu, P. B. Littlewood, and T. F. Rosenbaum, *Order parameter fluctuations at a buried quantum critical point*, Proceedings of the National Academy of Sciences **109** (2012), no. 19, 7224–7229. See page: 43

[GCH03]   O. Gunnarsson, M. Calandra, and J. E. Han, *Colloquium: Saturation of electrical resistivity*, Rev. Mod. Phys. **75** (2003), 1085–1099. See page: 8

[GDC+98]   G.-H. Gweon, J. D. Denlinger, J. A. Clack, J. W. Allen, C. G. Olson, E. DiMasi, M. C. Aronson, B. Foran, and S. Lee, *Direct observation of complete fermi surface, imperfect nesting, and gap anisotropy in the high-temperature incommensurate charge-density-wave compound* $smte_3$, Phys. Rev. Lett. **81** (1998), 886–889. See page: 43

[Gie12]   H. Gies, *Introduction to the functional rg and applications to gauge theories*, Renormalization Group and Effective Field Theory Approaches to Many-Body Systems, Springer, Berlin, 2012, pp. 287–348. See page: 65

[Gin61]   V. L. Ginzburg, *Some remarks on phase transitions of the second kind and the microscopic theory of ferroelectric materials*, Soviet Phys. Solid State **2** (1961), 1824–1834. See page: 1

[GMPS16]   O. Goulko, A. S. Mishchenko, N. Prokofév, and B. Svistunov, *Dark continuum in the spectral function of the resonant fermi polaron*, Phys. Rev. A **94** (2016), no. 5, 051605. See pages: 116 and 117

[GNRC69]   J. Gavoret, P. Nozieres, B. Roulet, and M. Combescot, *Optical absorption in degenerate semiconductors*, J. Phys. **30** (1969), no. 11-12, 987–997. See page: 86

[GPC+07]   A. Gabbay, Y. Preezant, E. Cohen, B. M. Ashkinadze, and L. N. Pfeiffer, *Fermi edge polaritons in a microcavity containing a high density two-dimensional electron gas*, Phys. Rev. Lett. **99** (2007), 157402. See pages: 84 and 89

[GS08]   K. B. Gubbels and H. T. C. Stoof, *Renormalization group theory for the imbalanced fermi gas*, Phys. Rev. Lett. **100** (2008), no. 14, 140407. See page: 139

[H. 09]   H. Haug and S. W. Koch, *Quantum theory of the optical and electronic properties of semiconductors*, World Scientific, 2009. See page: 89

[Hal83]   F. D. M. Haldane, *Nonlinear field theory of large-spin heisenberg antiferromagnets: Semiclassically quantized solitons of the one-dimensional easy-axis néel state*, Phys. Rev. Lett. **50** (1983), 1153–1156. See page: 56

[Hal19]   J. Halbinger, *Quantum phase transition to incommensurate* $2k_f$ *charge density wave order*, Master's thesis, Ludwig Maximilians University of Munich, 2019. See page: 43

[Has04]   M. B. Hastings, *Lieb-schultz-mattis in higher dimensions*, Phys. Rev. B **69** (2004), no. 10, 104431. See page: 56

[Haw91]   P. Hawrylak, *Optical properties of a two-dimensional electron gas: Evolution of spectra from excitons to fermi-edge singularities*, Phys. Rev. B **44** (1991), 3821–3828. See page: 86

[Her76]   J. A. Hertz, *Quantum critical phenomena*, Phys. Rev. B **14** (1976), 1165–1184. See page: 13

[HM12]   T. Holder and W. Metzner, *Incommensurate nematic fluctuations in two-dimensional metals*, Phys. Rev. B **85** (2012), 165130. See page: 11

[HM14]   ———, *Non-fermi-liquid behavior at the onset of incommensurate $2k_F$ charge- or spin-density wave order in two dimensions*, Phys. Rev. B **90** (2014), 161106. See page: 43

