

Crosstalk in multiqubit fluxonium architectures with transmon couplers

Martijn F. S. Zwanenburg^{*} and Christian Kraglund Andersen[†]
*QuTech and Kavli Institute of Nanoscience, Delft University of Technology,
 2628 CJ, Delft, The Netherlands*

 (Received 31 March 2026; revised 29 June 2026; accepted 27 July 2026; published 31 August 2026)

In recent years, several architectures have been proposed for implementing two-qubit operations on fluxonium superconducting qubits. A particularly promising approach, which was demonstrated experimentally by Refs. [L. Ding *et al.* *Phys. Rev. X* **13**, 031035 (2023), S. Singh *et al.* *Phys. Rev. Appl.* **25**, 024020 (2026)], employs a transmon superconducting qubit as a tunable coupler between the fluxonium qubits. These experiments have shown that the transmon coupler enables fast, high-fidelity two-qubit operations while suppressing unwanted ZZ crosstalk between the fluxonium qubits. In this work, we numerically study the scalability of this architecture. We find that, when trivially scaling this architecture, crosstalk from spectator qubits limits the gate fidelity to below 90%. We show that these spectator errors can be reduced to below 10^{-4} by reducing the coupling strength and by dynamically tuning transmons that are not used for a two-qubit operation to an off position. We further investigate the resilience of the operation to direct capacitive coupling between the transmon couplers and to microwave crosstalk.

DOI: [10.1103/rf7g-md41](https://doi.org/10.1103/rf7g-md41)

I. INTRODUCTION

Recent progress with two-qubit operations [1–7], single-qubit operations [8,9], reset [10], and readout [11,12] on fluxonium superconducting qubits [13] has made this qubit a promising building block for larger-scale quantum devices. One key area that has not yet received sufficient attention is the scaling of fluxonium qubits beyond two-qubit architectures, and in this work, we are specifically interested in scaling the fluxonium-transmon-fluxonium (FTF) architecture realized experimentally by Refs. [1,2]. In larger-scale systems, the frequencies of one FTF pair may depend on the state of other fluxonium qubits (spectators), which gives rise to crosstalk errors [14–16]. In this work, we numerically investigate whether larger-scale systems, such as building blocks for 2D surface codes [17–20] in which a two-qubit operation may be influenced by six nearest-neighbor fluxoniums, can effectively suppress these crosstalk errors. Previous studies of crosstalk errors on larger-scale superconducting systems were focused on transmon qubits [21], did not explicitly analyze the two-qubit gate errors arising in such

systems [22], or restricted their analysis to the influence of one spectator [7].

In the FTF architecture, a transmon qubit [23] is used as a tunable coupling element between two fluxonium qubits. The transmon suppresses unwanted ZZ coupling between the fluxoniums, while at the same time enabling the implementation of a conditional phase gate (CZ). The ZZ suppression is realized by balancing the direct coupling between the fluxoniums and the indirect coupling mediated by the transmon [1,2]. The two-qubit operation is applied by conditionally driving a 2π rotation between a computational and noncomputational state. During this operation, the state picks-up a geometrical phase of π , as required for the CZ gate. Given that the two-qubit gate relies directly on the resonant driving of a specific transition, it is clear that frequency shifts induced by nearby spectator qubits may indeed cause gate errors. In this setting, we are interested in the spectator-dependent transition frequencies of the transmon coupler. We study the gate errors that arise from spectator effects in the 1D and 2D systems shown in Fig. 1 with system parameters inspired by the experimental implementations by Refs. [1,2]. Our results show that large crosstalk errors arise in these systems that limit the two-qubit gate fidelity to below 90%. On the other hand, by dynamically tuning transmon couplers that are not actively used for a two-qubit operation down in frequency, we show that these crosstalk errors can be suppressed to below 10^{-4} .

In the remainder of this work, we first summarize the most relevant properties of the FTF system in Sec. II. Next,

^{*}Contact author: m.f.s.zwanenburg@tudelft.nl

[†]Contact author: c.k.andersen@tudelft.nl

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

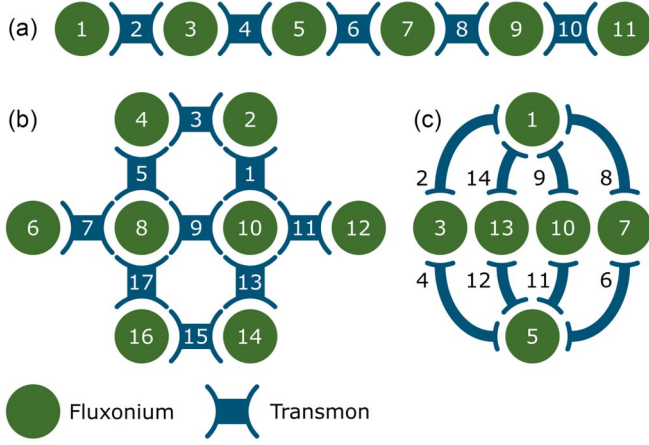


FIG. 1. Schematic drawing of the systems studied in this work. The numbers indicate the site numbering used throughout the main text. (a) 1D system, (b) 2D grid system, and (c) 2D system for the $[[4, 2, 2]]$ error correcting code with two auxiliary qubits.

in Sec. III we study the crosstalk errors in the systems shown in Fig. 1. Finally, in Sec. IV we analyze the effect of direct capacitive coupling between the transmon couplers and microwave crosstalk on the two-qubit gate operation.

II. FTF ARCHITECTURE

We consider the case where the fluxonium qubits are capacitively coupled to the transmon coupler, see Fig. 2(a). The individual fluxonium qubits are described by their Hamiltonian

$$\hat{H}_0^{F_i} = 4E_C^{F_i} \hat{n}_{F_i}^2 + \frac{1}{2} E_L^{F_i} \hat{\phi}_{F_i}^2 - E_J^{F_i} \cos(\hat{\phi}_{F_i} - 2\pi\varphi_{\text{ext}}^{F_i}), \quad (1)$$

where $E_C^{F_i}$, $E_L^{F_i}$, and $E_J^{F_i}$ are the capacitive, inductive, and Josephson energy of fluxonium i , respectively, and where $\hat{n}_{F_i}$ and $\hat{\phi}_{F_i}$ are the charge and phase operators of the fluxonium. Within this work, we only consider the case where the fluxonium qubits are flux-biased to their sweet spot, i.e., the reduced external flux is fixed at $\varphi_{\text{ext}}^{F_i} = 0.5$. The transmon is similarly described in terms of its charge operator $\hat{n}_T$ and phase operator $\hat{\phi}_T$ by the Hamiltonian

$$\hat{H}_0^T = 4E_C^T \hat{n}_T^2 - E_J^T(\varphi_{\text{ext}}^T) \cos(\hat{\phi}_T), \quad (2)$$

where E_C^T is the fixed charging energy and where

$$E_J^T(\varphi_{\text{ext}}^T) = (E_{J1}^T + E_{J2}^T) \sqrt{\cos(\pi\varphi_{\text{ext}}^T)^2 + d^2 \sin(\pi\varphi_{\text{ext}}^T)^2}, \quad (3)$$

is the effective Josephson energy of the transmon's SQUID loop that is threaded by a reduced external flux φ_{ext}^T [23]. Here, $d = |E_{J1}^T - E_{J2}^T| / (E_{J1}^T + E_{J2}^T)$ denotes the SQUID

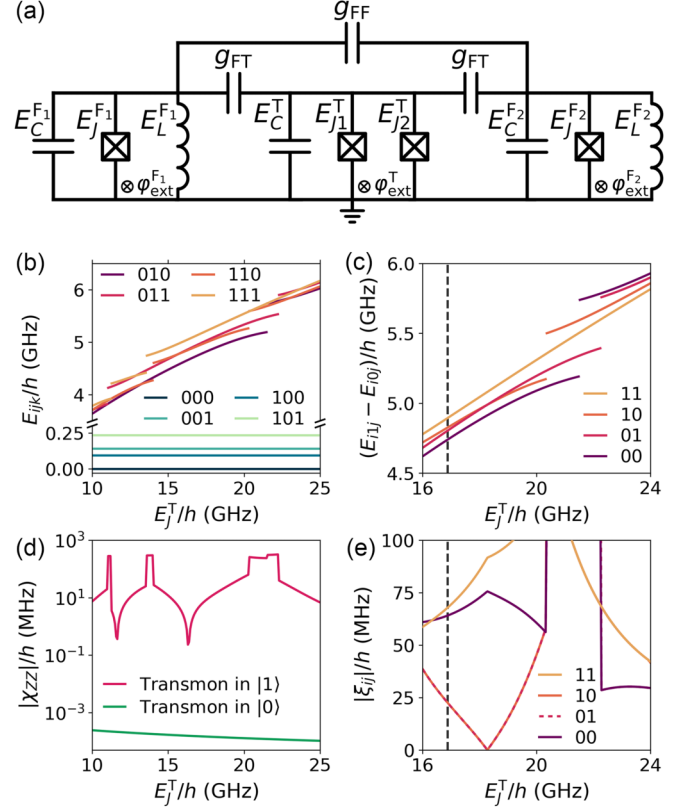


FIG. 2. (a) Schematic circuit diagram of the three-node FTF system. (b) Fluxonium and transmon energy levels as a function of E_J^T/h . The green-shaded and red-shaded lines correspond to states with the transmon in its ground state and first excited state, respectively. (c) Transition energies of the transmon's qubit level as a function of E_J^T/h . (d) Residual ZZ coupling between the fluxonium qubits with the transmon in the ground or excited state. (e) Minimum frequency separation of the transmon's qubit transition. The vertical dashed line in panels (c), (e) indicates the E_J^T experimentally realized in Ref. [2].

asymmetry of the transmon. Given the capacitive coupling between the systems, the coupling Hamiltonians are given by

$$\begin{aligned} \hat{H}_{\text{FT}} &= \hbar g_{\text{FT}} (\hat{n}_{F_1} \hat{n}_T + \hat{n}_T \hat{n}_{F_2}), \\ \hat{H}_{\text{FF}} &= \hbar g_{\text{FF}} \hat{n}_{F_1} \hat{n}_{F_2}, \end{aligned} \quad (4)$$

where g_{FT} is the coupling strength between the transmon and each fluxonium, and g_{FF} is the coupling strength between the fluxoniums. The total Hamiltonian is given by

$$\hat{H} = \sum_{i \in \{F_1, T, F_2\}} \hat{H}_0^i + \hat{H}_{\text{FT}} + \hat{H}_{\text{FF}}. \quad (5)$$

In this work, we will denote eigenstates of this three-node Hilbert space as $|F_1 T F_2\rangle$ with corresponding eigenenergies $E_{F_1 T F_2}$.

