Tunable inductive coupler for high fidelity gates between fluxonium qubits

Helin Zhang,1,2,∗ Chunyang Ding,1,2,3,∗ D. K. Weiss,4,5 Ziwen Huang,4,† Yuwei Ma,1,2 Charles Guinn,6 Sara Sussman,6 Sai Pavan Chitta,4 Danyang Chen,4 Andrew A. Houck,6 Jens Koch,4 and David I. Schuster1,2,3,7,‡

1James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA
2Department of Physics, University of Chicago, Chicago, Illinois 60637, USA
3Department of Physics and Applied Physics, Stanford University, Stanford, California 94305
4Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA
5Yale Quantum Institute, Yale University, New Haven, CT 06511, USA
6Department of Physics, Princeton University, Princeton, New Jersey 08544, USA
7Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA

Corresponding author: dschus@stanford.edu

Beyond single qubit operations, several recent works have demonstrated high-fidelity two-qubit fluxonium gates using either fixed capacitive coupling [18–21] or a tunable capacitive coupler [15, 22]. While these two-qubit gate schemes are promising, they either populate states outside of the computational subspace or use high-frequency fluxonium qubits. In this work, we realize a tunable inductive coupler, and perform (> 99.9%) fidelity two-qubit gates. Our inductive coupler, similar to the g-mon coupler [23, 24], realizes a large interaction strength even for low-frequency qubits, without involving any higher energy levels. This enables the gate to take advantage of the full coherence of the fluxonium while avoiding any leakage to states outside the logical subspace.

In Section II, we provide theoretical analysis of our tunable inductive coupler, showing the origins of XX coupling. We also find that both the XX and unwanted ZZ coupling can be turned off, allowing for an operational spot for single qubit gates. In Section III, we report the characterization of our device near this operational spot, measuring the ZZ coupling to be < 100 Hz. Finally, in Section IV, we construct √SWAP and √BSWAP gates using this XX interaction. We demonstrate the tuning of these gates and estimate their fidelities using various benchmarking protocols.

I. INTRODUCTION

Superconducting circuits are a promising platform for the development of scalable, error-corrected quantum computation, heralded by many recent advances towards large scale quantum processors [1, 2] and continued improvements on performances of qubits and gate operations [3–5]. These advances have relied on the transmon qubit [6], and through collective effort, these qubits have achieved excellent coherence times [5, 7] and gate fidelities approaching the quantum error correction threshold [8–10].

A promising alternative to the transmon is the fluxonium qubit [11], which has attractive properties including a nearly degenerate ground state and a large anharmonicity. Compared to other qubits fabricated with similar material qualities, low frequency qubits have less dielectric loss and thus a slower decoherence rate. The large anharmonicity mitigates the speed limit for qubit operations in transmons, and enables quantum gates as fast or even faster despite low qubit frequencies [12]. Recent works have demonstrated ~ 1ms coherence times and > 99.9% fidelity single-qubit gates using fluxonium qubits [13–15]. In addition, one can flux-tune the fluxonium qubit to be noise biased, increasing energy coherence (T1 > 5ms) at the expense of being more phase sensitive [16, 17].

Beyond single qubit operations, several recent works

II. INDUCTIVELY COUPLED HEAVY-FUXLONUM QUBITS

In this letter, we use a tunable inductive coupler, inspired by designs from [23–26], to realize strong coupling between two fluxonium qubits. Each qubit is individually

* These authors contributed equally to this work
† Current address: Superconducting Quantum Materials and Systems Center, Fermi National Accelerator Laboratory (FNAL), Batavia, IL 60510, USA
‡ Corresponding author: dschus@stanford.edu
described by the Hamiltonian

$$H_f = -4E_C \frac{d^2}{d\varphi^2} - E_J \cos(\varphi) + \frac{1}{2}E_L \left( \varphi + 2\pi \frac{\Phi_{\text{ext}}}{\Phi_0} \right)^2,$$

where $E_C = e^2/(2C_q)$ denotes the charging energy, $C_q$ the qubit’s total shunting capacitance, $E_J$ the Josephson energy of the small junction, $E_L = \Phi_0^2/(4\pi^2L_{JA})$ the inductive energy and $L_{JA}$ the total inductance of the superinductor, $\Phi_0 = h/2e$ is the superconducting flux quantum, and $\Phi_{\text{ext}}$ denotes the flux threading the loop.