[HPP19]   J. Halbinger, D. Pimenov, and M. Punk, *Incommensurate $2k_F$ density wave quantum criticality in two-dimensional metals*, Phys. Rev. B **99** (2019), 195102. See page: 44

[Hus08]   N. E. Hussey, *Phenomenology of the normal state in-plane transport properties of high-tc cuprates*, J. Phys. Cond. Mat. **20** (2008), no. 12, 123201. See page: 7

[INH09]   K. Ishida, Y. Nakai, and H. Hosono, *To what extent iron-pnictide new superconductors have been clarified: A progress report*, Journal of the Physical Society of Japan **78** (2009), no. 6, 062001. See page: 7

[Kad66]   L. P. Kadanoff, *Scaling laws for ising models near t_c*, Physics Physique Fizika **2** (1966), no. 6, 263. See page: 1

[Kau11]   Ribhu K Kaul, *Quantum criticality in su (3) and su (4) antiferromagnets*, Physical Review B **84** (2011), no. 5, 054407. See page: 67

[KBS10]   P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the functional renormalization group*, Lect. Notes Phys. 798, Springer, Heidelberg, 2010. See pages: 63 and 140

[KL17]   B. Kain and H. Y. Ling, *Hartree-fock treatment of fermi polarons using the lee-low-pine transformation*, Phys. Rev. A **96** (2017), 033627. See page: 138

[KM65]   W. Kohn and C. Majumdar, *Continuity between bound and unbound states in a fermi gas*, Phys. Rev. **138** (1965), A1617–A1620. See page: 134

[Kos74]   J. M. Kosterlitz, *The critical properties of the two-dimensional xy model*, J. of Phys. C: Solid State Phys. **7** (1974), no. 6, 1046. See page: 1

[KPV$^+$12]   M. Koschorreck, D. Pertot, E. Vogt, B. Fröhlich, M. Feld, and M. Köhl, *Attractive and repulsive fermi polarons in two dimensions*, Nature **485** (2012), no. 7400, 619. See page: 84

[KS12]   R. K. Kaul and A. W. Sandvik, *Lattice model for the su (n) néel to valence-bond solid quantum phase transition at large n*, Phys. Rev. Lett. **108** (2012), no. 13, 137201. See page: 67

[KSF01]   H. Kleinert and V. Schulte-Frohlinde, *Critical properties of ϕ^4-theories*, World Scientific, Singapore, 2001. See pages: 4, 5, 16, 17, 18, and 19

[KvD18]   F. B. Kugler and J. von Delft, *Multiloop functional renormalization group that sums up all parquet diagrams*, Phys. Rev. Lett. **120** (2018), 057403. See pages: 86 and 143

[KZJ$^+$12]   C. Kohstall, M. Zaccanti, M. Jag, A. Trenkwalder, P. Massignan, G. M. Bruun, F. Schreck, and R. Grimm, *Metastability and coherence of repulsive polarons in a strongly interacting fermi mixture*, Nature **485** (2012), no. 7400, 615. See pages: 84 and 117

[Lai13]   H.-H. Lai, *Possible spin liquid states with parton fermi surfaces in the su (3) ring-exchange model on the triangular lattice*, Phys. Rev. B **87** (2013), no. 20, 205131. See page: 67

[Lan37]   L. D. Landau, *On the theory of phase transitions*, Ukr. J. Phys. **11** (1937), 19–32. See page: 1

[Lee09]   S.-S. Lee, *Low-energy effective theory of fermi surface coupled with u(1) gauge field in $2+1$ dimensions*, Phys. Rev. B **80** (2009), 165102. See page: 14

[Lee18]   _____ , *Recent developments in non-fermi liquid theory*, Ann. Rev. Condes. Matter Phys. **9** (2018), no. 1, 227–244. See page: 15

[Lev59]   A. P. Levanyuk, *Contribution to the theory of light scattering near the second-order phase-transition points*, Sov. Phys. JETP **9** (1959), no. 3, 571–576. See page: 1

