Cavity grid for scalable quantum computation with superconducting circuits

F. Helmer¹, M. Mariantoni²,³, A. G. Fowler⁴, J. von Delft¹, E. Solano¹,⁵ and F. Marquardt¹(a)

¹ Department of Physics, CeNS, and ASC, Ludwig-Maximilians-Universität - Theresienstrasse 37, 80333 Munich, Germany, EU
² Walther-Meißner-Institut, Bayer. Akademie der Wissenschaften - Walther-Meißner-Str. 8, 85748 Garching, Germany, EU
³ Department of Physics, Technische Universität München - James-Franck-Str., 85748 Garching, Germany, EU
⁴ Institute for Quantum Computing, University of Waterloo - Waterloo, ON, Canada
⁵ Departamento de Química Física, Universidad del País Vasco - Euskal Herriko Unibertsitatea - 48080 Bilbao, Spain, EU

received 4 December 2008; accepted 16 February 2009
published online 20 March 2009

PACS 03.67.Lx – Quantum computation architectures and implementations
PACS 42.50.Pq – Cavity quantum electrodynamics; micromasers
PACS 85.25.-j – Superconducting devices

Abstract – We propose an architecture for quantum computing based on superconducting circuits, where on-chip planar microwave resonators are arranged in a two-dimensional grid with a qubit at each intersection. This allows any two qubits on the grid to be coupled at a swapping overhead independent of their distance. We demonstrate that this approach encompasses the fundamental elements of a scalable fault-tolerant quantum-computing architecture.

Introduction. – Superconducting circuits are promising candidates for scalable quantum information processing [1–10]. This route was further strengthened with the advent of circuit quantum electrodynamics. Starting with early proposals for implementing the quantum-optical Jaynes-Cummings model in the context of superconducting circuits [11–13], this research direction became a major topic after it was pointed out that on-chip microwave transmission line resonators could be coupled to superconducting qubits [14]. Since then, a series of groundbreaking experiments have demonstrated these concepts [15–17], including achievements like dispersive qubit readout [18], photon number splitting [19], single-photon generation [20], or lasing by a single artificial atom [21].

Recent experiments [22,23] have advanced to coupling two qubits via the cavity, yielding a flip-flop (XY) interaction permitting two-qubit gates. If multiple qubits share one cavity, arbitrary qubit pairs could be selectively coupled [24], which outperforms nearest-neighbor setups (no swapping overhead or disruption by single unusable qubits). However, moving towards more qubits requires suitable novel architectures.

In this letter, we present and analyze an architecture that builds on these elements and extends them into the second dimension, by forming a crossbar-like geometry of orthogonal microwave resonators, with qubits sitting at the intersections (fig. 1). The global coupling within each row and column makes this setup distinct from existing proposals for array-like arrangements in ion traps [25,26], optical lattices [27], semiconductor spins [28] or superconductors [1,29]. We show i) how to couple any two qubits on the grid, ii) with minimal swapping overhead using iii) an appropriately chosen (“Sudoku”-style) frequency distribution, and iv) suggest a scalable fault-tolerant quantum-computing architecture.

Before we turn to a description of our proposal, we note that experiments right now are obviously still struggling to improve the fidelity of single- and two-qubit operations for superconducting qubits, and this painstaking work is crucial for further progress in the whole field. Nevertheless, the effort going into this endeavor is ultimately justified by the long-term goal of implementing large-scale circuits able to perform nontrivial quantum computation tasks, where the numbers of qubits may run into the

(a)E-mail: Florian.Marquardt@physik.lmu.de
For definiteness we consider charge (or transmon) qubits, unless noted otherwise. Then the couplings \( g_{ij}^{A(B)} \) between the horizontal (vertical) cavity mode \( i \) (\( j \)) and the dipole operator \( \hat{n}_{ij} \) of qubit \((i, j)\) depend on the detailed electric-field distribution and geometry of the qubit. Equation (1) leads to the Jaynes-Cummings model and the cavity-mediated interaction between qubits [14]. It can be realized in different ways: A capacitive coupling was demonstrated for charge [15] (or “transmon” [19,23]) and phase qubits [22], while for flux qubits [16] \( \hat{n}_{ij} \) is the magnetic-moment coupling to the magnetic field.