We design heavy-fluxonium qubits [Fig. 1(a)] with large $E_J/E_C$, such that at half-integer flux bias, our qubits, labeled as $a, b$, have small splittings ($\omega_a/2\pi = 48$ MHz, $\omega_b/2\pi = 61$ MHz) and large anharmonicities ($\alpha_a = 4.6$ GHz, $\alpha_b = 5.2$ GHz). See Fig. 1(c) for the qubit level structure at the sweet spot ($\Phi_{\text{ext}} = \Phi_0/2$) and Table I for the qubit parameters.

These two heavy-fluxonium qubits are linked via an inductive coupler, as shown in Fig. 1(b). This full circuit, analyzed in Appendix A, consists of four degrees of freedom - one harmonic and one fluxonium-like. Ignoring the negligible cross-capacitance, these two fluxonium qubits do not directly interact with each other; instead, they each share an inductance ($E_L$) with the tunable coupler. Although the galvanic coupling is relatively strong, the coupler excitation energies (9.5GHz) are well above the energies associated with qubit excitations, leading to a dispersive interaction. Thus, near half-integer flux for the qubits, we can use a perturbative treatment, creating an effective Hamiltonian up to 4th order [27]:

$$H_{\text{eff}} = -\sum_{\mu=a,b} \frac{\omega_\mu}{2} \sigma_{x\mu}^\mu - \Omega_\mu \sigma_x^\mu + J\sigma_z^a \sigma_z^b + \zeta \sigma_z^a \sigma_z^b,$$

where $\sigma_x^a, \sigma_z^a$ are the qubit Pauli operators in the basis of symmetric and anti-symmetric wavefunctions.

The first term in the effective Hamiltonian describes the qubit frequencies, including the Lamb shifts induced by the coupler. The second term captures the effect of qubit flux bias, which we use for single-qubit gates [12]. This term additionally incorporates a first-order perturbative shift due to the fluxonium-like coupler degree of freedom. The third term is the desired XX coupling, where $J$ is a function of the coupler flux $\Phi_{\text{ext},c}$. Virtual exchanges through coupler excitations leads to a $\varphi_a \varphi_b$ term, which can be truncated to $\sigma_z^a \sigma_z^b$. There are two contributions here: a static one, due to the harmonic coupler degree of freedom, and a flux-tunable one, from the fluxonium-like coupler degree of freedom (see Ref. [27]).
Because these contributions have opposite signs, we can null $J$ at a particular $\Phi_{\text{ext},c}$, which we term the "off position." The final term describes a small unwanted $ZZ$ interaction that comes from fourth-order perturbation theory.

This effective Hamiltonian governs our operation of the two-qubit device: a parametric drive controls $J$, allowing us to achieve two-qubit gates, and a particular $\Phi_{\text{ext},c}$ bias fully turns off the qubit-qubit coupling. However, since $\Omega_\mu$ is coupler-flux dependent, this tuning causes shifts in qubit frequencies. This can be compensated by an accompanying change in individual qubit fluxes ($\Phi_{\text{ext},a}, \Phi_{\text{ext},b}$) to hold each qubit at its sweet spot. This set of constraints defines a curve in the 3D space of flux parameters, which we term the "sweet-spot contour". Along the sweet-spot contour, $\zeta$ is generally non-zero, but we find numerically $\zeta < 3$kHz. Effectively, this $ZZ$ term can be cancelled via small ($< 10^{-5}\Phi_0$) flux shifts away from the qubit sweet spot. We find the flux noise at this bias point not to be the dominant dephasing mechanism, as it would only limit the pure qubit dephasing time to 1ms, see Appendix C.

