Scalable Quantum Circuit and Control for a Superconducting Surface Code

R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo

Phys. Rev. Applied **8**, 034021 — Published 25 September 2017

DOI: 10.1103/PhysRevApplied.8.034021

Scalable quantum circuit and control for a superconducting surface code

R. Versluis,[1,2] S. Poletto,[2] N. Khammassi,[3] B. Tarasinski,[2,4] N. Haider,[1,2]
D. J. Michalak,[5] A. Bruno,[2,4] K. Bertels,[3,4] and L. DiCarlo[2,4]

[1]*Netherlands Organisation for Applied Scientific Research (TNO), P.O. Box 155, 2600 AD Delft, The Netherlands*
[2]*QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*
[3]*Computer Engineering, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*
[4]*Kavli Institute of Nanoscience, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*
[5]*Components Research, Intel Corporation, 2501 NW 229th Ave, Hillsboro, OR 97124, USA*

(Dated: August 14, 2017)

We present a scalable scheme for executing the error-correction cycle of a monolithic surface-code fabric composed of fast-flux-tuneable transmon qubits with nearest-neighbor coupling. An eight-qubit unit cell forms the basis for repeating both the quantum hardware and coherent control, enabling spatial multiplexing. This control uses three fixed frequencies for all single-qubit gates and a unique frequency detuning pattern for each qubit in the cell. By pipelining the interaction and readout steps of ancilla-based X- and Z-type stabilizer measurements, we can engineer detuning patterns that avoid all second-order transmon-transmon interactions except those exploited in controlled-phase gates, regardless of fabric size. Our scheme is applicable to defect-based and planar logical qubits, including lattice surgery.

I. INTRODUCTION

The scaling of small quantum processors [1–5] into large qubit arrays capable of fault-tolerant quantum computation (FTQC) [6] is an outstanding challenge for leading experimental quantum information platforms [7, 8]. Modular [9] and monolithic [10] approaches require a systems approach that simultaneously and compatibly addresses challenges in all layers of the quantum computer stack [11]: from the quantum hardware at the low level, through classical control electronics in the middle, to software at the high level (i.e., micro-instruction sets, compilers, and high-level programming languages).

Currently, the surface code [10, 12, 13] provides an experimentally attractive paradigm for FTQC owing to its modest requirements on the quantum hardware: only nearest-neighbor coupling is needed between qubits, and the error threshold falls robustly close to 1% across a range of error models and error-decoding strategies, signficantly higher than those of Steane and Shor codes [6]. In superconducting quantum integrated circuits based on circuit QED (cQED) [14], the error rate of single-qubit gates has reached $< 0.1\%$ [15–17], while those of two-qubit conditional-phase (CZ) gates and measurement are 0.6% [15] and $\sim 1\%$ [18, 19], respectively.

The scalability of monolithic systems hinges on the ability to copy-paste a unit cell in the quantum plane, with suitable quantum interconnect between cells, and suitable classical interconnect to and from the control plane. The latter pursuit is very active, with several groups developing vertical (rather than the traditional lateral) interconnection of input/output (I/O) signals using through-the-wafer coaxial lines [20], electromechanical sockets [21], and bump-bonding in flip-chip configuration [22].

For true scalability, it is crucial that the unit cell also extend into the classical control plane. A unit cell for control signals opens the door to hardware simplification through spatial multiplexing, i.e., the selective routing of control signals (with minimal customization) to spatially separated components. While frequency-division multiplexing is already heavily exploited in cQED [3, 19, 23], spatial multiplexing is in its infancy. Precision control of same-frequency qubits using a microwave-frequency vector switch matrix (VSM) for pulse multicasting has only recently been demonstrated [24].

In this paper, we propose a scalable scheme for the QEC cycle of a monolithic superconducting surface code by defining a concrete unit cell for both the quantum hardware and the control signals. We focus on a fabric of fast-flux-tunable transmon qubits interacting with nearest neighbors via flux-controlled conditional-phase ($\mathrm{CZ_f}$) gates [25, 26] realized by pulsing into the resonator-mediated $|11\rangle \leftrightarrow |02\rangle$ avoided crossing of the interacting transmon pair (numbers indicate excitation level). Our approach is compatible with adiabatic [26], sudden [27] and fast-adiabatic [15, 28] use of these crossings. Our eight-qubit unit cell uses three fixed frequencies for all single-qubit control and eight detuning sequences for two-qubit gates. This approach to classical control allows significant control hardware savings via spatial multiplexing. By pipelining the measurement of the two types of stabilizers of the surface code, we engineer detuning patterns avoiding all second-order transmon-transmon interactions except those exploited in $\mathrm{CZ_f}$ gates, regardless of fabric size.

Our scheme allows changing the weight of stabilizer measurements by simple on/off masking of detuning pulses, making it applicable to both defect-based and planar logical qubits [10], including lattice surgery [29].