It is well known [14,23] that the Hamiltonian (1) induces an effective flip-flop interaction of strength \( J_{\alpha\beta} = g_{\alpha\beta}(\Delta_{\alpha} + \Delta_{\beta})/(2\Delta_{\alpha}\Delta_{\beta}) \) between each pair of qubits \((\alpha, \beta)\) in the same cavity (for couplings \( g_{\alpha\beta} \) and detunings from the cavity \( \Delta_{\alpha(\beta)} \), in the dispersive limit \( |g| \ll |\Delta| \)).

\[
\hat{H}_{\text{flip-flop}} = J_{\alpha\beta} \left( \hat{\sigma}_{\alpha+} \hat{\sigma}_{\beta-} + \text{h.c.} \right). \tag{2}
\]

In the computational “idle state” these interactions have to be effectively turned off by detuning all the qubits from each other. This requires a detuning \( \delta \omega \gg J \) to avoid spurious two-qubit operations. Thus, the number \( N \) of qubits in a linear array is strongly restricted [24], since a frequency interval of order \( N\delta \omega \) is required. In the present 2D architecture, this constraint is considerably relaxed. The required frequency range is reduced from \( N\delta \omega \) to \( \sqrt{N} \delta \omega \) (where \( N \) is the total number of qubits), while still ensuring a spacing of \( \delta \omega \) within each cavity (the constraints being similar to the rules of the game “Sudoku”). This allows for grids with more than \( 20 \times 20 = 400 = N \) qubits, for realistic parameters. Figure 1 shows an acceptable frequency distribution. An extension to a fully scalable setup is discussed at the end of this paper.

**One-qubit operations.** – We briefly review some ingredients that have already been implemented [15–19]. Operations on a selected qubit can be performed via Rabi oscillations [18] using a microwave pulse resonant with the qubit at \( \epsilon_{ij}/\hbar \) but detuned from the cavity and all other qubits in the same cavity. Rotations around the \( z \)-axis can be performed via \( AC \) Stark shift [23], or by tuning the qubit frequency temporarily (see below). The cavities can be used for fast dispersive QND readout [18] of single qubits tuned close to the readout frequency or multiplexed readout of several qubits at once [14,23].

**Tunability.** – Additional charge and flux control lines (fig. 3) reaching each qubit are needed for tunability. For split-junction charge qubits [1,3], locally changing the magnetic flux sweeps the energy splitting \( \epsilon_{ij} = E_3(\Phi_{ij}) \) (see [22,30]), while keeping the qubit at the charge degeneracy point (to which it has been tuned via a separate charge gate line). This ensures maximum coherence through weak coupling to \( 1/f \) noise, although this requirement is relaxed in the new “transmon” design [31]. Individual addressability introduces some hardware overhead, thousands. While present-day experiments are still very far removed from this goal, it is worthwhile to develop architectures that couple more than a handful of qubits in a nontrivial setup, and which represent a challenging medium-term goal for the experiments to strive for. We will demonstrate that parameters (dephasing times, coupling strengths etc.) near those that are available nowadays would allow for a first proof-of-principle experiment in our proposed architecture, and further progress in the perfection of single qubits will enable truly useful larger-scale versions. The basic ideas behind our scheme are sufficiently general so as to permit replacing individual building blocks (particular qubit types, two-qubit gates etc.) by improved versions that might be developed within the coming years.

In addition, we would like to emphasize that even though any working set of universal one- and two-qubit operations permits to implement arbitrary computations in principle, it is by no means clear that the resulting generic implementation is efficient. Rather, in order to make the most efficient use of resources, it is mandatory to come up with larger-scale schemes that exploit the particular features of a given physical realization. In this sense, our proposal is similar in spirit to previous proposals for other physical systems that envisaged how well-known elementary operations could be extended to an efficient two-dimensional architecture [25–28].

**Basic architecture.** – The cavity grid consists of cavity modes belonging to \( N_A \) horizontal (A) and \( N_B \) vertical (B) cavities, \( \hat{H}_{\text{cav}} = \sum_{j=1}^{N_A} \hbar \omega_j A_j^+ A_j + \sum_{j=1}^{N_B} \hbar \omega_j B_j^+ B_j \), coupled to one qubit of frequency \( \epsilon_{ij} \) at each intersection \((i, j)\), generalizing [14] to a 2D architecture:

\[
\hat{H}_{\text{cav-qb}} = \sum_{i,j} \hat{n}_{ij} \left[ g_{ij}^A (\hat{a}_i + \hat{a}_i^+) + g_{ij}^B (\hat{b}_j + \hat{b}_j^+) \right]. \tag{1}
\]
but is essential both for two-qubit gates and for compensating fabrication spread.