### III. DEVICE CHARACTERIZATION

We measure the properties of a fabricated 2D superconducting device, validating our theoretical analysis of the tunable inductive coupler and calibrating it for single- and two-qubit gates. We use a tantalum base layer for increased qubit coherence [7, 28], and use double-angle aluminum evaporation [29] to fabricate the Josephson junctions (see Appendix F). The geometry of the device is shown in Figure 1(a). There are 205 Josephson junctions in each qubit’s superinductor ($E_{La}, E_{Lb}$) and 17 junctions in each superinductor of the tunable coupler ($E_L$). There are three dedicated control lines, one for each of the qubits and one for the coupler, which are used for both DC biasing as well as RF flux driving (see Appendix E).

There are two lumped LC resonators capacitively coupled to the fluxoniums for readout. These readout resonators have linewidths $\approx 0.7$ MHz and dispersive shifts $\approx 1.5$MHz from the computational levels. Since the qubit frequencies correspond to a temperature (~3mK) that is lower than the environment temperature, we must always initialize the qubit. We use a measurement-based active reset protocol for initialization [30, 31], enabled by the QICK controlled RFSoC FPGA [32]. The reset is carried out by measuring the qubit state and conditionally playing a single-qubit $\pi$ pulse, completed within 800ns. With this method, we initialize the qubits with $\sim 95\%$ fidelity, primarily limited by our measurement infidelity (which could be improved by using quantum-noise limited amplifiers, see Appendix G).

To determine the coupler parameters, we scan the coupler flux over the range of $\Phi_{\text{ext},c}/\Phi_0 = 0.24$ to $\Phi_{\text{ext},c}/\Phi_0 = 0.34$ and measure qubit frequencies while staying on the "sweet-spot contour", shown in Fig. 2(a). Since $\Omega_{a,b}$ is zero along this contour, we determine the $XX$ coupling strength $J$ from observing the qubit frequency shifts in the dressed Hamiltonian. We measure the $ZZ$ coupling strength by taking the difference of qubit $A$’s frequency when qubit $A$ is at $|0\rangle$ and $|1\rangle$. The frequency difference $(f_{11} - f_{00}) - (f_{01} - f_{10})$, extracted using Ramsey fits with different qubit initial states, measures the $ZZ$ coupling strength to be $< 100$Hz. (c) We perform this same measurement on qubit $A$ while sweeping its flux away from the "sweet-spot contour". We find that $ZZ$ coupling strength becomes non-zero, which we use as an indicator for fine-tuning qubit flux biases.

**FIG. 2.** Measurements of $XX$ and $ZZ$ coupling strength. (a) Top panel: Measured $|0\rangle$, $|1\rangle$, and $|1\rangle$ frequencies as a function of coupler flux as the qubits are tuned to be along the "sweet-spot contour". The dashed lines are from full numerical simulations with parameters extracted from qubit spectroscopy, see Appendix A. Bottom panel: $XX$ and $ZZ$ coupling strength (diamonds) extracted from the data (dots). Numerical simulation predictions are plotted as the dashed lines. $ZZ$ coupling strength is $< 3$kHz across the entire coupler flux range. (b) Ramsey experiment of qubit B at the coupler "off position", with the qubits initialized in either $|0\rangle$ or $|1\rangle$. The frequency difference $(f_{11} - f_{00}) - (f_{01} - f_{10})$ extracted using Ramsey fits with different qubit initial states, measures the $ZZ$ coupling strength to be $< 100$Hz. (c) We perform this same measurement on qubit $A$ while sweeping its flux away from the "sweet-spot contour". We find that $ZZ$ coupling strength becomes non-zero, which we use as an indicator for fine-tuning qubit flux biases.
To measure the precise ZZ coupling strength, we performed Ramsey experiments on qubit B, with qubit A in its ground or excited states, see Fig. 2(b). We then vary the flux near the qubit sweet spots, measuring the ZZ coupling strength. As predicted by our theory, we find a < 100kHz ZZ coupling strength point at the maximum $T_2$ point, see Fig 2(c), implying an off-on contrast $> 10^5$. At this coupler off position, qubit A (qubit B) $T_1$ is 180µs (300µs), and the $T_2$ is 250µs (300µs), see Table I.