When using the system parameters realized by Ref. [2], the spectrum of the low-energy subspace readily reveals

the four computational levels, while we see a higher energy subspace consisting of four noncomputational states when the transmon is excited, see Fig. 2(b). As E_J^T increases, the resulting transmon transition frequencies become resonant with the $|1\rangle \leftrightarrow |2\rangle$ and the $|0\rangle \leftrightarrow |3\rangle$ transitions of the fluxoniums, which we see in Fig. 2(b) as avoided crossings in the noncomputational states. Avoided crossing with the $|0\rangle \leftrightarrow |2\rangle$ and $|1\rangle \leftrightarrow |3\rangle$ transitions of the fluxoniums do not appear, as these transitions are parity forbidden when the fluxonium is biased to its flux symmetry point. Notice that, for simplicity, we have omitted the explicit external flux dependence of E_J^T . In particular, we see in Fig. 2(c) that the transmon transition frequencies are different for each of the four possible two-qubit states of the fluxoniums, which enables a microwave-activated conditional phase gate between the fluxonium qubits. The vertical, dashed line in Fig. 2(c) marks the operating point of Ref. [2]. The residual ZZ coupling between the fluxonium qubits, defined as $\chi_{ZZ} = E_{1i1} - E_{0i1} - E_{1i0} + E_{0i0}$, remains below 1 kHz for a wide range of transmon frequencies provided that the transmon is in its ground state, see Fig. 2(d). When the transmon is in $|1\rangle$ the ZZ coupling is significantly larger as a result of the coupling between the noncomputational states.

It is important to note that, in principle, different choices exist with regards to the exact transition that is driven for the conditional 2π rotation that implements the CZ operation. For instance, in Refs. [1,2] the gate was implemented by driving the $|101\rangle \leftrightarrow |111\rangle$ transition, but the gate can also be implemented by driving the $|101\rangle \leftrightarrow |201\rangle$ or $|101\rangle \leftrightarrow |102\rangle$ transitions as also done in Ref. [11]. In this work, we will focus on implementing the CZ by driving a transmon transition, i.e., $|i0j\rangle \leftrightarrow |i1j\rangle$. As also discussed in Ref. [2], the speed of the operation is limited by residual, undesired driving of other transitions that are close in energy to the targeted transition, which results in leakage. When driving a transmon transition $|i0j\rangle \leftrightarrow |i1j\rangle$, there are numerous leakage transitions that can be driven off-resonantly, for instance the other three transmon transitions shown in Fig. 2(c). Therefore, a key parameter that influences the speed of the operation is the minimum frequency separation ξ_{ij}

$$\xi_{ij} = \min_{kl \neq ij} |(E_{k1l} - E_{k0l}) - (E_{i1j} - E_{i0j})|, \quad (6)$$

which specifies the minimum frequency difference of the targeted transmon transition to the other transmon transitions. A larger minimum frequency separation indicates smaller leakage errors and hence a potentially faster gate. In the experimental implementations in Refs. [1,2], it was irrelevant which transition, specifically, has the smallest detuning to the conditionally driven transition, as it did not influence the employed mitigation strategies that correct for the leakage errors. In Fig. 2(e), we see that the largest minimum frequency separation for the $|101\rangle \leftrightarrow |111\rangle$

transition is achieved when tuning the transmon as close as possible to the avoided crossings. However, ξ_{ij} does not provide a complete picture of the total leakage in the operation. For instance, tuning too close to the avoided crossings would result in leakage to the third excited state of the fluxonium through the transitions $|i00\rangle \leftrightarrow |i03\rangle$ and $|00j\rangle \leftrightarrow |30j\rangle$. Furthermore, the extent to which a leakage transition is driven not only depends on the frequency difference, but also on how strongly it couples to the drive operator.

In this two-qubit gate scheme, the residual ZZ interaction remains suppressed at the operating point. Therefore, in the experimental implementations of Refs. [1,2] it was not required to dynamically flux-tune the transmon from an *off* to an *on* point, as often required in tunable-coupler schemes [24,25]. A potential downside is that the interaction between the noncomputational states is always on, which may lead to crosstalk errors.

III. SPECTATOR ERRORS

To determine the spectator errors that arise when scaling-up the systems in Refs. [1,2], we first study the 1D architecture shown in Fig. 1(a) consisting of six fluxoniums and five transmon couplers. We consider the application of a two-qubit operation on the center FTF pair, such that there are two spectator fluxoniums to the left and right. The parameters of the center FTF pair are set to the parameters realized in Refs. [1,2]. The parameters for the other fluxoniums and transmons are chosen in a ladder-like fashion, such that each pair of fluxonium qubits has a similar detuning, see Figs. 3(a), and 3(b). Note that this approach does not scale indefinitely since there may be a practical upper or lower bound to the qubit frequencies. The precise system parameters used for the simulations are detailed in Appendix A. We highlight that the key difference between the parameter regimes used in Refs. [1,2] is the coupling strengths, which are $g_{FT}/2\pi = 550$ MHz and $g_{FT}/2\pi = 236$ MHz, respectively.

To analyze the spectator errors, we first need to compute the eigenvalues and eigenvectors of the system. We compute the spectrum for each individual fluxonium and transmon and truncate their Hilbert spaces to the lowest six and three levels, respectively. For the 1D system, the dimension of the resulting combined truncated Hilbert space is approximately 11×10^6 , which is roughly equivalent to the Hilbert space dimension of 23 qubits. It is therefore highly challenging to compute the spectrum with exact diagonalization, and instead we employ a tensor network algorithm in which the eigenstates are represented as a matrix-product state (MPS) [26,27]. Specifically, we make use of the density matrix renormalization group for excited states (DMRG-X) algorithm [28], which is a variational algorithm in which each site of the MPS is updated iteratively until it converges to an eigenstate of the system.

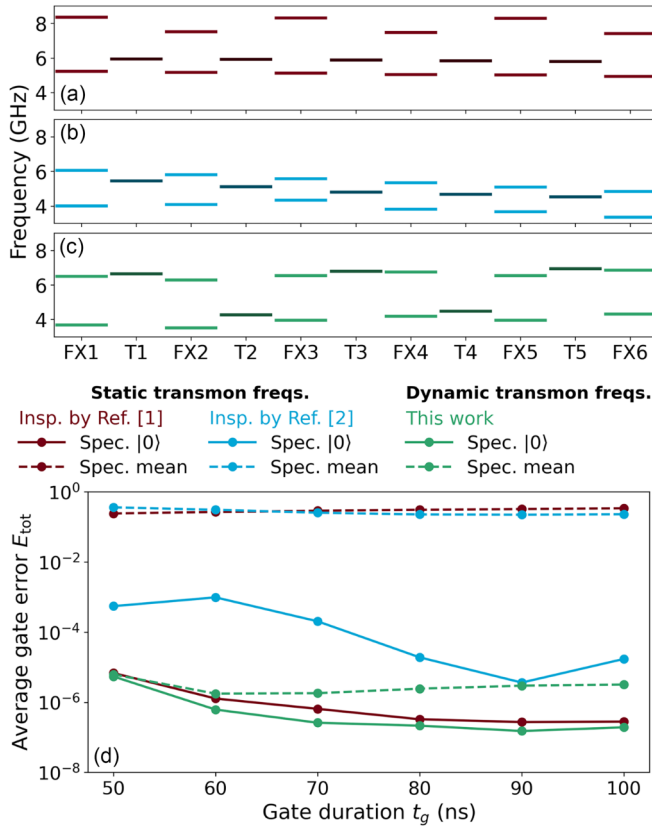


FIG. 3. (a)–(c) Energy level diagrams of the 1D systems with the parameters inspired by Refs. [1,2] and for the parameters proposed in this work, respectively. For the fluxoniums, the bare frequencies of the $|1\rangle \leftrightarrow |2\rangle$ and $|0\rangle \leftrightarrow |3\rangle$ transitions are shown, and for the transmons the bare frequency of the qubit transition is shown. (d) Spectator errors in a 1D system consisting of six fluxoniums and five transmon couplers as a function of the gate duration t_g . For each system, the two-qubit operation is calibrated with all spectators in $|0\rangle$ (solid lines), and then, the error is averaged over the four spectator states with $s_l, s_r \in \{0, 1\}$ (dashed lines). In red and blue, we plot the errors for the parameters inspired (Insp.) by Refs. [1,2], respectively, and in green we plot the results for the parameters proposed in this work. For the parameters inspired by Refs. [1,2], all transmons are statically positioned at their on point, while for the parameters proposed in this work the frequencies of inactive transmons are dynamically tuned to an off position.

The full implementation of this algorithm is detailed in Appendix B.