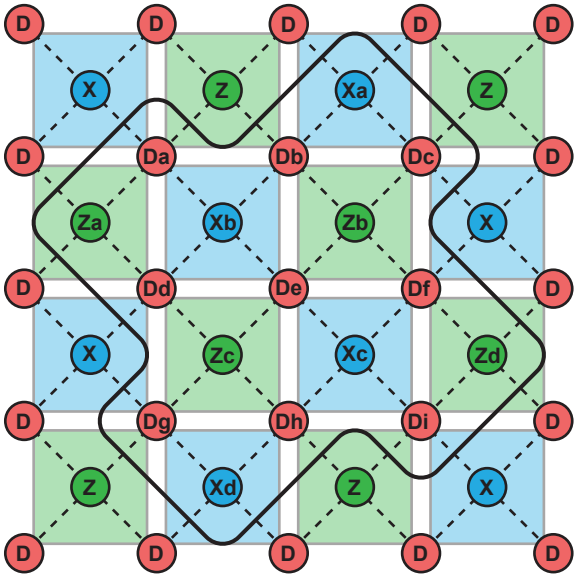

FIG. 1. Layout of a surface-code fabric. Red circles with D labels represent data qubits. Blue (green) circles with X (Z) labels represent ancillas performing X-type (Z-type) quantum parity checks of their nearest-neighbor data qubits. Each check is realized as an indirect quantum measurement, consisting of a coherent step involving pairwise interactions (dashed lines) followed by ancilla measurement. The delineated fabric of nine data qubits (D_a through D_i) and eight ancillas (X_a through X_d and Z_a through Z_d) constitutes the distance-3 planar logical qubit named Surface-17.

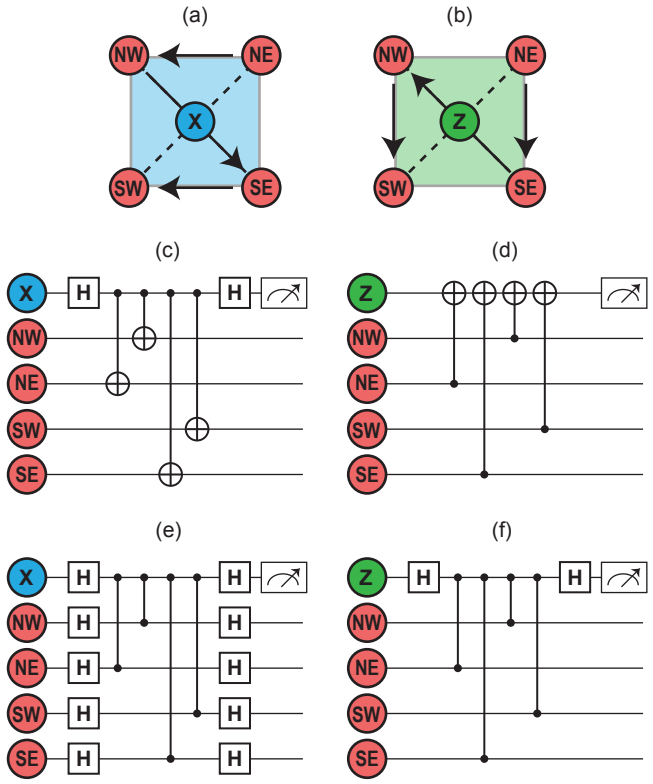

FIG. 2. X-type (a) and Z-type (b) plaquettes. Data qubits are labelled according to their position relative to the ancilla (NE=northeast, NW=northwest, SE=southeast, and SW=southwest). Standard circuits for measuring X-type (c, e) and Z-type (d, f) stabilizers indirectly using ancillas, using CNOT (c, d) or CZ (e, f) as the primitive data-ancilla interaction. The order of two-qubit gates, NE-NW-SE-SW (NE-SE-NW-SW) for X-type (Z-type), ensures that all data qubits common to adjacent plaquettes do their interactions with one ancilla before the other, and also provides resilience to ancilla errors in Surface-17 [31]. Using the relations $H = Y_{+90}Z = ZY_{-90}$, one can see that the opening and closing H gates can be replaced by Y_{-90} and Y_{+90} rotations, respectively.

II. BACKGROUND

A. Surface code QEC cycle

A surface-code fabric consists of the two-dimensional square lattice of data-carrying qubits shown in Fig. 1. The stabilizers of this code are the X-type (Z-type) parity operators $\prod_i X_i$ ($\prod_i Z_i$), where i denotes data qubits on the corners of the blue (green) plaquettes. Conventionally, these stabilizers are measured indirectly using ancilla qubits positioned at the center of the plaquettes, forming a second square lattice. Standard circuits for measuring X- and Z-type stabilizers, shown in Fig. 2, combine a sequence of coherent interactions of the ancilla with its nearest-neighbor data qubits, followed by projective ancilla measurement.