We bias the tunable coupler at the off position and realize both single and two-qubit controls with RF flux drives. Due to significant ($\approx 20\%$) geometric DC and RF flux crosstalk in our system, we have off-resonant drives on both qubits. This effectively causes dynamic qubit frequency shifts as well as unwanted XZ, ZX and ZZ drive terms (see Appendix D). Therefore, we develop a method to calibrate the RF flux crosstalk by measuring its effect on qubit frequencies (See Appendix H) and apply crosstalk pulse cancellations in our two-qubit gates.

To dynamically activate the XX coupling, we RF drive the coupler near the difference ($f_{01} - f_{10}$) and sum ($f_{01} + f_{10}$) frequencies of the two qubits, see Fig. 3. We sweep the frequency and length of this pulse at drive amplitudes of 0.5%Φ$_0$ (1.1%Φ$_0$) and observe chevrons of correlated oscillations between $|01\rangle$ and $|10\rangle$ ($|00\rangle$ and $|11\rangle$), finding an oscillation rate of 0.83MHz (2.50MHz). The offset from the expected frequencies is small, indicating that the crosstalk cancellation is effective.

IV. SINGLE AND TWO-QUBIT GATES

We develop fast, high-fidelity single and two-qubit gates with our device. For the single qubit gates, we flux modulate the qubits at their corresponding bare qubit frequencies. With qubit frequencies at 48.4 and 61.8 MHz, we calibrated single qubit π/2 gates using Gaussian pulses with length 83.3 and 65.1 ns, which is approximately 4 qubit Larmor periods. We update the phases of subsequent drive pulses to implement virtual Z gates [33]. We use simultaneous randomized benchmarking [34] to measure average single qubit Clifford gate fidelities, initially measuring qubit A (qubit B) fidelities of 99.90% (99.88%), limited primarily by coherent errors. Thus, we further optimize the single-qubit gates using DRAG shaping [35], finding that the single-qubit gate fidelities increase to 99.94% and 99.95% respectively for each qubit, see Fig. 4(a).

After pulse shaping, the remaining gate infidelity is primarily from qubit decoherence. Using a master-equation to simulate gate infidelities, we estimate contributions of qubit decay and dephasing to be $4 \times 10^{-4}$ and $3 \times 10^{-4}$, respectively. With our current pulse length, the pulse energy deviation from different carrier envelope phases is not the limiting factor of our single qubit gate fidelities. The error from such deviations is less than $1 \times 10^{-5}$, far lower than the dominant decoherence contribution. We also evaluated the contribution of the Bloch-Siegert shift at this drive strength to be $5.6 \times 10^{-5}$. Complete error analysis, including other sources of error, is performed in Appendix J.

For two-qubit gates, due to the negligibly small ZZ interaction strength and the tunable XX coupling, the intuitive choices for our native two-qubit gates are

$$ \sqrt{\text{SWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & i/\sqrt{2} & 0 \\ 0 & i/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (3) $$

$$ \sqrt{\text{BSWAP}} = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 & i/\sqrt{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ i/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \end{pmatrix}. \quad (4) $$

These gates are fully entangling and have been used in demonstrating quantum advantage [1]. Furthermore, by embedding a π pulse between $\sqrt{\text{SWAP}}$ or $\sqrt{\text{BSWAP}}$ gates, we can compile CNOT and CZ gates while also echoing out low-frequency noise.

Using data shown in Fig. 3, we calibrate a $\sqrt{\text{SWAP}}$ ($\sqrt{\text{BSWAP}}$) gate that takes 25.78ns (101.6ns). Because the drive frequency of $\sqrt{\text{SWAP}}$ ($\sqrt{\text{BSWAP}}$) is 13.4MHz (110.2MHz), there are only 4 (11) oscillation periods in a single gate pulse. In the case of the $\sqrt{\text{SWAP}}$ pulse,
We implement cross entropy benchmarking by interleaving √iSWAP (√iSWAP) gates between random single qubit gates. To subtract the errors of the single qubit gates, we first measure a reference sequence that has the same number of layers, where each layer consists of a π/2 gate and a Z gate on each qubit, both with randomly selected phases (ϕ = nπ/4, n ∈ [0, 7]). The depolarizing error [36, 38] is found by varying the number of layers and measuring the qubit state population. Then, we repeat the process with the interleaved sequence. Because the reference gates are randomly sampled, we can calculate the depolarizing error of the two-qubit gate as

\[ p = \frac{P_{\text{interleaved}}}{P_{\text{reference}}} \]

where \( p \) is the depolarizing error of the two-qubit gate, \( D \) is the dimension of the Hilbert space, and \( F \) is the gate fidelity.