In the FTF architecture, spectator errors arise through the dependence of the frequency and drive matrix element of the conditionally driven transition and of the dominant leakage channels on the state of spectator qubits. To quantify the resulting crosstalk from spectator qubits in terms of gate errors, we numerically solve the time evolution of the two-qubit gate. Due to the size of the system and the strong interactions, it is infeasible to time evolve the full system in a tensor network formalism, and

instead we perform a truncation to the center FTF pair where the operation is performed. First, we transform the three-node Hamiltonian in Eq. (5) to the eigenframe and add a drive Hamiltonian to obtain

$$\begin{aligned} \hat{H}/\hbar &= \sum_i \omega_i |i\rangle\langle i| + \hat{H}_D/\hbar, \\ \hat{H}_D/\hbar &= \Omega \mathcal{E}_I(t) \cos(\omega_d t) \hat{\tilde{D}}, \end{aligned} \quad (7)$$

where ω_i specifies the energy of each level, Ω is the drive strength, $\mathcal{E}_I(t)$ is the pulse envelope, ω_d is the drive frequency, and $\hat{\tilde{D}}$ is the drive operator. Here, we drive the operation through the transmon's charge operator and neglect microwave crosstalk, such that $\hat{\tilde{D}} = \hat{n}_T$. The $\tilde{\cdot}$ denotes that the operator has been transformed to the eigenbasis. Using the DRMG-X algorithm, we compute the eigenenergies ω_i and we can further calculate the matrix elements of the drive operator $\hat{\tilde{D}}$ in the full 1D system. As we expect the spectator effects to be dominated by changes in the transition frequency and drive matrix elements of the transmon's qubit transition, we only compute the energies for the states $|\psi(s_l, i, k, j, s_r)\rangle = |s_l 0 s_r 0 i k j 0 s_r 0 s_r\rangle$ and matrix elements $\langle \psi(s_l, i, 1, j, s_r) | \hat{\tilde{D}} | \psi(s_l, i, 0, j, s_r) \rangle$ with $i, j, k \in \{0, 1\}$ for a specific configuration of the left (s_l) and right (s_r) spectator fluxoniums. The notation of the kets $|\cdot\rangle$ follows the site ordering of the 1D system shown in Fig. 1(a). For the systems with the parameters inspired by Refs. [1,2], strong delocalization of the eigenvectors causes the DRMG-X algorithm to be computationally expensive, see also Appendix B. To alleviate the required computational resources, we assume that the left (right) spectators are in the same state s_l (s_r). Using these DRMG-X results, we can readily reconstruct the Hamiltonian in Eq. (7) to obtain the effective three-node Hamiltonian for the center FTF pair for a certain spectator configuration. The resulting system is solved numerically to obtain the time evolution U . We highlight that this approach neglects any leakage outside the center FTF pair and neglects any spectator-dependent leakage to the second and third excited states of the transmon and fluxoniums. For systems where the eigenstates delocalize strongly, this approach may therefore only provide a lower bound on the actual crosstalk errors.

We calculate the gate error for the 1D system with the system parameters inspired by Refs. [1,2] as a function of the gate duration t_g to elucidate the trade-off between fast driving and potential errors arising from nearby off-resonant transitions, see Fig. 3(d). Specifically, we calculate the total gate error as $E_{\text{tot}} = E_{\text{phase}} + E_{\text{leakage}}$, where we have defined the coherent phase error E_{phase} and the average leakage error E_{leakage} as

$$E_{\text{phase}} = \frac{3}{10}(1 - \cos(\tilde{\theta}_c)),$$

$$E_{\text{leakage}} = 1 - \frac{1}{4} \sum_{\psi=000,001,100,101} |\langle \psi | U | \psi \rangle|^2. \quad (8)$$

Here, $\tilde{\theta}_c = \theta_c - \pi$ is the conditional phase error, with θ_c the conditional phase defined as

$$\theta_c = \text{Arg}(\langle 101 | U | 101 \rangle) - \text{Arg}(\langle 001 | U | 001 \rangle) \\ - \text{Arg}(\langle 100 | U | 100 \rangle) + \text{Arg}(\langle 000 | U | 000 \rangle), \quad (9)$$

where $\text{Arg}(\cdot)$ denotes the argument. We implement the CZ gate by driving the $|101\rangle \leftrightarrow |111\rangle$ transition with a cosine pulse defined as

$$\mathcal{E}_I(t) = \frac{1}{2} \left(1 - \cos\left(\frac{2\pi}{t_g} t\right) \right). \quad (10)$$

Importantly, we numerically calibrate the drive frequency ω_d and drive strength Ω to minimize the gate error when all spectator qubits are in the state $|0\rangle$. More details regarding the drive pulse and calibration are discussed in Appendix C. Next, we calculate the error averaged over four different spectator states $s_l, s_r \in \{0, 1\}$. We observe that, while the error with the spectators in $|0\rangle$ is low, the error averaged over the spectator states is above 0.1 for all gate durations. The individual contributions of the phase and leakage errors to the total error are plotted in Appendix C, from which we conclude that, while the phase errors are the limiting error source, both leakage and phase errors are above or around 1%.

The results in Fig. 3(d) show that trivially scaling the parameters of Refs. [1,2] is not feasible due to large spectator errors. We expect these errors to increase even further when considering a 2D architecture, where the two-qubit operation is subject to six nearest-neighbor spectator qubits instead of the two nearest-neighbor spectator qubits in 1D. Note that we have only considered a simple pulse shape here and we would expect that pulse shaping techniques such as derivative removal by adiabatic gate (DRAG) and related techniques [29–31] could reduce errors for the calibrated gate with all spectators in $|0\rangle$ [2]. Importantly, however, we do not expect such pulse shaping techniques to reduce spectator errors since they are not designed to be robust against changes in the primary transition frequency.

Next, our goal is to suppress the spectator errors. We recall that the spectator errors arise from the spectator dependence of the frequency and matrix elements of the transitions relevant for the operation. For the 1D systems with the parameters inspired by Refs. [1,2], the frequency of the $|101\rangle \leftrightarrow |111\rangle$ transition changes by 8.6 and 5.7 MHz, respectively, when the spectator states change from $s_l = s_r = 0$ to $s_l = s_r = 1$. We readily understand

that the large spectator errors arise from the transmon-like modes being delocalized over multiple couplers. Using perturbation theory, it can be shown that the spectator dependence of this delocalization scales as $g_{\text{FT}}^5/\Delta_{\text{TT}}^2$, where Δ_{TT} is the detuning between nearest transmons. Thus, we will achieve the suppression by reducing the coupling g_{FT} and by increasing the frequency spacing between the transmon couplers. The full parameter regime, which we found to reduce the spectator effects, is shown schematically in Fig. 3(c) and detailed further in Appendix A. Most notably, we reduce the coupling to $g_{\text{FT}}/2\pi = 200$ MHz, and we set $g_{\text{FF}}/2\pi = 20$ MHz to suppress the static ZZ coupling to below 0.2 kHz for all pairs. Furthermore, we dynamically tune the frequency of inactive transmons, i.e., transmons that are not actively used to perform a two-qubit gate, away from the $|0\rangle \leftrightarrow |3\rangle$ transition of the fluxoniums to suppress the residual coupling between these levels. Specifically, we tune these transmons down in frequency to approximately 300 MHz above the fluxoniums' $|1\rangle \leftrightarrow |2\rangle$ transition. Experimentally, this can be realized by flux-pulsing the transmon. For the 1D system, only the transmons that sit directly next to the center FTF pair are tuned to this off point, such that simultaneous two-qubit operations can be performed on the middle and two outer FTF pairs. We highlight that, at their off position, the transmon couplers still suppress the ZZ interaction between the fluxoniums to below 0.2 kHz. Finally, to maximize the frequency spacing between active and inactive transmons, we park the active transmons above the $|0\rangle \leftrightarrow |3\rangle$ transitions of the fluxoniums. Caution must be taken at this point, since the dynamical flux tuning of the transmons in this parameter regime means that the transmon's $|0\rangle \leftrightarrow |1\rangle$ transition will cross the $|0\rangle \leftrightarrow |3\rangle$ transitions of the fluxoniums. However, we emphasize that since the transmon $|1\rangle$ state and the fluxonium $|3\rangle$ states are both noncomputational states, and therefore unoccupied, we do not expect residual errors from this crossing.

In this improved parameter regime, the dependence of the frequency of the conditionally driven transition on the state of spectator fluxoniums is only approximately 10 kHz, which is an improvement of two orders of magnitude when compared to the systems inspired by Refs. [1,2]. As a result, the error averaged over the spectator states shown in Fig. 3(d) remains below 10^{-5} . Crucially, in this case we have again numerically calibrated the gate only when the spectator qubits are in the $|0\rangle$ state. As shown in Appendix B, the eigenstates remain well localized at the center FTF pair in this parameter regime, and we therefore expect negligible leakage to states outside the center FTF pair. In this parameter regime, the minimum frequency separation ξ_{ij} is only 42 MHz, while for the parameters used in Refs. [1,2] the minimum frequency separation is 152 and 68 MHz, respectively. Naturally, this results in a larger leakage rate. To suppress this leakage, we make use

of the hybridization structure of the eigenstates. We apply the gate by driving the $|000\rangle \leftrightarrow |010\rangle$ transition through the charge operator of one of the fluxoniums instead of through the transmon's charge operator. We always drive the fluxonium, which has the largest detuning between its $|0\rangle \leftrightarrow |3\rangle$ transition and the transmon's qubit transition, which, for the purpose of explaining how the operation works, we assume to be the left fluxonium. That implies we set $\hat{D} = \hat{n}_{F1}$ in Eq. (7). By driving the left fluxonium, the leakage channels $|10i\rangle \leftrightarrow |11i\rangle$ are naturally suppressed, as they do not hybridize with the left fluxonium and therefore do not couple through its charge operator. To suppress the leakage channel $|001\rangle \leftrightarrow |011\rangle$, we will employ the DRAG technique [29,30], which is applicable here since we aim to suppress a single nearly resonant transition frequency in the spectral density of the pulse. We further suppress leakage of $|100\rangle \leftrightarrow |110\rangle$ by adding a small linear compensation drive to the right fluxonium. Full details regarding this drive scheme and its calibration are detailed in Appendix C. We reiterate that, in principle, the operation can also be performed by selectively driving the $|1\rangle \leftrightarrow |2\rangle$ or $|0\rangle \leftrightarrow |3\rangle$ transitions of the fluxonium [1]. However, for the system parameters proposed in this work, the conditional frequency of these transitions is typically smaller than for the transmon transition, such that we would expect the (spectator) errors to be larger when driving these other transitions.