Two-qubit gates. — We use the effective flip-flop interaction of eq. (2) [see (13,32–34)] to induce two-qubit gates. In the “idle state”, the interaction is ineffective, since the qubits are out of resonance, $|e_\alpha - e_\beta| \gg J$. During the gate, the two qubit frequencies are tuned into mutual resonance near the cavity frequency to increase $J$, see fig. 1(b). After a waiting time $t = \hbar \pi / [2 |J|]$, this realizes the universal two-qubit iSWAP gate (demonstrated experimentally in [23]), which can be used to construct CNOT and SWAP. Each SWAP($\alpha, \beta$) operation in the protocol (fig. 2) can be decomposed into three iSWAP gates between qubits $\alpha$ and $\beta$ [35]:

$$\text{SWAP} = \text{iSWAP} \cdot R_\beta \cdot \text{iSWAP} \cdot R_\alpha \cdot \text{iSWAP} \cdot R_\beta.$$  \hspace{1cm} (3)

Here $R_\alpha$ rotates qubit $\alpha$ by an angle $-\pi/2$ around the $x$-axis via a Rabi pulse. Arbitrary gates between any two qubits ($e.g.$, 1 and 3) in different cavities can be implemented via an intermediate qubit 2 at the junction of two orthogonal cavities containing 1 and 3 (see fig. 2). The sequence

$$\text{SOPS}(1,3) \equiv \text{SWAP}(1,2) \text{OP}(2,3) \text{SWAP}(1,2),$$  \hspace{1cm} (4)

leaves qubit 2 unchanged and performs the desired operation “OP” between 1 and 3.

We simulated such an operation (fig. 2) for realistic parameters. Relaxation and pure dephasing for each qubit $\alpha$, with rates $\gamma$ and $\gamma^\ast$, are modeled by a Lindblad master equation (where $\hat{P}_\alpha = |e_\alpha\rangle \langle e_\alpha|$ projects onto the excited state of qubit $\alpha$):

$$\dot{\hat{\rho}} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}] + \sum_\alpha (L^{\alpha}_\rho + L^{\alpha \ast}_\rho) \hat{\rho},$$  \hspace{1cm} (5)

$$L^{\alpha \rho}_\rho \hat{\rho} = \gamma^\rho \left[ 2 \hat{P}_\alpha \hat{P}_\rho \hat{P}_\alpha - \hat{P}_\rho \hat{P}_\rho - \hat{P}_\rho \hat{P}_\rho \right],$$  \hspace{1cm} (6)

$$L^{\alpha \rho \rho}_\rho \hat{\rho} = \gamma \left[ \hat{\sigma}_\alpha \hat{\rho} \hat{\sigma}_\alpha^\ast - \frac{1}{2} \hat{\sigma}_\alpha \hat{\sigma}_\alpha \hat{\rho} - \frac{1}{2} \hat{\rho} \hat{\sigma}_\alpha \hat{\sigma}_\alpha^\ast \right].$$  \hspace{1cm} (7)

We consider three qubits, where (1,2) and (2,3) are coupled via flip-flop terms (see eq. (2)) after adiabatic elimination of the cavities. During two-qubit gates, the qubit energy is ramped and will cross other qubit energies (fig. 1), potentially leading to spurious population transfer to other qubits if the process is too slow, while ramping too fast would excite higher qubit levels. For a 10 ns switching time (during which a sweep over $\delta \epsilon = 2 \pi \cdot 10$ GHz is accomplished), the probability of erroneous transfer during one crossing is estimated to be less than $10^{-2}$ from the Landau-Zener tunneling formula, and thus could be safely disregarded for the present simulation, where energies were instead switched instantaneously. Although several crossings may occur during one sweep, the scalable setup to be introduced further below keeps this kind of error under control by having only eight qubits per cavity. For the simulation we used the following parameters: Initially, the qubit transition frequencies are at $\epsilon = \hbar \omega_c = 2 \pi \cdot 4, 5, 6$ GHz. A resonant classical drive yields a Rabi frequency of $\Omega = 150$ MHz. A qubit-cavity coupling $g = 2 \pi \cdot 150$ MHz and a detuning $\Delta = 2 \pi \cdot 15$ GHz (from a cavity at $2 \pi \cdot 15$ GHz) produce $J / h = 2 \pi \cdot 21$ MHz. The dephasing and decay rates are $\gamma^\rho = 0.16$ MHz and $\gamma = 0.6$ MHz ($i.e.$ $T_1 = 1.7 \mu$s and $T_\rho = 6.3 \mu$s), consistent with recent experiments on transmon qubits [31]. Note that in the idle state the actual $J$ is reduced by a factor of 10 (due to larger detuning from the cavity). Employing a qubit spacing of $\delta \omega \sim 500$ MHz, this yields a residual coupling strength of $J^2 / \hbar^2 \delta \omega \sim 0.4$ MHz, which may be reduced further by refocusing techniques. To check the accuracy of adiabatic elimination, we performed an additional simulation of an iSWAP operation between two qubits taking the cavity fully into account, observing an error below the level brought about by dissipation.