Using controlled-not (CNOT) gates as the fundamental interaction, X-type and Z-type stabilizer measurements can be fully parallelized with circuit depth seven. We define circuit depth as the number of operations on each ancilla per QEC cycle, counting in measurement but excluding ancilla initialization [we assume Pauli frame updating (PFU) [13, 30] is used for data and ancilla qubits]. The order of two-qubit gates in Fig. 2 is important for two reasons [31]. First, data qubits common to adjacent plaquettes must do all their interactions with one ancilla before the other. Second, the S (N) pattern for X-type (Z-type) stabilizers provides resilience to single ancilla-qubit errors even in small distance-three surface codes such as Surface-17. This circuit consists of the patch delineated in Fig. 1, with nine data qubits (labelled D_a to D_i), four ancillas (X_a to X_d) for X-type stabilizer measurements, and four ancillas (Z_a to Z_d) for Z-type stabilizer measurements.

When the two-qubit gate is CZ, parallelizing the stabilizer measurements of Surface-17 requires depth nine because of non-commutation between Hadamard (H) gates and CZ gates. The full circuit for the parallelized QEC cycle of Surface-17 using CZ gates is shown in Fig. 3. Using gate and measurement times from recent experiments ($\tau_\mathrm{1Q} = 20$ ns for single-qubit gates, $\tau_\mathrm{2Q} = 40$ ns for $\mathrm{CZ_f}$ gates, and 500 ns for ancilla readout and photon depletion in readout resonator), the QEC cycle will complete in 740 ns.

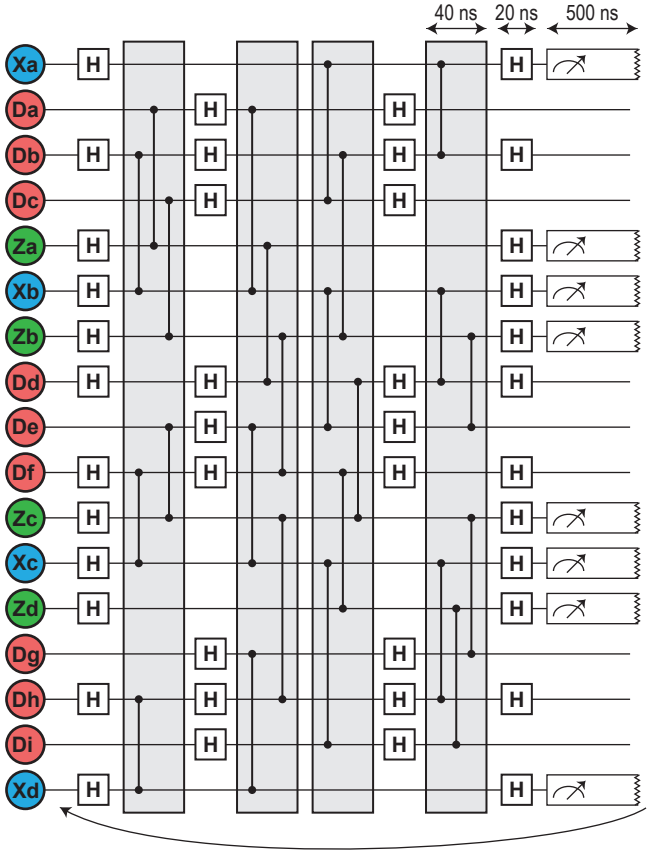

FIG. 3. Depth-nine quantum circuit for parallelized X- and Z-stabilizer measurements in Surface-17 using CZ gates. The six CZ gates inside each gray box are executed simultaneously. Typical values of gate and readout times are indicated at the top. The bottom arrow represents the looping of QEC cycles. Qubits are labelled as in Fig. 1.

B. Limitations of fully parallelized X- and Z-type stabilizer measurements using $\mathrm{CZ_f}$ gates

On paper, it is straightforward to compose a depth-nine quantum circuit for the fully parallelized QEC cycle of a surface-code fabric of arbitrary size. However, to the best of our knowledge following numerous failed attempts, the full parallelization of X- and Z-type stabilizer measurements makes it impossible to realize a scalable implementation with $\mathrm{CZ_f}$ gates that satisfies all of the following desirable properties:

- Microwave pulses for single-qubit gates should be applied at a fixed, small number of frequencies.
- Transmons should maximally exploit their coherence sweetspot [32].
- Flux-pulsed transmons should not cross any other interaction zones on their way to or from the intended $|11\rangle \leftrightarrow |02\rangle$ avoided crossings realizing the $\mathrm{CZ_f}$ gate.
- The flux-pulsing schemes should be extensible to a surface code of arbitrary size using a fixed number of detuning sequences and a fixed detuning range.
- The implementation should be compatible with logical qubit operations.