Next, we analyze the crosstalk errors in two 2D qubit architectures. We study crosstalk errors from spectator qubits in a 2D grid that could be used as a building block for error correcting surface codes [17–20], and a 2D architecture that can be used for a $[[4, 2, 2]]$ error correcting code with two auxiliary qubits [32,33]. These architectures are shown in Figs. 1(b) and 1(c), respectively, and the full set of parameters used in this work is detailed in Appendix A. For both architectures, we also set $g_{FT}/2\pi = 200$ MHz. As also done for the 1D architecture, we tune inactive transmons down in frequency and place them around 300 MHz above the $|1\rangle \leftrightarrow |2\rangle$ transition of the fluxoniums. In the 2D grid, we consider a two-qubit operation performed on the center FTF pair (sites 8, 9 and 10), and for the $[[4, 2, 2]]$ architecture we consider a two-qubit operation performed on the left-top FTF pair (sites 1, 2, and 3). We calibrate the operation with all spectators in $|0\rangle$ and compute the error averaged over all spectator configurations (i.e., 2^6 configurations for the 2D grid, and 2^4 configurations for the $[[4, 2, 2]]$ architecture). The results of these simulations are shown in Fig. 4, where we see that the spectator errors remain well below 10^{-4} for all gate durations $50 \text{ ns} \leq t_g \leq 100 \text{ ns}$, thus, providing strong evidence that our proposed parameter regime can serve as a scalable architecture for fluxonium qubits.

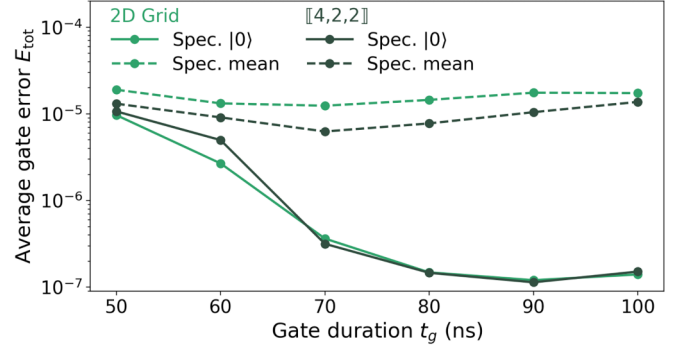


FIG. 4. Spectator errors for the 2D qubit architectures illustrated in Fig. 1. For each system, the two-qubit operation is calibrated with all spectators in $|0\rangle$ (solid lines), and then, the error is averaged over all spectator configurations (dashed lines).

IV. MODEL IMPERFECTIONS

In this section, we analyze the influence of direct capacitive coupling between the transmon couplers and microwave crosstalk on the gate error. First, we add a direct capacitive coupling between all pairs of nearest-neighbor transmons such that the Hamiltonian $\hat{H}_{TT}^{ij} = \hbar g_{TT} \hat{n}_{T_i} \hat{n}_{T_j}$ directly couples transmons i and j with a coupling strength g_{TT} . For the 2D grid, we see in Fig. 5(a) that the error, when averaged over all spectator states, remains below 10^{-4} even when the transmons are coupled strongly to each other. Here, we have numerically calibrated the gates independently for each g_{TT} with all spectator qubits in the $|0\rangle$ state. Noticeably, for $g_{TT}/2\pi \geq 20$ MHz the gate error improves for shorter gate durations. This occurs because the derivative of the conditional phase with respect to the drive detuning increases for longer gate durations. As a result, frequency deviations lead to larger conditional phase errors for longer gates. Additionally, we

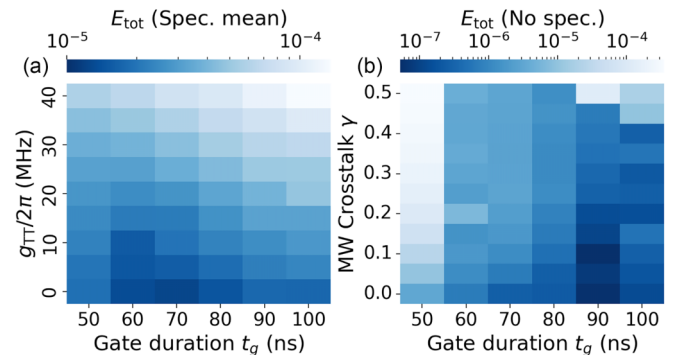


FIG. 5. (a) Total error E_{tot} averaged over all spectator configurations for the 2D grid system as a function of the gate duration t_g and the capacitive coupling between neighboring transmons g_{TT} . (b) Total error E_{tot} of the bare center FTF pair as a function of the gate duration t_g and the microwave (MW) crosstalk strength γ .

find that the static ZZ interaction between the fluxoniums remains suppressed when increasing g_{TT} . Specifically, the maximum ZZ interaction strength as a function of all spectator configurations for the center fluxoniums is 27 and 58 Hz for $g_{\text{TT}} = 0$ and $g_{\text{TT}}/2\pi = 40$ MHz, respectively.

We recall that in the gate scheme proposed in this work, the operation is applied by driving one of the fluxoniums and a small compensation drive is applied to the other fluxonium. In a realistic device architecture, we expect some capacitive coupling between each fluxonium's drive line and the transmon coupler [1,2], which causes effective microwave crosstalk. Here, we simulate the gate error as a function of the microwave crosstalk strength γ , which is defined as the ratio between the coupling strength between the charge line and the transmon and between the charge line and the fluxonium. We assume γ to be the same for both fluxoniums. The exact drive Hamiltonian is detailed in Appendix C.

To isolate the effect of microwave crosstalk from other spectator effects, we compute the total error E_{tot} of the bare center FTF pair of the 2D grid system (i.e., no spectator systems are present in these simulations). As shown in Fig. 5(b), the influence of microwave crosstalk depends on the gate duration. For $t_g = 50$ ns, the gate error increases significantly to $\gtrsim 10^{-4}$, and the dominant error channels in this operation are leakage from $|000\rangle \rightarrow |003\rangle$ and leakage from $|101\rangle \rightarrow |111\rangle$. For the 60, 70, and 80 ns gate duration, the error increases only slightly with increasing γ , but remains below 10^{-5} . For 90 and 100 ns, the gate error also remains low, but increases when $\gamma \gtrsim 0.4$. We conclude that the gate scheme is moderately resilient against microwave crosstalk, and that the precise effects may depend strongly on the gate duration.

V. DISCUSSION

In this work, we have investigated spectator errors that arise when scaling-up fluxonium qubit architectures with transmon couplers. We found that spectator errors limit the two-qubit gate fidelity to below 90% when trivially scaling-up the FTF architecture based on the parameters used in Refs. [1,2]. By modifying the parameter regime and dynamically tuning inactivate transmons to an off point, we showed that these spectator errors can be suppressed to below 10^{-4} in 1D and 2D qubit architectures where each two-qubit operation is influenced by up to six spectator fluxoniums. We further showed that the two-qubit gate operation is robust against direct capacitive coupling between the transmons and moderately resilient against microwave crosstalk.

ACKNOWLEDGMENTS

The authors acknowledge support from the Dutch Research Council (NWO). The authors further acknowledge

the use of computational resources of the DelftBlue supercomputer, provided by Delft High Performance Computing Centre [34]. The authors thank all members of the Andersen lab, Isidora Araya Day, Sebastian Miles, Warre Missiaen, and Anton Akhmerov for insightful discussions.

DATA AVAILABILITY

All numerical data is available through Ref. [35] and the code used for the numerical simulations and data processing is available through Ref. [36].

APPENDIX A: SYSTEM PARAMETERS

The parameters for the 1D system, 2D grid system, and $[[4,2,2]]$ system are detailed in Tables I, II, and III, respectively. Tables II and III, in addition to the E_J and E_{01} values used for the DRMG-X simulations, also list the on-parameters E_J^{on} and E_{01}^{on} for each transmon. Figures 6(a)–(d) show the (transition) energies, drive matrix elements, and frequency detuning of the center FTF pair of the 1D system for varying effective Josephson energies of the transmon. Similarly to Figs. 2(b) and 2(c), Figs. 6(a) and 6(b) show the relevant level energies and transmon qubit frequencies, respectively. As detailed further in Appendix C, the two-qubit operation in the systems with the parameters proposed in this work suppresses leakage through the hybridization structure of the eigenstates. Figure 6(c) shows the charge matrix elements of the left fluxonium (solid lines) and transmon (dashed lines) for the transmon's qubit transition. Importantly, while the transmon's matrix elements are large for all transitions, the fluxonium's matrix elements are small when the left fluxonium is in $|1\rangle$. Figure 6(d) shows the detuning Δ between the conditionally driven transition $|000\rangle \leftrightarrow |010\rangle$ and potential leakage states. The vertical, dashed line indicates the proposed on point for this transmon that is used in the main text. For these parameters, there are only three states with a detuning smaller than 200 MHz, which are the three other transmon transitions. We further notice that tuning too close to the avoided crossings may increase the leakage rate to the fluxoniums' third excited state and the transmon's second excited state. Finally, the parameters used for the 2D grid system are shown graphically in Fig. 7.