We continuously perform cross entropy benchmarking for both √iSWAP and √iSWAP gates for more than 20 hours, as shown in Fig. 4. For √iSWAP, we interleaved it in reference sequences consisting of single qubit gates made of Gaussian pulses. We measured an average reference sequence depolarizing error \( \sim 3.2 \times 10^{-3} \) and an average interleaved sequence depolarizing error \( \sim 6 \times 10^{-3} \). Using equation 5, we derive an average √iSWAP fidelity 99.72 ± 0.04%. For √iSWAP, we implemented DRAG pulses on the single-qubit gates to improve the reference sequence fidelity. We measured depolarizing error \( \sim 1.3 \times 10^{-3} \) of reference sequences and total depolarizing error \( \sim 2.4 \times 10^{-3} \) of interleaved sequences, which results in a √iSWAP gate fidelity 99.91 ± 0.01%.

We compute the error-budget of our single- and two-qubit gates, taking into account decoherence and other error channels, in Table II. For all gates, qubit decoherence is the dominant source of error, while the beyond RWA errors and carrier envelope errors caused by using slow qubits are much lower, which shows the feasibility of high fidelity gates within the low frequency regime.

Finally, we constructed a CNOT gate with two √iSWAP gates and five single qubit gates, and measured a CNOT fidelity of 99.5% with two qubit interleaved randomized benchmarking. This result is consistent with our previously measured single qubit and √iSWAP gate fidelities, and can be further improved with better circuit compilation.

V. CONCLUSION

We have presented a tunable coupler design for heavy-fluxonium qubits, which utilizes inductive coupling to take advantage of the large phase matrix elements. This
The full Hamiltonian of the circuit shown in Fig. 1(b) is

\[ H = H_{\text{qubit}} + H_{\text{coupler}} + V, \]

where

\[ H_{\text{qubit}} = \sum_{\mu=a,b} [4E_{C\mu}n_{\mu}^2 + \frac{1}{2}E_{L\mu}\varphi_{\mu}^2 + E_{J\mu}\cos(\varphi_{\mu} + \pi)] \]

(A1a)

\[ H_{\text{coupler}} = 4E_{C-}n_{-}^2 + \frac{1}{2}E_{Lc}\varphi_{-}^2 - E_{Jc}\cos(\varphi_{-} + 2\pi\Phi_c/\Phi_0) \]

+ 4E_{C+}n_{+}^2 + \frac{1}{2}E_{Le}\varphi_{+}^2,

(A1b)

\[ V = -\frac{E_{La}}{2}[\varphi_{+}(\varphi_{+} + \varphi_{-})] - \frac{E_{Lb}}{2}[\varphi_{b}(\varphi_{+} - \varphi_{-})] \]

(A1c)

\[ + \sum_{\mu=a,b} \frac{E_{L\mu}}{2} \delta\phi_{\mu}[-2\varphi_{\mu} + \varphi_{+} + (-1)^\mu\varphi_{-}], \]

where \([\varphi_{\mu}, n_{\mu}] = i\delta_{\mu\mu}, \mu = a, b, -, +, \) and we have defined \(E_{Lc} = \frac{1}{2}[\left|E_{La}\right| + \left|E_{Lb}\right| + \left|E_{Le}\right|], E_{C+} = 2E_{Cc}, \) and \(E_{C-} = (1/E_{Cc} + 1/[2E_{Cc}])^{-1}. \) All other circuit parameters can be read off from Fig. 1(b). To account for the need to tune the qubit flux as a function of the coupler flux as discussed in the main text, we define \(\delta\phi_{\mu} = \phi_{\mu} - \pi. \) See Ref. [27] for further details on the derivation of Eq. (A1).