We have found frequency arrangements and flux-pulse sequences that meet the first three criteria. However, all of these solutions require a growing number of detuning sequences and detuning ranges as the fabric expands, in order to avert all other interactions on the way to and from the $|11\rangle \leftrightarrow |02\rangle$ avoided crossings of $\mathrm{CZ_f}$ gates. Furthermore, these solutions seem practically infeasible already for distance five (Surface-49 [29]). To our knowledge, no fully parallel solution exists with a fixed number of detuning sequences and a fixed detuning range. In the next section, we introduce a pipelined (rather than parallelized) version of the QEC cycle that simultaneously meets the five desirable properties for a fabric of arbitrary size.

III. THE PIPELINED QEC CYCLE

Our scalable scheme, which we term *pipelined QEC cycle*, combines four key elements:

A. Repeating unit cells of eight qubits;

B. Pipelined X- and Z-type stabilizer measurements;

C. Three frequencies for single-qubit control;

D. Eight detuning sequences implementing the requisite $\mathrm{CZ_f}$ gates, realizable by on/off masking of three flux-pulse primitives.

We now introduce these elements in detail.

A. Unit cell

The first element is a unit cell (Fig. 4) from which a surface code of arbitrary size can be assembled by repetition (and truncation at boundaries). A unit cell contains four data qubits (D_1 to D_4) and four ancillas (X_1, X_2, Z_1, and Z_2). Crucially, the cell is the fundamental unit of repetition not just for the quantum hardware. It is also the unit of repetition for all coherent control.

B. Pipelining of X-type and Z-type stabilizer measurements

The second element is the pipelined execution of the X- and Z-type stabilizer measurements. The pipelining concept is illustrated in Fig. 5(a). While stabilizer measurements of one type always run simultaneously, the coherent and readout steps of ancillas of the other type are

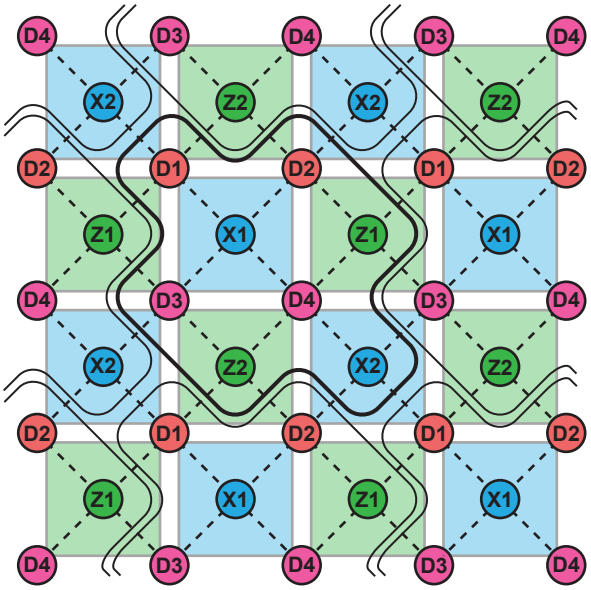

FIG. 4. Composing the surface-code fabric by repetition of 8-qubit unit cells. Red and pink circles represent data qubits, blue (green) circles represent ancillas for X-type (Z-type) stabilizer measurements, and dashed lines represent nearest-neighbor couplings. Dot colors also indicate the frequency for single-qubit microwave control (red for f_1, green and blue for f_2, and pink for f_3). Contours delineate unit cells (with qubits named D_1 to D_4, X_1, Z_1, X_2, and Z_2).

interleaved. In other words, ancillas of one type undergo coherent steps while ancillas of the other type are measured. Time slots A and B (D and E) are for single-qubit gates pertinent to the X-type (Z-type) stabilizer measurements, while slots 1 to 4 (5 to 8) are for two-qubit gates. Note that nine of the CZ gates involve two qubits within the cell, while fourteen involve one qubit from a neighboring unit cell.

Generally, ancilla measurement (including any photon depletion of the readout resonator) will take longer than the coherent steps, leaving time to perform operations on the data qubits in steps C and F while all ancillas are measured. Possible operations include logical gates, refocusing pulses, or single-qubit gates performing error correction. Clearly, performing such operations during steps C or F would not increase the QEC cycle time.

Pipelining offers several advantages. First, it compresses the stabilizer measurements to depth seven, two single-qubit-gate steps less than fully-parallelized quantum circuits (such as Fig. 3 for Surface-17). A second and more crucial advantage is the ability to scale without increasing the number of frequencies for single-qubit control or qubit detuning sequences, as explained next.