The parameters proposed in this work are informed by the following five design rules, which are based on a numerical investigation of the gate scheme for various parameters detailed in Appendix C. We reiterate that we apply the operation by driving the $|000\rangle \leftrightarrow |010\rangle$ transition by driving the fluxonium that has the largest detuning between its $|0\rangle \leftrightarrow |3\rangle$ transition and the transmon's qubit transition. Here, for the purpose of explaining

TABLE I. Parameters used for the 1D system consisting of six fluxoniums and five transmons. For Ref. [1], the parameters of sites 5, 6, and 7 are copied from device B at the flux-point $\varphi_{\text{ext}} = 0$. For Ref. [2], the parameters of sites 5, 6, and 7 are copied from the flux-point $\varphi_{\text{ext}} = 0$. For Ref. [1], the coupling strengths are $g_{\text{FT}}/2\pi = 550$ MHz and $g_{\text{FF}}/2\pi = 120$ MHz, for Ref. [2], $g_{\text{FT}}/h = 236$ MHz and $g_{\text{FF}}/2\pi = 40$ MHz, and in this work, we use $g_{\text{FT}}/2\pi = 200$ MHz and $g_{\text{FF}}/2\pi = 20$ MHz. All parameters are in units of GHz.

Site system	1 FX1	2 T1	3 FX2	4 T2	5 FX3	6 T3	7 FX4	8 T4	9 FX5	10 T5	11 FX6
Ref. [1]											
E_C/h	1.41	0.30	1.33	0.30	1.41	0.30	1.33	0.30	1.41	0.30	1.33
E_J/h	5.85	16.4	5.55	16.2	5.7	16.0	5.4	15.8	5.55	15.6	5.25
E_L/h	0.88		0.60		0.88		0.60		0.88		0.66
E_{01}/h	0.397	5.957	0.256	5.919	0.422	5.880	0.273	5.841	0.449	5.802	0.292
E_{12}/h	5.233		5.166		5.124		5.049		5.016		4.932
E_{03}/h	8.354		7.525		8.320		7.474		8.287		7.425
Ref. [2]											
E_C/h	0.88	0.186	0.88	0.186	0.88	0.186	0.88	0.186	0.88	0.186	0.88
E_J/h	5.170	21.5	5.015	19.0	4.993	16.87	4.335	16.0	4.000	15.0	3.550
E_L/h	0.95		0.75		0.50		0.50		0.40		0.35
E_{01}/h	0.213	5.463	0.160	5.124	0.095	4.817	0.142	4.685	0.138	4.530	0.162
E_{12}/h	4.016		4.102		4.346		3.835		3.672		3.365
E_{03}/h	6.062		5.827		5.587		5.347		5.101		4.860
This work											
E_C/h	1	0.25	1	0.25	1	0.25	1	0.25	1	0.25	1
E_J/h	4.481	23.874	4.118	10.308	4.889	24.920	5.330	11.360	4.889	25.989	5.572
E_L/h	0.95		0.85		0.95		1.05		0.95		1.1
E_{01}/h	0.416	6.650	0.432	4.274	0.333	6.800	0.308	4.501	0.333	6.950	0.293
E_{12}/h	3.701		3.529		3.974		4.199		3.974		4.326
E_{03}/h	6.500		6.300		6.550		6.750		6.550		6.850

TABLE II. Parameters used for the 2D grid system. For this system, we also use $g_{\text{FT}}/2\pi = 200$ MHz and $g_{\text{FF}}/2\pi = 20$ MHz. All parameters are in units of GHz.

Site system	1 T1	2 FX1	3 T2	4 FX2	5 T3	6 FX3	7 T4	8 FX4	9 T5	10 FX5	11 T6	12 FX6	13 T7	14 FX7	15 T8	16 FX8	17 T9
E_C/h	0.25	1	0.25	1	0.25	1	0.25	1	0.25	1	0.25	1	0.25	1	0.25	1	0.25
E_J/h	11.350	4.481	23.874	4.118	10.308	5.236	10.357	4.889	24.920	5.330	11.350	5.176	11.360	4.889	25.989	5.572	11.961
E_L/h		0.95		0.85		1.25		0.95		1.05		0.95		0.95		1.1	
E_{01}/h	4.499	0.416	6.650	0.432	4.274	0.435	4.285	0.333	6.800	0.308	4.499	0.285	4.501	0.333	6.950	0.293	4.626
E_{12}/h		3.701		3.529		3.985		3.974		4.199		4.176		3.974		4.326	
E_{03}/h		6.500		6.300		7.000		6.550		6.750		6.600		6.550		6.850	
E_J^{on}/h	24.920		23.874		23.531		27.080		24.920		24.920		24.920		25.989		25.989
E_{01}^{on}/h	6.800		6.650		6.600		7.100		6.800		6.800		6.800		6.950		6.950

TABLE III. Parameters used for the $\llbracket 4, 2, 2 \rrbracket$ system. For this system, we also use $g_{\text{FT}}/2\pi = 200$ MHz and $g_{\text{FF}}/2\pi = 20$ MHz. All parameters are in units of GHz.

Site system	1 FX1	2 T1	3 FX2	4 T2	5 FX3	6 T3	7 FX4	8 T4	9 T5	10 FX5	11 T6	12 T7	13 FX6	14 T8
E_C/h	1	0.25	1	0.25	1	0.25	1	0.25	0.25	1	0.25	0.25	1	0.25
E_J/h	5.330	24.920	4.889	11.960	5.572	25.989	4.889	11.349	11.359	5.176	11.970	11.980	4.481	11.369
E_L/h	1.05		0.95		1.1		0.95			0.95			0.95	
E_{01}/h	0.308	6.800	0.333	4.626	0.293	6.950	0.333	4.499	4.501	0.285	4.628	4.630	0.416	4.503
E_{12}/h	4.199		3.974		4.326		3.974			4.176			3.701	
E_{03}/h	6.750		6.550		6.850		6.550			6.600			6.500	
E_J^{on}/h		24.920		25.989		25.989		24.920	24.920		25.989	25.989		24.920
E_{01}^{on}/h		6.800		6.950		6.950		6.800	6.800		6.950	6.950		6.800

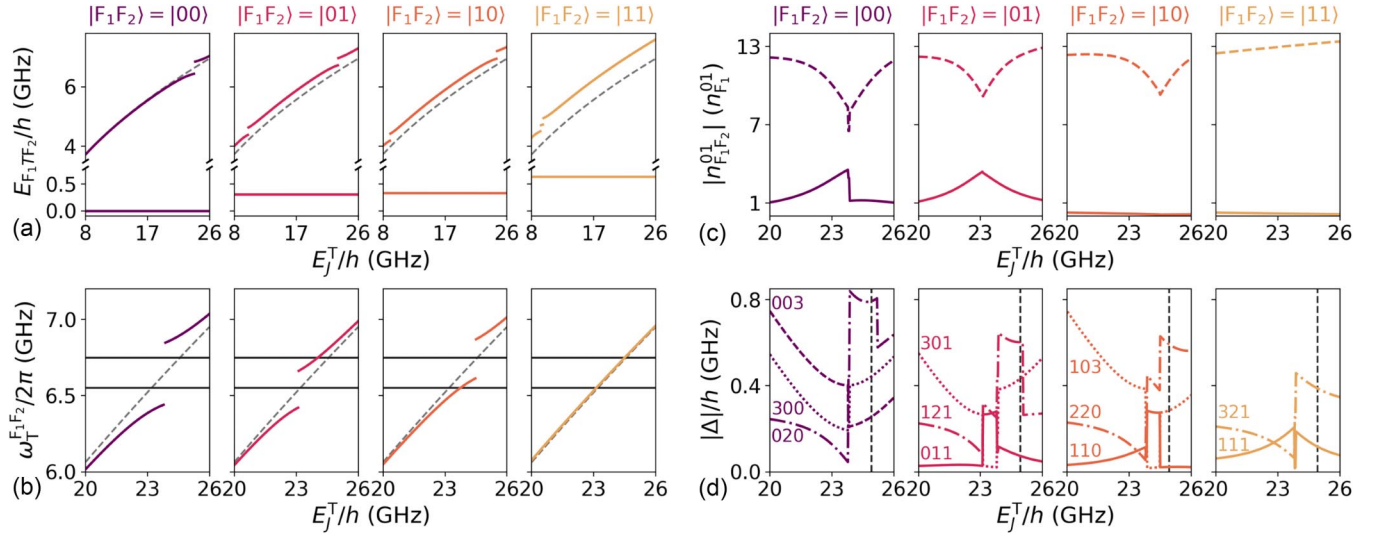


FIG. 6. System parameters of the center FTF pair in the 1D system as a function of the effective Josephson energy E_J^T of the transmon. Each of the four columns in panels (a), (b) and (c), (d) corresponds to a specific configuration of the fluxoniums. (a), (b) Energies and transition frequencies analogous to Figs. 2(b) and 2(c). The gray, dashed lines in panels (a), (b) indicate the bare transmon (transition) frequencies. The solid, horizontal lines in panel (b) indicate the fluxoniums' bare $|0\rangle \leftrightarrow |3\rangle$ energies. (c) Charge matrix elements of the left fluxonium (solid lines) and the transmon (dashed lines) for the transmon's qubit transition. The matrix elements are plotted in units of the bare charge matrix element of the qubit transition of the left fluxonium. (d) Detuning of the conditionally driven transition to the nearest leakage transitions. The leakage levels involve the first excited state of the transmon (solid lines), second excited state of the transmon (dashed-dotted lines), and the third excited state of the fluxoniums (dashed and dotted lines). The vertical, dashed line indicates the effective Josephson energy of this transmon used in the main text.

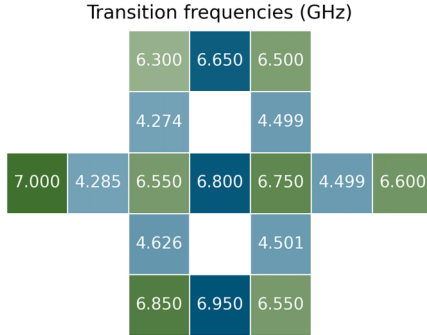


FIG. 7. Visualization of the energy parameters used for the 2D grid. For the fluxoniums, the $|0\rangle \leftrightarrow |3\rangle$ transition energy is plotted on a green color scale, and for the transmons the qubit transition frequency is plotted on a blue color scale.

the design rules, we assume that we are driving the operation through the left fluxonium.