We comment briefly on how the effective Hamiltonian Eq. (2) arises from Eq. (A1). The XX operator content of the effective coupling between the qubits can be traced back to the operator form of the coupling between the qubits and the couplers \(\varphi_{\mu}\varphi_{\pm}. \) The phase operator in the qubit subspace of a fluxonium qubit biased at the sweet spot can be written as \((1|\varphi_{+}\rangle\sigma_x \| \rangle\). Thus, one application of the perturbing term \(\varphi_{\mu}\varphi_{\pm}\) swaps e.g. an excitation of qubit \(a\) into the coupler, and an application of \(\varphi_{b}\varphi_{\pm}\) returns this excitation to the qubit subspace in the form of an excitation of qubit \(b. \) For further details, we refer the reader to Ref. [27].

Using this Hamiltonian we locate the sweet-spot contour by minimizing the eigenenergy \(E_{1100}\) as a function of the qubit fluxes, keeping the coupler-flux fixed. The off position is then found by performing the same minimization along the sweet-spot contour as a function of coupler flux.

**Acknowledgments**

The authors would like to thank Sho Uemera, Leandro Stefanazzi, Gustavo Cancelo for their help with the RFSoC, and Kevin He, Kan-Heng Lee, Ziqian Li, Tanay Roy, Rachel Dey for useful discussions. This work was supported by the Army Research Office under Grant No. W911NF1910016. This work is funded in part by EPiQC, an NSF Expedition in Computing, under grant CCF1730449. This work was partially supported by the University of Chicago Materials Research Science and Engineering Center, which is funded by the National Science Foundation under award number DMR-1420709. Devices were fabricated in the Pritzker Nanofabrication Facility at the University of Chicago, which receives support from Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205), a node of the National Science Foundation’s National Nanotechnology Coordinated Infrastructure. Sara Sussman was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

**Appendix A: Full circuit model**

To determine our system’s physical parameters, we measure the two-tone spectroscopies of higher (~ 5GHz) frequency transitions. Each qubit is charge-driven through the readout resonator while the resonator is being probed. We see a sharp change in the resonator transmission when a qubit transition is driven on resonance. Thus, by sweeping different flux biases, we can see a frequency-flux 2D qubit spectroscopy. We measured the
transition frequencies from the ground state to first and second plasmon states ([20] and [30]) for qubit A while changing the qubit A flux bias. Because [10] has significant thermal population around the $\Phi_{ext,a} = 0.5\Phi_0$ bias point, we could also see transitions from it to the higher levels (see Fig. 5(Top)). Similarly, we measured transition frequencies for qubit B (see Fig. 5(Middle and Bottom). Using the package acqubits [40, 41], we subsequently fitted the spectroscopy data with the full Hamiltonian Eq. (A1), as shown in overlay lines in Fig. 5. Because our DC flux crosstalk matrix is not perfectly calibrated across this range, we could not keep the coupler and the other qubit flux biases precisely at the same point during a flux sweep, leading to slight mismatches between the theory curves and experimental data.

![Graphs showing transition frequencies](image)

**FIG. 5.** Plasmon spectroscopy features of both qubits. Top panel: $|00\rangle \rightarrow |20\rangle$ and $|00\rangle \rightarrow |30\rangle$ transition frequency vs qubit A flux bias. Middle panel: $|00\rangle \rightarrow |02\rangle$ and $|00\rangle \rightarrow |03\rangle$ transition frequency vs qubit B flux bias. Bottom panel: zoomed in plot around $\Phi_{ext,b} = 0$. The data was taken with coupler flux around 0, and the other qubit’s flux fixed at 0.