C. Single-qubit control and detuning sequences

The third and fourth elements are best described together. Figure 5(b) presents our choice of frequencies for single-qubit control and the qubit-specific detuning sequences for realizing the two-qubit QEC cycle interactions. Single-qubit gates on data qubits (steps A, B, D, and E) are performed at frequencies f_1 and f_3 (alternating in data-qubit rows), while those on ancillas are performed at intermediate frequency f_2. Note that with only nearest-neighbor coupling, two distinct frequencies (one for ancilla qubits and one for data qubits) reduce the exchange coupling between same-frequency qubits to fourth order (qubit-resonators, resonator-qubit, qubit-resonator, resonator-qubit). When extending to the proposed three frequencies, this also allows engineering the detuning sequences so that no transmon crosses any other second-order interaction zone on the way to or from the $|11\rangle \leftrightarrow |02\rangle$ avoided crossings exploited in the $\mathrm{CZ_f}$ gates.

During steps 1-4 and 5-8, transmons are flux pulsed to a discrete set of frequencies, depending on whether they interact, idle, or are measured: D_1 and D_2 to f_1 or f_1^{Int}; ancillas to f_2, f_2^{Park} or f_2^{Int}; and D_3 and D_4 to f_3 or f_3^{Park}. CZ gates occur between transmons at f_1^{Int} and f_2, and between transmons at f_2^{Int} and f_3. The exact value of the frequencies shall be chosen such that a compromise is reached between gate speed and residual interactions, which are due to the finite detuning when transmons are at f_1, f_2^{Park}, and f_3^{Park}.

In the case of perfect qubit manufacturing, same-labelled qubits would be identical for all unit cells. In that case, the frequency detuning patterns during interaction steps 1 through 4 and 5 through 8 can be synthesized by on/off masking of three flux-pulse primitives using a switch matrix: A first primitive detuning data qubits of type D_1 and D_2 from f_1 to f_1^{Int}, a second one detuning ancillas from f_2 to f_2^{Int}, and a third one detuning data qubits of type D_3 and D_4 from f_3 to f_3^{Park}. For example, the detuning sequence for D_2 in Fig. 5(b) can be synthesized by masking the pulse primitive on (off) at steps 1, 4, 6, and 7 (2, 3, 5, and 8).

D. Frequency arrangement variations

There exist other possible frequency arrangements than that shown in Fig. 5(b). For example, consider the inverted arrangement with all data qubits at f_2 and the ancillas at the outer frequencies. Figure 6 shows one of these configurations, with X_1 and Z_1 (X_2 and Z_2) at f_1 (f_3). It is straightforward to modify the detuning sequences for this arrangement to also avert all unwanted interactions. However, upon comparing this alternative to the original arrangement, we observe a key difference making the original preferable for a cQED implementation with flux-tuneable transmons. Specifically, the original exactly balances the number of interaction steps in which qubits can remain at their upper frequency (i.e., at or closest to their coherence sweetspot), while the flipped arrangement allows this on just two (out of eight) steps for data qubits and zero or four (out of four) steps for ancilla qubits. The reduced data-qubit dephasing during