1. The on position of a transmon should be 50–150 MHz above the $|0\rangle \leftrightarrow |3\rangle$ transition of the right fluxonium to ensure a sufficiently large detuning between the $|000\rangle \leftrightarrow |010\rangle$ and $|001\rangle \leftrightarrow |011\rangle$ transitions. This ensures that the leakage channel $|001\rangle \rightarrow |011\rangle$ can be efficiently suppressed using DRAG.
2. The on position of a transmon should be at least 250 MHz above the $|0\rangle \leftrightarrow |3\rangle$ transition of the left fluxonium. As this fluxonium is being driven to

implement the CZ gate, a too small frequency spacing results in leakage from $|00i\rangle \rightarrow |30i\rangle$. However, the frequency spacing should not be too large to ensure a sufficiently large coupling of the conditionally driven transition $|000\rangle \leftrightarrow |010\rangle$ through the fluxonium's charge operator.

3. The on position of next-to-nearest-neighbor transmons should be detuned by at least 150 MHz. A smaller detuning may cause resonances between transmons that are at their on position simultaneously. Nearest-neighbor transmons can have the same on frequency, as they are never used simultaneously to apply a two-qubit operation.
4. When a transmon is not used to apply a two-qubit operation, it is tuned down in frequency to approximately 300 MHz above the $|1\rangle \leftrightarrow |2\rangle$ transition of the fluxoniums.
5. For the individual fluxoniums, we fixed $E_C/h = 1$ GHz and picked the E_J and E_L parameters of the fluxoniums to ensure that the frequency gap between the $|1\rangle \leftrightarrow |2\rangle$ and $|0\rangle \leftrightarrow |3\rangle$ is sufficiently large, as this determines the detuning between on and off transmon couplers.

APPENDIX B: DMRG-X IMPLEMENTATION

In this work, we use the DMRG-X [28] algorithm to compute the eigenvalues and eigenstates of the system. In this algorithm, the state is represented as an MPS, which is

updated iteratively until it converges to an eigenstate of the system. Throughout this iterative optimization, the bond dimensions of the MPS are increased until the optimization converges in the variance of the Hamiltonian $\text{Var}(\hat{H}, \psi) = \langle \psi | \hat{H}^2 | \psi \rangle - (\langle \psi | \hat{H} | \psi \rangle)^2$. We truncate the optimization at $\text{Var}(\hat{H}) < 10^{-9} \text{ GHz}^2$. Numerically, we find that this truncation corresponds to an absolute error in the eigenenergy below 0.1 Hz. Most of the computational complexity of DMRG-X arises from the Lanczos algorithm [37,38], which is an iterative eigensolver that is used to compute the eigenstates of the effective two-site Hamiltonians. We speed-up the calculation by using CuPy [39] and by running the algorithm on an NVIDIA A100 GPU provided by Ref. [34].

To generate the matrix-product operator that represents the Hamiltonian of the system, we first diagonalize each individual transmon and fluxonium. The transmons are diagonalized in the charge-number basis with a charge-number cutoff of 50 (101 charge states total). The fluxoniums are diagonalized in the harmonic-oscillator basis with 100 levels. After diagonalization, the systems are transformed to the eigenbasis, keeping only the lowest three eigenstates for the transmons and the lowest six eigenstates of the fluxoniums.

Due to the strong delocalization of the noncomputational eigenstates of the systems inspired by Refs. [1,2], convergence of DMRG-X (to the correct eigenstate) is not guaranteed. We employ several techniques to ensure and verify that the algorithm converges to the correct eigenstate. To ensure reliable convergence, it is imperative to provide DMRG-X with an accurate guess of the eigenstate. In the 1D systems studied in this work, we generate these guesses by computing the eigenvectors of smaller subsystems using exact diagonalization. Specifically, we split the system into five subsystems: the left-most FTF pair, the center FTF pair, the right-most FTF pair, and the two remaining transmons that sit in between these three FTF pairs. The eigenstates of these five subsystems are then inserted into a single MPS, which forms the initial guess for DMRG-X. A similar methodology is employed for the 2D systems studied in this work.

For the parameters corresponding to Refs. [1,2], significant delocalization of the noncomputational states causes the guessed MPS to not be sufficiently close to the actual eigenstate, and the algorithm tends to converge to a different eigenstate. To resolve this challenge, we adiabatically ramp-up the coupling strengths between the five subsystems from 0% to 100%. This implies that we ramp g_{FT} between FX2 and T2, T2 and FX3, FX4 and T4, and T4 and FX5 and g_{FF} between FX2 and FX3 and between FX4 and FX5 (for the site labeling, we refer to Table I). When the algorithm detects that it does not converge or has converged to an incorrect state, the step-size of the ramp is decreased. To improve the convergence rate of the algorithm, it is important to detect that the

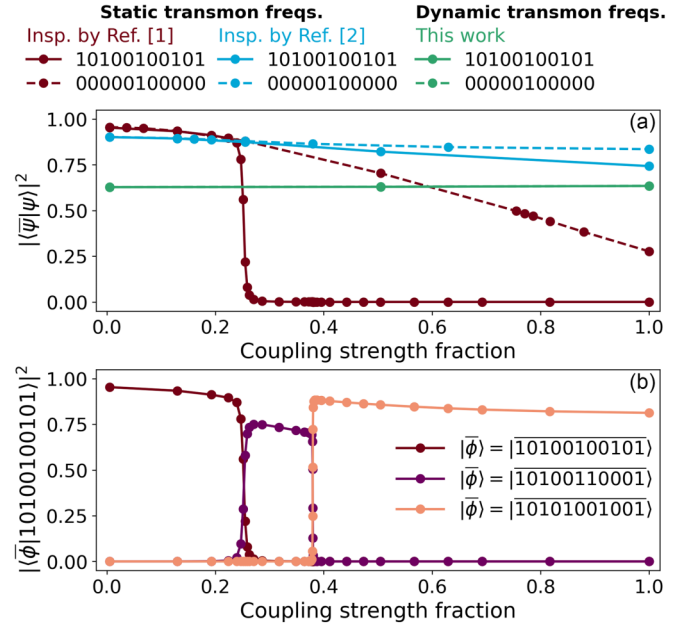


FIG. 8. (a) Overlap between the computed eigenstate $|\psi\rangle$ and the bare eigenstate $|\bar{\psi}\rangle$ as a function of the coupling strength fraction for two different states. (b) Hybridization of the $|10100100101\rangle$ state for the parameters based on Ref. [1] as a function of the coupling strength fraction. As a result of resonances, the state hybridizes very strongly with two other bare eigenstates.

algorithm will not converge (to the correct state) as early as possible. We ensure this by (1) setting an upper bound on the variance of the guessed state and (2) setting an upper bound on the allowed change in the eigenstate of the local two-site effective Hamiltonians.

Throughout the optimization, we monitor the overlap between the optimized MPS and the initial guess and between the optimized MPS and the bare eigenstate of the system. In Fig. 8(a), we plot the overlap between the optimized eigenstate $|\psi\rangle$ and the bare eigenstate $|\bar{\psi}\rangle$ as a function of the coupling strength fraction between the five subsystems. Here, the bare eigenstate $|\bar{\psi}\rangle$ is an eigenstate of the system without any coupling terms. The notation used for the kets $|\cdot\rangle$ follows the site ordering for the 1D system shown in Fig. 1(a). We plot the overlap for two states that both have the center FTF pair in $|010\rangle$, but have all spectator fluxoniums either in $|0\rangle$ or in $|1\rangle$. For the parameters based on Refs. [1,2], the hybridization of the states changes significantly when the coupling strength increases. Furthermore, the hybridization depends on the state of the spectator qubits, indicating that the matrix element of the state with the drive operator will also depend on the state of spectator qubits, which results in spectator errors. For the parameters proposed in this work, we see that the eigenstate does not depend on the coupling strength nor on the state of the spectator qubits, indicating that it does not delocalize strongly outside the center FTF pair.

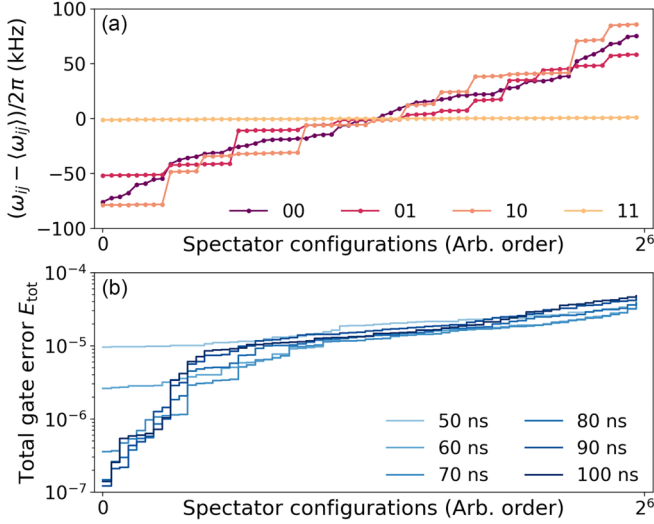


FIG. 9. (a) Deviations of the transmon's qubit transition frequency ω_{ij} as a function of the state of the fluxonium qubits ij for all spectator configurations for the 2D grid system. (b) Total gate error E_{tot} for all gate durations and for all spectator configurations. In panels (a), (b), the order of the spectator configurations is arbitrary (Arb.) as each trace is individually sorted on the y axis.

Figure 8(a) shows that the $|10100100101\rangle$ state for the parameters inspired by Ref. [1] completely loses character and has a negligible overlap with its bare state after a certain coupling strength. This is most likely due to resonances that occur in the system and further emphasizes the need for the adiabatic ramp in this implementation of DMRG-X. In Fig. 8(b), we plot the overlap of this state with two other bare states and see that it hybridizes significantly with the $|10100110001\rangle$ and $|10101001001\rangle$ states.