[LL14]   Z. Lan and C. Lobo, *A single impurity in an ideal atomic fermi gas: current understanding and some open problems*, arXiv preprint arXiv:1404.3220 (2014). See page: 134

[LO65]   A. I. Larkin and Y. N. Ovchinikov, *Inhomogeneous state of superconductors*, Sov. Phys. JETP **20** (1965), 762. See page: 21

[LS04]   M. Levin and T. Senthil, *Deconfined quantum criticality and néel order via dimer disorder*, Phys. Rev. B **70** (2004), no. 22, 220403. See pages: 61 and 62

[LSB07]   J. Lou, A. W. Sandvik, and L. Balents, *Emergence of u(1) symmetry in the 3d xy model with Z_q anisotropy*, Phys. Rev. Lett. **99** (2007), 207203. See page: 62

[LSBK15]   S. Lederer, Y. Schattner, E. Berg, and S. A. Kivelson, *Enhancement of superconductivity near a nematic quantum critical point*, Phys. Rev. Lett. **114** (2015), 097001. See page: 8

[LSM61]   E. Lieb, T. Schultz, and D. Mattis, *Two soluble models of an antiferromagnetic chain*, Annals of Physics **16** (1961), no. 3, 407–466. See page: 56

[LWD$^+$07] R. Lortz, Y. Wang, A. Demuer, P. H. M. Böttger, B. Bergk, G. Zwicknagl, Y. Nakazawa, and J. Wosnitza, *Calorimetric evidence for a fulde-ferrell-larkin-ovchinnikov superconducting state in the layered organic superconductor κ-(bedt-ttf) 2 cu (ncs) 2*, Phys. Rev. Lett. **99** (2007), no. 18, 187002. See pages: 8 and 22

[LWMA17]   M. Lajkó, K. Wamer, F. Mila, and I. Affleck, *Generalization of the haldane conjecture to su(3) chains*, Nucl. Phys. B **924** (2017), 508 – 577. See page: 67

[Mah00]   G.D. Mahan, *Many-particle-physics*, 3rd ed., Kluwer Academic/Plenum Publishers, New York and London, 2000. See pages: 7, 83, 85, and 86

[Man16]   I. Mandal, *Superconducting instability in non-fermi liquids*, Phys. Rev. B **94** (2016), no. 11, 115138. See page: 20

[MC04]   C. Mora and R. Combescot, *Nature of the fulde-ferrell-larkin-ovchinnikov phases at low temperature in 2 dimensions*, Europhys. Lett. **66** (2004), no. 6, 833. See page: 22

[MGB⁺19]   B. Michon, C. Girod, S. Badoux, J. Kačmarčík, Q. Ma, M. Dragomir, H. A. Dabkowska, B.. D. Gaulin, J.-S. Zhou, S. Pyon, et al., *Thermodynamic signatures of quantum criticality in cuprate superconductors*, Nature **567** (2019), no. 7747, 218. See page: 8

[Mih11]   B. Mihaila, *Lindhard function of a d-dimensional fermi gas*, arXiv preprint arXiv:1111.5337 (2011). See page: 43

[Mil93]   A. J. Millis, *Effect of a nonzero temperature on quantum critical points in itinerant fermion systems*, Phys. Rev. B **48** (1993), 7183–7196. See page: 13

[ML15]   I. Mandal and S.-S. Lee, *Ultraviolet/infrared mixing in non-fermi liquids*, Phys. Rev. B **92** (2015), no. 3, 035141. See page: 15

[MMSS15]   M. A. Metlitski, D. F. Mross, S. Sachdev, and T. Senthil, *Cooper pairing in non-fermi liquids*, Phys. Rev. B **91** (2015), 115111. See page: 8

[MPH11]   C. J. M. Mathy, M. M. Parish, and D. A. Huse, *Trimers, molecules, and polarons in mass-imbalanced atomic fermi gases*, Phys. Rev. Lett. **106** (2011), no. 16, 166404. See pages: 134 and 135