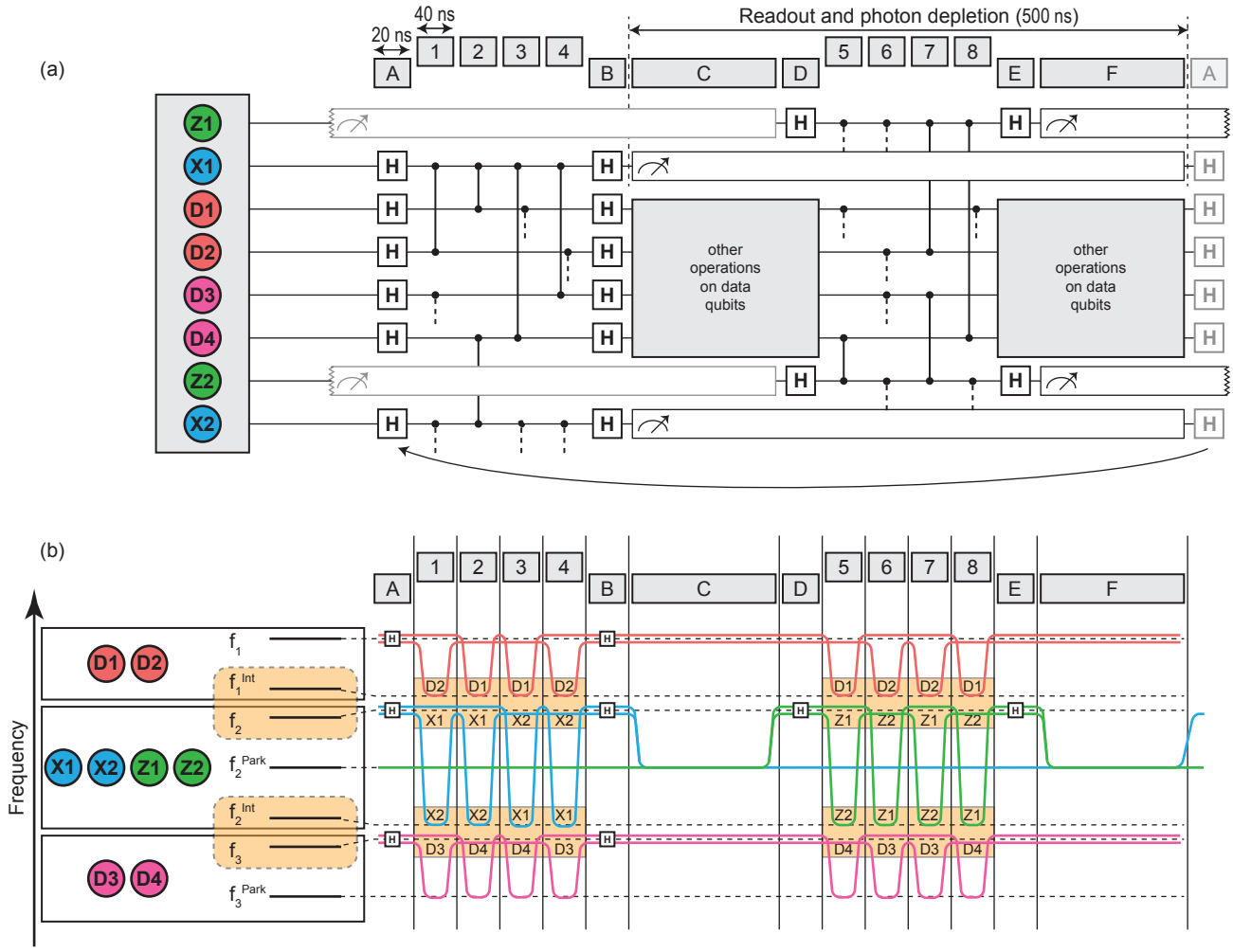

FIG. 5. (a) Unit cell quantum circuit for pipelined X- and Z-type stabilizer measurements. Qubits are labelled as in Fig. 4. Time slots 1 to 4 (5 to 8) are for X-type (Z-type) stabilizer CZ gates. Time slots A and B (D and E) are for X-type (Z-type) stabilizer single-qubit gates. Time slots C and F allow (optional) operations on data qubits while X-ancillas (Z-ancillas) are measured during time slots D through F (time slots A through C). The bottom arrow indicates the repetition of QEC cycles. Elements in transparent gray belong to the previous or next QEC cycle. (b) Frequency arrangement and detuning sequences for qubits in the unit cell: single-qubit gates on D_1 and D_2 (D_3 and D_4) are performed at f_1 (f_3), while those on ancillas are performed at f_2. Ancilla measurements are performed at $f_2^{\rm Park}$. $\mathrm{CZ_f}$ gates are performed between qubits at $f_1^{\rm Int}$ and f_2 or at $f_2^{\rm Int}$ and f_3. No interactions take place at f_1 or the parking frequencies $f_2^{\rm Park}$ and $f_3^{\rm Park}$. Small offsets are added to some detuning sequences to clarify the distinction between sequences for D_1 and D_2, X_1 and X_2, Z_1 and Z_2, and D_3 and D_4.

the coherent steps will lead to a lower logical error rate. Note that this advantage of the original arrangement is made possible by lowering the ancillas to $f_2^{\rm Park}$ for their measurement, at which the additional dephasing is innocuous in view of the measurement-induced projection.

To reduce residual single-qubit gate cross-talk between D_1 and D_2 (D_3 and D_4), another variation can be implemented by breaking the degeneracy in frequency f_1 (f_3), which requires increasing the number of primitive pulses from three to five, or even in f_2, further increasing the number of primitive pulses to eight.

IV. COMPATIBILITY WITH LOGICAL QUBIT OPERATIONS

Two types of logical qubits can be envisioned for surface code: defect-based [10] and planar [29]. Defect-based logical qubits are introduced by stopping the measurement of one or two stabilizers (X-type for *rough* logical qubits, and Z-type for *smooth* ones [10]). In our scheme, turning stabilizer measurements fully off can be accomplished in either of two ways. One is to mask off the H gates of the corresponding ancilla, without changing the detuning sequence or stopping the ancilla measurement. If the ancilla is in $|0\rangle$, all its $\mathrm{CZ_f}$ gates are inactive and