For the 2D grid system, we computed the eigenvalues and eigenstates relevant for the two-qubit gate operation for all 2^6 spectator configurations. In Fig. 9(a), we plot the transmon's transition frequency $\omega_{ij} = (E_{i1j} - E_{i0j})/\hbar$ for all spectator configurations. For the states that have at least one fluxonium in the ground state, the transmon's qubit frequency deviates by at most 100 kHz from its frequency averaged over all spectator configurations $\langle \omega_{ij} \rangle$. When the fluxoniums are in $ij = 11$, the frequency deviations remain approximately below 1 kHz. In Fig. 9(b), we plot the gate error for all spectator configurations after calibrating the operation with all spectators in $|0\rangle$. For all gate durations and spectator configurations, the error remains below 10^{-4} .

APPENDIX C: GATE CALIBRATION

To suppress crosstalk errors, we reduced the coupling strength g_{FT} and parked the transmon above the $|0\rangle \leftrightarrow |3\rangle$ transition of the fluxoniums. This choice of parameter

regime results in a smaller minimum frequency separation ξ_{ij} , which increases the leakage rate of the operation. While the minimum frequency separation of the systems in Refs. [1,2] is 152 and 68 MHz, respectively, the minimum frequency separation for the parameter regime proposed in this work is only 42 MHz. As a result, the operation is limited by three leakage channels: $|i0j\rangle \leftrightarrow |i1j\rangle$ for ij not equal to the conditionally driven transition. It is unclear how these leakage channels can be suppressed simultaneously with conventional pulse shaping techniques such as DRAG [29–31], as they are often designed to correct for only one leakage channel.

In this work, we leverage the delocalization of the transmon's wavefunction to suppress two of the three leakage channels. Since the transmon's frequency is close in energy to the $|0\rangle \leftrightarrow |3\rangle$ transitions of the fluxoniums, but far detuned from the fluxoniums' $|1\rangle \leftrightarrow |2\rangle$ transitions, its wavefunction will delocalize to a fluxonium when the fluxonium is in the $|0\rangle$ state, and not when it is in the $|1\rangle$ state, see also Fig. 6(d). Therefore, if we implement the two-qubit operation by driving the $|000\rangle \leftrightarrow |010\rangle$ through the charge operator of the left fluxonium, the leakage transitions $|100\rangle \leftrightarrow |110\rangle$ and $|101\rangle \leftrightarrow |111\rangle$ are immediately suppressed as these transitions do not couple through the charge operator of the left fluxonium. This leaves $|001\rangle \leftrightarrow |011\rangle$ as the only leakage transition, which can be suppressed using DRAG. A numerical investigation of this gate scheme revealed that small leakage from $|100\rangle \rightarrow |110\rangle$ can still remain, most likely as a result of the direct capacitive coupling between the fluxoniums. We further suppress this leakage channel by adding a small compensation drive to the right fluxonium.

The operation could be applied by driving any of the two fluxoniums. However, the leakage is maximally suppressed by driving the fluxonium that has the largest detuning between the $|0\rangle \leftrightarrow |3\rangle$ transition and the transmon's $|0\rangle \leftrightarrow |1\rangle$ transition, as this maximizes the detuning between the $|000\rangle \leftrightarrow |010\rangle$ transition and the leakage transition that is suppressed using DRAG. By driving the fluxonium instead of the transmon, we introduce two more leakage channels into the operation, namely $|000\rangle \leftrightarrow |300\rangle$ and $|001\rangle \leftrightarrow |301\rangle$. We ensure these leakage transitions are suppressed by designing them to be sufficiently far detuned in frequency from the driven transition.

The full drive Hamiltonian of this system previously defined in Eq. (7) is given by

$$\hat{H}_D/\hbar = \Omega(\mathcal{E}_I(t) \cos(\omega_d t) + \lambda \mathcal{E}_Q(t) \sin(\omega_d t)) \hat{D}. \quad (\text{C1})$$

Here, Ω is the drive strength, λ is the DRAG parameter, ω_d is the drive frequency, $\hat{D}$ is the drive operator, $\mathcal{E}_I(t)$ is the in-phase pulse envelope previously defined in Eq. (10) and $\mathcal{E}_Q(t)$ is the quadrature pulse envelope defined as

$$\mathcal{E}_Q(t) = \frac{\partial \mathcal{E}_I(t)}{\partial t} = \frac{\pi}{t_g} \sin\left(\frac{2\pi}{t_g} t\right). \quad (\text{C2})$$

We reiterate that the $\tilde{\cdot}$ indicates that the operator has been transformed to the eigenbasis. When the left fluxonium is driven, the drive operator $\hat{\tilde{D}}$ is equal to

$$\hat{\tilde{D}} \mapsto \hat{n}_{F_1} + \eta \hat{n}_{F_2} + \gamma(1 + \eta) \hat{n}_T. \quad (\text{C3})$$

Here, η is the relative drive strength used to suppress the $|100\rangle \leftrightarrow |110\rangle$ transition, and γ is the microwave crosstalk strength used in Sec. IV.

The following procedure is used to calibrate the gate operation, assuming that the left fluxonium is being driven:

1. Calibration of the drive frequency and drive strength. For a range of detunings, the drive strength is optimized to minimize the leakage $1 - |\langle 000|U|000\rangle|^2$. For each detuning, the conditional phase is computed, and the results are fitted with a linear function to find the approximate detuning that provides a conditional phase of π . The procedure is repeated until the conditional phase error converges to below 10^{-7} .
2. The DRAG parameter λ is optimized to minimize leakage $1 - |\langle 001|U|001\rangle|^2$.
3. The relative drive strength η is optimized to minimize the leakage $1 - |\langle 100|U|100\rangle|^2$.
4. The drive frequency and drive strength are optimized with the same procedure as step 1.

Step 2 is skipped if the leakage from $|001\rangle$ is already below 10^{-7} . Similarly, step 3 is skipped if the leakage from $|100\rangle$ is below 10^{-7} . Steps 2–4 are repeated at most two times. Calibration for a system in which the right fluxonium is driven is, of course, analogous, with the notable change that the DRAG parameter is used to minimize leakage from the $|100\rangle$ state and the relative drive strength to minimize leakage from the $|001\rangle$ state.

We emphasize that we only employ this calibration scheme for the systems with the parameters proposed in this work. For the time evolutions performed on the systems with the parameters inspired by Refs. [1,2], we only calibrate a simple cosine pulse envelope and drive the $|101\rangle \leftrightarrow |111\rangle$ transition directly though the transmon's charge operator, i.e., $\lambda = \eta = 0$ and $\hat{\tilde{D}} = \hat{n}_T$. The pulse is subsequently calibrated by optimizing the drive strength to minimize the leakage $1 - |\langle 101|U|101\rangle|^2$ and optimizing the drive frequency to ensure the conditional phase is π with a similar procedure as described in step 1 of the calibration protocol detailed above. For the 1D architecture, the individual contributions from coherent phase errors and leakage errors to the total gate error are shown in Fig. 10.

To demonstrate the versatility of the gate scheme to the varying parameters of each FTF pair, we optimize a 60 ns

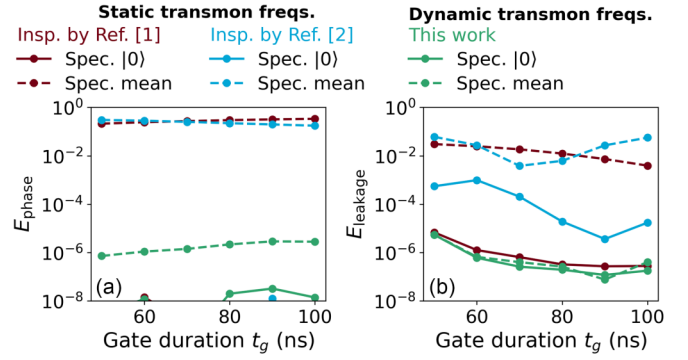


FIG. 10. (a), (b) Individual contributions from the phase error and leakage error, respectively, to the total error plotted in Fig. 3(d). The phase errors for the calibrated operations with all spectrators in $|0\rangle$ are $\lesssim 10^{-8}$ and are therefore not fully visible in panel (a).

CZ gate for each bare FTF pair in the 2D grid system of which the results are shown in Fig. 11, where we see that the gate error is below 10^{-5} for each FTF pair. We further plot the optimized pulse parameters in Figs. 11(b)–11(e). We plot the optimized drive strength Ω in units of the drive strength $\Omega_{2\pi}$, which is the drive strength that would perform a 2π rotation of the same duration on the qubit transition of the driven fluxonium (if it were a perfect, uncoupled two-level system). The DRAG parameter λ and

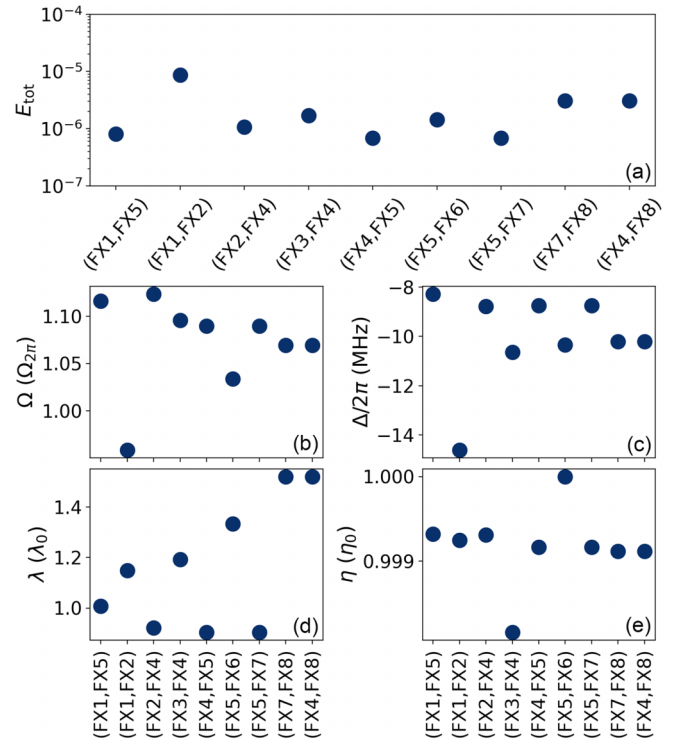


FIG. 11. (a) Total error E_{tot} for all bare FTF pairs in the 2D grid system. (b)–(e) Calibrated drive parameters. In panels (b), (d), and (e), the drive parameters are plotted in units of $\Omega_{2\pi}$, λ_0 , and η_0 , respectively, which are defined in the main text.