A measure of the fidelity of the operation is obtained [36] by computing $F(\hat{\rho}_{\text{real}}(t), \hat{\rho}_{\text{ideal}}(t))$, where $F(\hat{\rho}_1, \hat{\rho}_2) \equiv \text{tr}(\sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}})^2$, and $\hat{\rho}_{\text{ideal}}$ denotes the time-evolution in the absence of dissipation. Figure 2 shows a fidelity of

![Fig. 2](link-to-image)
Fig. 3: (Colour on-line) A possible multilayer architecture. The layers 2 and 3 with coplanar wave guides are positioned above a “control line layer” 1. The qubit positions are indicated as red dots within each layer only for reference (they would be fabricated above layer 3).

about 95%, confirming that presently achievable parameters suffice for a first proof-of-principle experiment.

We emphasize that the swapping overhead does not grow with the distance between the qubits. Furthermore, multiple operations may run in parallel, even if they involve the same cavities, provided no qubit is affected simultaneously by two of the operations and the qubit pairs are tuned to different frequencies.

Here we have chosen the dispersive two-qubit gate that relies on proven achievements. Faster resonant gates (e.g., CPHASE [24,37]) might be implemented, with a time scale on the order of $1/g$ instead of $\Delta/g^2$.

**Hardware.** – For illustration, we discuss one out of many conceivable setups (fig. 3). The cavities can be coplanar wave guides or microstrip resonators. Available multilayer technology allows the fabrication of thin films stacked on top of each other. For example, consider wave guide layers of Nb or Al, separated by 100 nm of dielectric, which can be optimized for good decoherence properties (e.g., [38]). The inner grid area for $10 \times 10$ wave guides (each with $20 \mu m$ inner conductor and $10 \mu m$ gaps) would have a width of about 1 mm, whereas the full resonator lengths are above 10 mm, allowing all the qubits to be placed near the cavity mode central field antinode, with comparable couplings (see fig. 3). The qubit-cavity coupling remains similar to single-layer designs, owing to the small layer thickness of only 100 nm. Good isolation between two orthogonal cavities was estimated in a previous theoretical work [39,40], and unwanted cross-talk may be reduced further by choosing different cavity frequencies. The qubits can be placed above all layers to minimize fabrication problems. Weak coupling between qubits and control lines (e.g., cross-capacitance $\sim 0.01 fF$) suppresses sufficiently unwanted Nyquist noise from these lines, which could lead to decoherence. Indeed, for a Cooper-pair box of total capacitance $C_{\Sigma}$, this cross-capacitance yields a relaxation rate $\sim (C_g/C_{\Sigma})^2 e^2 \omega Z/\hbar$ for radiation into a control line of impedance $Z$ at the qubit splitting frequency $\omega$, leading to estimates that are small compared to the intrinsic qubit relaxation rate for the present parameters (the same holds for dephasing).

**Scalable, fault-tolerant architecture.** – The cavity grid is a building block for a truly scalable, fault-tolerant architecture. Scalability means that, at a minimum, the physics of initialisation, readout, single- and two-qubit gates does not depend on the total number of qubits. Figure 4 shows a scalable architecture requiring only eight different qubit frequencies. In each unit cell of 64 qubits (fig. 4) we choose two arrays of seven data qubits and use each array to store a single logical qubit, employing the Steane quantum error correction code [41]. Clean logical states can be prepared in additional ancilla qubits. Moreover, errors in the data qubits can be copied into the ancillae, which are then measured, locating the errors and enabling correction [42]. All other qubits are placeholders,
which are crucial: Swapping a pair of data qubits directly could corrupt both if the SWAP gate fails, resulting in a pair of errors that may not be correctable by the seven qubit Steane code. Using three SWAP gates with a placeholder qubit for temporary data storage solves this problem. We ignore errors in placeholder qubits as they contain no data.

A logical CNOT gate is illustrated in fig. 4. The final arrangement of qubits differs from the initial one and can be returned to it by swapping. However, if all logical qubits undergo similar logical gates, explicitly swapping back may be unnecessary as subsequent gates will do this automatically. A broad range of single logical qubit gates are possible. Full details of our chosen set of logical gates and their associated circuits including error correction can be found elsewhere [43,44].

Conclusions. – In this letter, we have proposed a novel architecture for quantum computation using a 2D grid of superconducting qubits coupled to an array of on-chip microwave cavities. A “Sudoku”-type arrangement of qubit frequencies permits global coupling of a large number of qubits while suppressing spurious interactions. These basic ideas could be implemented in a wide variety of hardware implementations. Elementary operations within this scheme could be demonstrated in the near future on small grids, while the setup has the potential to form the basis for truly scalable fault-tolerant architectures.

***

We acknowledge useful discussions with R. J. Schoelkopf, A. Imamoğlu, A. Blais, A. Wallraff, H. J. Majer, F. Deppe, Y. Nakamura, D. Esteve, C. Wilson and R. Gross, and support by EuroSQIP and the DFG research networks SFB 631 and NIM, and the Emmy-Noether program (FM). ES thanks the Ikerbasque Foundation, the EU EuroSQIP project, and the UPV-EHU grant GIU07/40.

REFERENCES