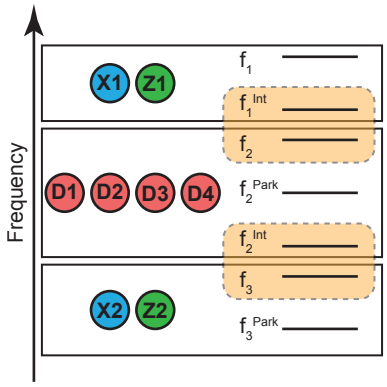

FIG. 6. Alternative frequency arrangement for qubits in the unit cell, with ancillas X_1 and Z_1 (X_2 and Z_2) at the outer frequency f_1 (f_3) and data qubits at the inner frequency f_2.

there is no net action on the logical qubit. If it starts in $|1\rangle$, the stabilizer operator (not its measurement) is applied. Although this performs a logical $X_{\rm L}$ ($Z_{\rm L}$) gate on a rough (smooth) qubit, the ancilla measurements allows to keep track of the action by Pauli frame updating. A second way to turn a stabilizer fully off is to mask off all the flux-pulse primitives in the interaction step, keeping the qubits at their sweetspot and minimizing flux noise. The corresponding H gates on the ancillas or data qubits could also be masked to further reduce qubit errors due to qubit control inaccuracies.

Logical operations, such as move and braiding operations on defect-based qubits [10], and lattice surgery on planar ones [29], also require dynamically changing the weight of specific stabilizer measurements, i.e., selectively removing specific data qubits from the quantum parity checks. In our scheme, this can easily be achieved by selective on/off masking of flux-pulse primitives. For example, removing a qubit of type D_2 from the X-type stabilizer measurement below it simply requires masking off the pulse primitive at step 1. The order of the two-qubit gates can also be changed by masking.

V. CONCLUSION AND OUTLOOK

We have presented a concrete scheme for the QEC cycle of an arbitary-size surface code implemented with flux-tuneable transmons. The scheme combines four key concepts: an eight-qubit unit cell as the basis for repetition of quantum hardware and control signals; pipelining of X- and Z-type stabilizer measurements; a fixed set of three frequencies for single-qubit control; and a fixed set of eight detuning sequences implementing the requisite controlled-phase gates. These eight detuning sequences can be composed by on-off masking of three flux-pulse primitives. We propose an implementation of this scheme with room temperature control systems to validate and test the method and its scalability, see Supplemental Material [33]. For the longer term a cryogenic implementation remains highly attractive.

ACKNOWLEDGEMENTS

We thank C. C. Bultink, M. A. Rol, X. Fu, C. García-Almudever, L. Riesebos, D. Deurloo, B. Criger, and T. E. O'Brien for helpful discussions. This research is supported by the Dutch Organisation for Scientific Research on Matter (FOM), Intel Corporation, and by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. Army Research Office grant W911NF-16-1-0071. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

[1] J. Kelly, R. Barends, A. Fowler, A. Megrant, E. Jeffrey, T. White, D. Sank, J. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.C. Hoi, C. Neill, P.J.J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A.N. Cleland, J.M. Martinis, State preservation by repetitive error detection in a superconducting quantum circuit, Nature **519**, 66 (2015).

[2] A. D. Corcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, Demonstration of a quantum error detection code using a square lattice of four superconducting qubits, Nat. Commun. **6**, 6979 (2015).

[3] D. Ristè, S. Poletto, M. Z. Huang, A. Bruno, V. Vesterinen, O. P. Saira, and L. DiCarlo, Detecting bit-flip errors in a logical qubit using stabilizer measurements, Nat. Commun. **6**, 6983 (2015).

[4] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, Demonstration of a small programmable quantum computer with atomic qubits, Nature **536**, 63 (2016).

[5] J. Cramer, N. Kalb, M. A. Rol, B. Hensen, M. Markham, D. J. Twitchen, R. Hanson, and T. H. Taminiau, Repeated quantum error correction on a continuously encoded qubit by real-time feedback, Nat. Commun. **5**, 11526 (2016).

[6] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).

[7] J. M. Martinis, Qubit metrology for building a fault-tolerant quantum computer, npj Quantum Inf. **1**, 15005 (2016).

[8] K. R. Brown, J. Kim, and C. Monroe, Co-designing a scalable quantum computer with trapped atomic ions, npj Quantum Inf. **2**, 16034 (2016).

[9] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, Freely scalable quantum technologies using cells of 5-to-50 qubits with very lossy and noisy photonic links, Phys. Rev. X **4**, 041041 (2014).

[10] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[11] N. C. Jones, R. Van Meter, A. G. Fowler, P. L. McMahon, J. Kim, T. D. Ladd, and Y. Yamamoto, Layered architecture for quantum computing, Phys. Rev. X **2**, 031007 (2012).

[12] S. B. Bravyi and A. Y. Kitaev, Quantum codes on a lattice with boundary, arXiv:quant-ph/9811052 (1998).

[13] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[14] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, Phys. Rev. A **69**, 062320 (2004).

[15] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, Superconducting quantum circuits at the surface code threshold for fault tolerance, Nature **508**, 500 (2014).