[MS07]   Y. Matsuda and H. Shimahara, *Fulde–ferrell–larkin–ovchinnikov state in heavy fermion superconductors*, J. Phys. Soc. Jpn. **76** (2007), no. 5, 051005. See page: 21

[MS10a]   M. A. Metlitski and S. Sachdev, *Quantum phase transitions of metals in two spatial dimensions. i. ising-nematic order*, Phys. Rev. B **82** (2010), no. 7, 075127. See pages: 12, 13, and 14

[MS10b]   ———, *Quantum phase transitions of metals in two spatial dimensions. ii. spin density wave order*, Phys. Rev. B **82** (2010), no. 7, 075128. See pages: 10 and 14

[MSH⁺12]   W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, *Functional renormalization group approach to correlated fermion systems*, Rev. Mod. Phys. **84** (2012), 299–352. See page: 63

[MT64]   K. Maki and T. Tsuneto, *Pauli paramagnetism and superconducting state*, Progress of Theoretical Physics **31** (1964), no. 6, 945–956. See page: 21

[NDD69]   P. Nozières and C. T. De Dominicis, *Singularities in the x-ray absorption and emission of metals. iii. one-body theory exact solution*, Phys. Rev. **178** (1969), 1097–1107. See pages: 84 and 86

[NGR69]   P. Nozières, J. Gavoret, and B. Roulet, *Singularities in the x-ray absorption and emission of metals. ii. self-consistent treatment of divergences*, Phys. Rev. **178** (1969), no. 3, 1084. See pages: 84 and 86

[Noz94]   P. Nozières, *The effect of recoil on edge singularities*, J. Phys. **4 (9)** (1994), 1275–1280. See page: 86

[NS06]   Y. Nishida and D. T. Son, *ϵ expansion for a fermi gas at infinite scattering length*, Physical review letters **97** (2006), no. 5, 050403. See page: 141

[NS07a]   P. Nikolić and S. Sachdev, *Renormalization-group fixed points, universal phase diagram, and 1/ n expansion for quantum liquids with interactions near the unitarity limit*, Phys. Rev. A **75** (2007), no. 3, 033608. See page: 140

[NS07b]   Y. Nishida and D. T. Son, *Fermi gas near unitarity around four and two spatial dimensions*, Physical Review A **75** (2007), no. 6, 063617. See page: 141

[Osh00]   M. Oshikawa, *Commensurability, excitation gap, and topology in quantum many-particle systems on a periodic lattice*, Phys. Rev. Lett. **84** (2000), no. 7, 1535. See page: 56

[PDZ09]   M. Punk, P. T. Dumitrescu, and W. Zwerger, *Polaron-to-molecule transition in a strongly imbalanced fermi gas*, Phys. Rev. A **80** (2009), no. 5, 053605. See pages: 116, 133, and 138

[PG18]   D. Pimenov and M. Goldstein, *Spectra of heavy polarons and molecules coupled to a fermi sea*, Phys. Rev. B **98** (2018), 220302. See pages: 84, 87, 88, 133, and 134

[Pim15]   D. Pimenov, *Fermi-edge polaritons with finite hole-mass*, Master's thesis, Ludwig Maximilians University Munich, 2015. See pages: 84 and 89

[PL13]   M. M. Parish and J. Levinsen, *Highly polarized fermi gases in two dimensions*, Phys. Rev. A **87** (2013), no. 3, 033616. See pages: 134, 135, 136, 137, and 138

[PMPP18]   D. Pimenov, I. Mandal, F. Piazza, and M. Punk, *Non-fermi liquid at the fflo quantum critical point*, Phys. Rev. B **98** (2018), 024510. See page: 21

[Pro93]   N.V. Prokof'ev, *Diffusion of a heavy particle in a fermi-liquid theory*, Int. J. of Mod. Phys. B **07** (1993), no. 19, 3327–3351. See page: 86

[PS08a]   N. V. Prokof'ev and B. V. Svistunov, *Bold diagrammatic monte carlo: A generic sign-problem tolerant technique for polaron models and possibly interacting many-body problems*, Phys. Rev. B **77** (2008), no. 12, 125101. See page: 116