relative drive strength are plotted in units of their theoretically predicted values, which, assuming the left fluxonium is being driven, are given by

$$\lambda_0 = \frac{1}{\omega_d - (\omega_{011} - \omega_{001})},$$

$$\eta_0 = -\frac{\langle 110 | \hat{n}_{F_1} | 100 \rangle}{\langle 110 | \hat{n}_{F_2} | 100 \rangle}. \quad (\text{C4})$$

- [1] L. Ding, M. Hays, Y. Sung, B. Kannan, J. An, A. Di Paolo, A. H. Karamlou, T. M. Hazard, K. Azar, D. K. Kim *et al.*, High-fidelity, frequency-flexible two-qubit fluxonium gates with a transmon coupler, *Phys. Rev. X* **13**, 031035 (2023).
- [2] S. Singh, E. Y. Huang, J. Hu, F. Yilmaz, M. F. S. Zwanenburg, P. Kumaravadivel, S. Wang, T. V. Stefanski, and C. K. Andersen, Fast microwave-driven two-qubit gates between fluxonium qubits with a transmon coupler, *Phys. Rev. Appl.* **25**, 024020 (2026).
- [3] K. N. Nesterov, I. V. Pechenezhskiy, C. Wang, V. E. Manucharyan, and M. G. Vavilov, Microwave-activated controlled-Z gate for fixed-frequency fluxonium qubits, *Phys. Rev. A* **98**, 030301(R) (2018).
- [4] I. N. Moskalenko, I. A. Simakov, N. N. Abramov, A. A. Grigorev, D. O. Moskalev, A. A. Pishchimova, N. S. Smirnov, E. V. Zikiy, I. A. Rodionov, and I. S. Besedin, High fidelity two-qubit gates on fluxoniums using a tunable coupler, *npj Quantum Inf.* **8**, 130 (2022).
- [5] E. Dogan, D. Rosenstock, L. Le Guevel, H. Xiong, R. A. Mencia, A. Somoroff, K. N. Nesterov, M. G. Vavilov, V. E. Manucharyan, and C. Wang, Two-fluxonium cross-resonance gate, *Phys. Rev. Appl.* **20**, 024011 (2023).
- [6] W.-J. Lin, H. Cho, Y. Chen, M. G. Vavilov, C. Wang, and V. E. Manucharyan, 24 days-stable CNOT gate on fluxonium qubits with over 99.9% fidelity, *PRX Quantum* **6**, 010349 (2025).
- [7] P. Zhao, G. Zhao, S. Li, C. Zha, and M. Gong, Scalable fluxonium-qubit architecture with tunable interactions between noncomputational levels, *Phys. Rev. Appl.* **25**, 044072 (2026).
- [8] D. A. Rower, L. Ding, H. Zhang, M. Hays, J. An, P. M. Harrington, I. T. Rosen, J. M. Gertler, T. M. Hazard, B. M. Niedzielski *et al.*, Suppressing counter-rotating errors for fast single-qubit gates with fluxonium, *PRX Quantum* **5**, 040342 (2024).
- [9] M. F. S. Zwanenburg, S. Singh, E. Y. Huang, F. Yilmaz, T. V. Stefanski, J. Hu, P. Kumaravadivel, and C. K. Andersen, Single-qubit gates beyond the rotating-wave approximation for strongly anharmonic low-frequency qubits, *Phys. Rev. Research* **7**, 043290 (2025).
- [10] T. Wang, F. Wu, F. Wang, X. Ma, G. Zhang, J. Chen, H. Deng, R. Gao, R. Hu, L. Ma *et al.*, Efficient initialization of fluxonium qubits based on auxiliary energy levels, *Phys. Rev. Lett.* **132**, 230601 (2024).
- [11] T. V. Stefanski, F. Yilmaz, E. Y. Huang, M. F. S. Zwanenburg, S. Singh, S. Wang, L. J. Splitthoff, and C. K. Andersen, Improved fluxonium readout through dynamic flux pulsing, [arXiv:2411.13437](https://arxiv.org/abs/2411.13437).
- [12] K. N. Nesterov and I. V. Pechenezhskiy, Measurement-induced state transitions in dispersive qubit-readout schemes, *Phys. Rev. Appl.* **22**, 064038 (2024).
- [13] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, Fluxonium: Single cooper-pair circuit free of charge offsets, *Science* **326**, 113 (2009).
- [14] S. Krinner, S. Lazar, A. Remm, C. K. Andersen, N. Lacroix, G. J. Norris, C. Hellings, M. Gabureac, C. Eichler, and A. Wallraff, Benchmarking coherent errors in controlled-phase gates due to spectator qubits, *Phys. Rev. Appl.* **14**, 024042 (2020).
- [15] N. Sundaresan, I. Lauer, E. Pritchett, E. Magesan, P. Jurcevic, and J. M. Gambetta, Reducing unitary and spectator errors in cross resonance with optimized rotary echoes, *PRX Quantum* **1**, 020318 (2020).
- [16] D. M. Zajac, J. Stehlik, D. L. Underwood, T. Phung, J. Blair, S. Carnevale, D. Klaus, G. A. Keefe, A. Carniol, M. Kumph, M. Steffen, and O. E. Dial, Spectator errors in tunable coupling architectures, [arXiv:2108.11221](https://arxiv.org/abs/2108.11221).
- [17] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code, *Nature (London)* **605**, 669 (2022).
- [18] J. F. Marques, B. M. Varbanov, M. S. Moreira, H. Ali, N. Muthusubramanian, C. Zachariadis, F. Battistel, M. Beekman, N. Haider, W. Vlothuizen, A. Bruno, B. M. Terhal, and L. DiCarlo, Logical-qubit operations in an error-detecting surface code, *Nat. Phys.* **18**, 80 (2022).
- [19] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, *Nature (London)* **614**, 676 (2023).
- [20] Google Quantum AI and Collaborators, Quantum error correction below the surface code threshold, *Nature (London)* **638**, 920 (2025).
- [21] S. González-García, A. Szasz, A. Pagano, D. Kafri, G. Vidal, and A. Di Paolo, Multi-target density matrix renormalization group X algorithm and its application to circuit quantum electrodynamics, [arXiv:2506.24109](https://arxiv.org/abs/2506.24109).
- [22] F. Lange, L. Heunisch, H. Fehske, D. P. DiVincenzo, and M. J. Hartmann, Cross-talk in superconducting qubit lattices with tunable couplers—comparing transmon and fluxonium architectures, *Quantum Sci. Technol.* **11**, 015020 (2026).
- [23] 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).
- [24] F. Yan, P. Krantz, Y. Sung, M. Kjaergaard, D. L. Campbell, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Tunable coupling scheme for implementing high-fidelity two-qubit gates, *Phys. Rev. Appl.* **10**, 054062 (2018).
- [25] J. An, H. Zhang, Q. Ding, L. Ding, Y. Sung, R. Winik, J. Kim, I. T. Rosen, K. Azar, P. Piñero *et al.*, ZZ-free two-transmon CZ gate mediated by a fluxonium coupler, *Phys. Rev. Appl.* **26**, 014076 (2026).
- [26] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys. (Amsterdam)* **326**, 96 (2011).

- [27] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys. (Amsterdam)* **349**, 117 (2014).
- [28] V. Khemani, F. Pollmann, and S. L. Sondhi, Obtaining highly excited eigenstates of many-body localized Hamiltonians by the density matrix renormalization group approach, *Phys. Rev. Lett.* **116**, 247204 (2016).
- [29] 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).
- [30] J. M. Gambetta, F. Motzoi, S. T. Merkel, and F. K. Wilhelm, Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator, *Phys. Rev. A* **83**, 012308 (2011).
- [31] E. Hyppä, A. Vepsäläinen, M. Papič, C. F. Chan, S. Inel, A. Landra, W. Liu, J. Luus, F. Marxer, C. Ockeloen-Korppi, S. Orbell, B. Tarasinski, and J. Heinsoo, Reducing leakage of single-qubit gates for superconducting quantum processors using analytical control pulse envelopes, *PRX Quantum* **5**, 030353 (2024).
- [32] D. Gottesman, Quantum fault tolerance in small experiments, [arXiv:1610.03507](https://arxiv.org/abs/1610.03507).
- [33] M. Takita, A. D. Córcoles, E. Magesan, B. Abdo, M. Brink, A. Cross, J. M. Chow, and J. M. Gambetta, Demonstration of weight-four parity measurements in the surface code architecture, *Phys. Rev. Lett.* **117**, 210505 (2016).
- [34] Delft High Performance Computing Center (DHPC), DelftBlue supercomputer (phase 2) (2024).
- [35] Data repository, [10.4121/76fc34c4-6af6-4cc3-b234-ff47d58d6de2](https://doi.org/10.4121/76fc34c4-6af6-4cc3-b234-ff47d58d6de2).
- [36] GitHub repository, <https://github.com/andersenqubitlab/crosstalk-in-larger-fluxonium-processors>.
- [37] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.* **45**, 255 (1950).
- [38] Y. Saad, *Iterative methods for sparse linear systems* (SIAM, Philadelphia, PA, 2003), Chap. 7, pp. 229–258.
- [39] R. Okuta, Y. Unno, D. Nishino, S. Hido, and C. Loomis, CuPy: A NumPy-compatible library for NVIDIA GPU calculations, in *Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)* (2017).