[16] M. A. Rol, C. C. Bultink, T. E. O'Brien, S. R. de Jong, L. S. Theis, X. Fu, F. Luthi, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, D. Deurloo, R. N. Schouten, F. K. Wilhelm, and L. DiCarlo, Restless tuneup of high-fidelity qubit gates, Phys. Rev. Applied **7**, 041001 (2017).

[17] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, J.M. Gambetta, Efficient Z-gates for quantum computing, ArXiv:1612:00858 (2016).

[18] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, Feedback control of a solid-state qubit using high-fidelity projective measurement, Phys. Rev. Lett. **109**, 240502 (2012).

[19] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, Fast accurate state measurement with superconducting qubits, Phys. Rev. Lett. **112**, 190504 (2014).

[20] A. Bruno, S. Poletto, N. Haider, and L. DiCarlo, "X48.00004: Extensible circuit QED processor architecture with vertical I/O," APS March Meeting 2016.

[21] J. H. Béjanin, T. G. McConkey, J. R. Rinehart, C. T. Earnest, C. R. H. McRae, D. Shiri, J. D. Bateman, Y. Rohanizadegan, B. Penava, P. Breul, S. Royak, M. Zapatka, A. G. Fowler, and M. Mariantoni, Three-dimensional wiring for extensible quantum computing: The quantum socket, Phys. Rev. Applied. **6**, 044010 (2016).

[22] D. Rosenberg, D. R. Yost, R. Das, D. Hover, L. Racz, S. Weber, J. Yoder, A. Kerman, and W. D. Oliver, "X48.00003: 3D integration for superconducting qubits," APS March Meeting 2016.

[23] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, Partial-measurement backaction and nonclassical weak values in a superconducting circuit, Phys. Rev. Lett. **111**, 090506 (2013).

[24] S. Asaad, C. Dickel, S. Poletto, A. Bruno, N. K. Langford, M. A. Rol, D. Deurloo, and L. DiCarlo, Independent, extensible control of same-frequency superconducting qubits by selective broadcasting, npj Quantum Inf. **2**, 16029 (2016).

[25] F. W. Strauch, P. R. Johnson, A. J. Dragt, C. J. Lobb, J. R. Anderson, and F. C. Wellstood, Quantum logic gates for coupled superconducting phase qubits, Phys. Rev. Lett. **91**, 167005 (2003).

[26] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blas, S. M. Frunzio, L. an Girvin, and R. J. Schoelkopf, Demonstration of two-qubit algorithms with a superconducting quantum processor, Nature **460**, 240 (2009).

[27] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Preparation and measurement of three-qubit entanglement in a superconducting circuit, Nature **467**, 574 (2010).

[28] J. M. Martinis, Qubit metrology for building a fault-tolerant quantum computer, npj Quantum Inf. **1**, 15005 (2015).

[29] C. Horsman, A. G. Fowler, S. Devitt, and R. V. Meter, Surface code quantum computing by lattice surgery, New J. Phys. **14**, 123011 (2012).

[30] E. Knill, Quantum computing with realistically noisy devices, Nature **434**, 39 (2005).

[31] Y. Tomita and K. M. Svore, Low-distance surface codes under realistic quantum noise, Phys. Rev. A **90**, 062320 (2014).

[32] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Suppressing charge noise decoherence in superconducting charge qubits, Phys. Rev. B **77**, 180502 (2008).

[33] See supplemental material at [URL will be inserted by publisher] for a proposed implementation, which includes Refs. [34–41].

[34] T. E. O'Brien, B. Tarasinski, L. DiCarlo, Density-matrix simulation of small surface codes under current and projected experimental noise, arXiv:quant-ph/1703.04136 (2017).

[35] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[36] O. A. Mukhanov, Energy-efficient single flux quantum technology, IEEE Trans. Appl. Supercond. **21**, 760 (2011).

[37] J. M. Chow, J. I. Gambetta, E. Magesan, D. W. Abraham, A. W. Cross, B. R. Johnson, N. A. Masluk, C. A. Ryan, J. A. Smolin, S. J. Srinivasan, and M. Steffen, Implementing a strand of a scalable fault-tolerant quantum computing fabric, Nat. Commun. **5**, 4015 (2014).

[38] J. M. Hornibrook, J. I. Colless, I. D. Conway Lamb, S. J. Pauka, H. Lu, A. C. Gossard, J. D. Watson, G. C. Gardner, S. Fallahi, M. J. Manfra, and D. J. Reilly, Cryogenic control architecture for large-scale quantum computing, Phys. Rev. Applied **3**, 024010 (2015).

[39] S. Poletto *et al.*, in preparation.

[40] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[41] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Optimized driving of superconducting artificial atoms for improved single-qubit gates, Phys. Rev. A **82**, 040305 (2010).