[PS08b]   _____ , *Fermi-polaron problem: Diagrammatic monte carlo method for divergent sign-alternating series*, Phys. Rev. B **77** (2008), no. 2, 020408. See pages: 116 and 134

[PSS15]   A. A. Patel, P. Strack, and S. Sachdev, *Hyperscaling at the spin density wave quantum critical point in two-dimensional metals*, Phys. Rev. B **92** (2015), 165105. See page: 20

[Pun10]   M. Punk, *Many-particle-physics with ultracold gases*, Ph.D. thesis, Technical University of Munich, 2010. See pages: 88, 137, and 142

[PvDGG17]   D. Pimenov, J. von Delft, L. Glazman, and M. Goldstein, *Fermi-edge exciton-polaritons in doped semiconductor microcavities with finite hole mass*, Phys. Rev. B **96** (2017), no. 15, 155310. See pages: 84, 87, and 88

[PZS16]   F. Piazza, W. Zwerger, and P. Strack, *Fflo strange metal and quantum criticality in two dimensions: Theory and application to organic superconductors*, Phys. Rev. B **93** (2016), no. 8, 085112. See pages: 8, 21, and 22

[RDS90]   M. Randeria, J.-M. Duan, and L Shieh, *Superconductivity in a two-dimensional fermi gas: Evolution from cooper pairing to bose condensation*, Phys. Rev. B **41** (1990), 327–343. See page: 88

[Rei06]   J. Reiss, *Renormalised mean-field analysis of the 2d hubbard model*, Ph.D. thesis, University of Stuttgart, 2006. See page: 9

[RGN69]   B. Roulet, J. Gavoret, and P. Nozières, *Singularities in the x-ray absorption and emission of metals. i. first-order parquet calculation*, Phys. Rev. **178** (1969), 1072–1083. See pages: 84 and 86

[RK95]   A. Rosch and T. Kopp, *Heavy particle in a d-dimensional fermionic bath: A strong coupling approach*, Phys. Rev. Lett. **75** (1995), no. 10, 1988. See pages: 86 and 138

[Ros99]   A. Rosch, *Quantum-coherent transport of a heavy particle in a fermionic bath*, Advances in Physics **48** (1999), no. 3, 295–394. See pages: 86 and 87

[RSR87]   A. E. Ruckenstein and S. Schmitt-Rink, *Many-body aspects of the optical spectra of bulk and low-dimensional doped semiconductors*, Phys. Rev. B **35** (1987), 7551–7557. See page: 86

[RW94]   M. Reuter and C. Wetterich, *Exact evolution equation for scalar electrodynamics*, Nucl. Phys. B **427** (1994), no. 1, 291–324. See pages: 65 and 66

[Sac11]   S. Sachdev, *Quantum phase transitions*, Cambridge Univerisity Press, Cambridge, UK, 2011. See pages: 1, 5, 8, 12, 13, 16, and 59

[San07]   A. W. Sandvik, *Evidence for deconfined quantum criticality in a two-dimensional heisenberg model with four-spin interactions*, Physical review letters **98** (2007), no. 22, 227202. See page: 57

[SBC+17] M. Sidler, P. Back, O. Cotlet, A. Srivastava, T. Fink, M. Kroner, E. Demler, and A. Imamoglu, *Fermi polaron-polaritons in charge-tunable atomically thin semiconductors*, Nature Physics **13** (2017), no. 3, 255–261. See pages: 84, 89, and 117

[SBS+04] T. Senthil, L. Balents, S. Sachdev, A. Vishwanath, and M. P. A. Fisher, *Quantum criticality beyond the landau-ginzburg-wilson paradigm*, Phys. Rev. B **70** (2004), 144407. See pages: 2, 57, 58, 59, and 60

[SBS+05] _____, *Deconfined criticality critically defined*, J. of the Phys. Soc. of Japan **74** (2005), no. Suppl, 1–9. See page: 61

[SE11]   R. Schmidt and T. Enss, *Excitation spectra and rf response near the polaron-to-molecule transition from the functional renormalization group*, Phys. Rev. A **83** (2011), no. 6, 063620. See pages: 139, 140, and 141

[Sen07]   A. Sen, *Deconfined quantum critical points*, Master's thesis, Tata Institute for Fundamental Research, Mumbai, 2007. See page: 59

[Sen08]   T. Senthil, *Critical fermi surfaces and non-fermi liquid metals*, Phys. Rev. B **78** (2008), 035103. See page: 7

[Sha94]   R. Shankar, *Renormalization-group approach to interacting fermions*, Reviews of Modern Physics **66** (1994), no. 1, 129. See pages: 8 and 9

[SHM18] J. Sýkora, T. Holder, and W. Metzner, *Fluctuation effects at the onset of the $2k_F$ density wave order with one pair of hot spots in two-dimensional metals*, Phys. Rev. B **97** (2018), 155159. See page: 43

[SKI+18] R. Schmidt, M. Knap, D. A. Ivanov, J.-S. You, M. Cetina, and E. Demler, *Universal many-body response of heavy impurities coupled to a fermi sea*, Rep. Prog. Phys. **81** (2018), 024401. See pages: 84, 87, 117, and 134

[SL15] S. Sur and S.-S. Lee, *Quasilocal strange metal*, Phys. Rev. B **91** (2015), no. 12, 125136. See page: 20

[SL16] _____ , *Anisotropic non-fermi liquids*, Phys. Rev. B **94** (2016), 195135. See page: 19

[SLL17] A. Schlief, P. Lunts, and S.-S. Lee, *Exact critical exponents for the antiferromagnetic quantum critical metal in two dimensions*, Phys. Rev. X **7** (2017), 021010. See pages: 19 and 143

[SLL18] _____ , *Noncommutativity between the low-energy limit and integer dimension limits in the ϵ expansion: A case study of the antiferromagnetic quantum critical metal*, Phys. Rev. B **98** (2018), 075140. See page: 19

[SM06] K. V. Samokhin and M. S. Mar'enko, *Quantum fluctuations in larkin-ovchinnikov-fulde-ferrell superconductors*, Phys. Rev. B **73** (2006), no. 14, 144502. See page: 22

[SS09] T. Senthil and R. Shankar, *Fermi surfaces in general codimension and a new controlled nontrivial fixed point*, Phys. Rev. Lett. **102** (2009), no. 4, 046406. See page: 15

[SS13] A. Smerald and N. Shannon, *Theory of spin excitations in a quantum spin-nematic state*, Phys. Rev. B **88** (2013), no. 18, 184430. See page: 67

[Ste01] G. R. Stewart, *Non-fermi-liquid behavior in d- and f-electron metals*, Rev. Mod. Phys. **73** (2001), 797–855. See pages: 7 and 8

[SVB+04] T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, and M. P. A. Fisher, *Deconfined quantum critical points*, Science **303** (2004), no. 5663, 1490–1494. See page: 57

[SVM+17] F. Scazza, G. Valtolina, P. Massignan, A. Recati, A. Burchianti, C. Fort, M. Inguscio, M. Zaccanti, and G. Roati, *Repulsive fermi polarons in a resonant mixture of ultracold li 6 atoms*, Phys. Rev. Lett. **118** (2017), no. 8, 083602. See page: 84

[SWH+14] S. Smolka, W.Wuester, F. Haupt, S. Faelt, W. Wegschneider, and A. Imamoglu, *Cavity quantum electrodynamics with many-body state of a two-dimensional electron gas*, Science **346** (2014), no. 6207, 332–335. See pages: 84, 89, and 90

[SWSZ09] A. Schirotzek, C.-H. Wu, A. Sommer, and M. W. Zwierlein, *Observation of fermi polarons in a tunable fermi liquid of ultracold atoms*, Phys. Rev. Lett. **102** (2009), no. 23, 230402. See pages: 83, 84, and 116

[SYH$^+$13]   A. Soumyanarayanan, M. M. Yee, Y. He, J. van Wezel, D. J. Rahn, K. Rossnagel, E. W. Hudson, M. R. Norman, and J. E. Hoffman, *Quantum phase transition from triangular to stripe charge order in nbse2*, Proc. of the Nat. Acad. of Sciences **110** (2013), no. 5, 1623–1627. See page: 43

[TT87]   J.-C. Tolédano and P. Tolédano, *The landau theory of phase transitions*, World Scientific, Singapore, 1987. See pages: 4 and 56

[US90]   T. Uenoyama and L. J. Sham, *Effect of finite hole mass on edge singularities in optical spectra*, Phys. Rev. Lett. **65** (1990), 1048–1051. See page: 86

[Vas04]   A. N. Vasiliev, *The field theoretic renormalization group in critical behavior theory and stochastic dynamics*, CRC press, Boca Raton, 2004. See page: 16

[vLRVW07]   H. von Löhneysen, A. Rosch, M. Vojta, and P. Wölfle, *Fermi-liquid instabilities at magnetic quantum phase transitions*, Rev. Mod. Phys. **79** (2007), 1015–1075. See pages: 4, 5, 7, and 8

[VRVH13]   J. Vlietinck, J. Ryckebusch, and K. Van Houcke, *Quasiparticle properties of an impurity in a fermi gas*, Phys. Rev. B **87** (2013), no. 11, 115133. See page: 133

[WAA$^+$16]   Y. Wang, A. Abanov, B. L. Altshuler, E. A. Yuzbashyan, and A. V. Chubukov, *Superconductivity near a quantum-critical point: The special role of the first matsubara frequency*, Phys. Rev. Lett. **117** (2016), 157001. See page: 8

[WCG$^+$18]   G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B Urbaszek, *Colloquium: Excitons in atomically thin transition metal dichalcogenides*, Rev. Mod. Phys. **90** (2018), 021001. See page: 84

[Wen04]   X.-G. Wen, *Quantum field theory of many-body systems: from the origin of sound to an origin of light and electrons*, Oxford University Press, Oxford, 2004. See page: 62

[WF72]   K. G. Wilson and M. E. Fisher, *Critical exponents in 3.99 dimensions*, Phys. Rev. Lett. **28** (1972), no. 4, 240. See page: 5

[WK74]   K. G. Wilson and J. Kogut, *The renormalization group and the ϵ expansion*, Physics reports **12** (1974), no. 2, 75–199. See page: 1

[WMH$^+$08]   K. E. Wagner, E. Morosan, Y. S. Hor, J. Tao, Y. Zhu, T. Sanders, T. M. McQueen, H. W. Zandbergen, A. J. Williams, D. V. West, and R. J. Cava, *Tuning the charge density wave and superconductivity in cu$_x$tas$_2$*, Phys. Rev. B **78** (2008), 104520. See page: 43

[WNM$^+$17]   C. Wang, A. Nahum, M. A. Metlitski, C. Xu, and T. Senthil, *Deconfined quantum critical points: symmetries and dualities*, Phys. Rev. X **7** (2017), no. 3, 031051. See page: 59

[Wos18]   J. Wosnitza, *Fflo states in layered organic superconductors*, Annalen der Physik **530** (2018), no. 2, 1700282. See page: 22

[Xu12]   C. Xu, *Unconventional quantum critical points*, Int. J. of Mod. Phys. B **26** (2012), no. 18, 1230007. See pages: 2 and 4

[YRKK06]   H. Yao, J. A. Robertson, E.-A. Kim, and S. A. Kivelson, *Theory of stripes in quasi-two-dimensional rare-earth tellurides*, Phys. Rev. B **74** (2006), 245126. See page: 43

Acknowledgement

I would never have completed this thesis without continuous support and encouragement by many people. A big thank you goes to