### Appendix C: ZZ suppression

The suppression of static ZZ coupling is an attractive feature of our architecture, allowing for large on/off ratios. This suppression is achieved by cancelling out two contributions of ZZ by fine tuning the qubit flux bias. To see this, we start from the effective Hamiltonian

$$H_{eff} = - \sum_{\mu=a,b} \frac{\omega_{\mu}}{2} \sigma_{z}^{\mu} - \Omega_{\mu} \sigma_{x}^{\mu} + J \sigma_{x}^{a} \sigma_{x}^{b} + \zeta \sigma_{x}^{a} \sigma_{x}^{b}, \quad (C1)$$

where Pauli operators are defined in the eigenbasis of the system at the off position. The static ZZ shift $\zeta$ is small due to a number of reasons. First, $\zeta$ is generally suppressed in low-frequency fluxonium architectures, due to the large detuning between the qubit excitation energies and the energies of the higher-lying states [18]. Second, the galvanic-coupling architecture helps suppress $\zeta$ because the phase matrix elements of qubit states with higher-lying fluxonium states are very small relative to charge matrix elements. Finally, there is no direct coupling between the qubits in this architecture. Thus, second- and third-order contributions to $\zeta$ vanish, with the leading-order contribution arising only at fourth order in perturbation theory [9, 15, 42]. Sweeping the coupler flux while keeping qubits at their sweet spots ($\Omega_{\mu} = 0$), $|\zeta|$ is nonvanishing and on the order of 1 kHz.

However, $\zeta$ can be canceled with a non-zero $\Omega_{\mu}$ by qubit-flux dependent shifts away from the sweet-spot contour. To see this, we diagonalize the single-qubit terms in Eq. (C1) up to second order of $\Omega_{\mu}/\omega_{\mu}$ using the unitary transformation $U = \exp\left(-i \sum_{\mu=a,b} \frac{\omega_{\mu}}{2} \sigma_{z}^{\mu} \right)$, yielding

$$H'_{eff} = U^\dagger H_{eff} U = \sum_{\mu=a,b} \frac{\omega_{\mu}}{2} \sigma_{z}^{\mu} + \frac{2\Omega_{\mu}}{\omega_{\mu}} \sigma_{x}^{\mu} + J \left( \sigma_{x}^{a} + \frac{2\Omega_{a}}{\omega_{a}} \sigma_{z}^{a} \right) \left( \sigma_{x}^{b} + \frac{2\Omega_{b}}{\omega_{b}} \sigma_{z}^{b} \right) + \zeta \sigma_{x}^{a} \sigma_{x}^{b}, \quad (C2)$$

where we have expanded the result to second order in the small parameters $\Omega_{\mu}/\omega_{\mu}$ and ignored negligible corrections arising from the last term in Eq. (C2). Thus the overall ZZ shift is given by

$$\zeta_{tot} = \zeta + \zeta_s, \quad \text{where} \quad \zeta_s = J \frac{4\Omega_a \Omega_b}{\omega_a \omega_b} \quad (C3)$$
where both contributions effectively arise as fourth order perturbations. The sign of $\zeta_s$ depends upon the signs of $J$ and $\Omega_\mu$, where notably $J$ changes sign when moving across the off position in coupler flux, see Fig. 6(a). Thus with small adjustments in the qubit fluxes (so that the $\Omega_\mu$ are nonzero) and the coupler flux, $\zeta_s$ can be used to cancel $\zeta$. This leads to an overall vanishing ZZ shift $\zeta_{tot} = 0$.

We validate this understanding by calculating $\zeta_{tot}$ using the full Hamiltonian (A1)

$$\zeta_{tot}/2\pi = f_{11} - f_{01} - f_{10} + f_{00}, \quad (C4)$$

see Fig. 6(b). The off position and $\zeta_{tot} = 0$ point are within $10^{-5}\Phi_0$ in the qubit fluxes according to the full model, relaxing the constraint of operating exactly on the sweet-spot contour. We need not worry about tuning of the qubit fluxes away from the sweet spot, as flux tuning within $10^{-3}\Phi_0$ is the main contributor to $\zeta_{tot}$, and $\zeta_s$ can be tuned from positive to negative values to cancel out $\zeta$.

**FIG. 6.** The ZZ interaction strength $(a)$ $\zeta_s$ obtained from the effective model and $(b)$ $\zeta_{tot}$ from numerical diagonalization of the full model. Both quantities are plotted as a function of the qubit flux shifts $\Phi_s \equiv \Phi_{ext,a} - 0.5\Phi_0 = 0.5\Phi_0 - \Phi_{ext,b}$ and the coupler flux $\Phi_{ext,c}$. In both plots, the red circle marks the off position and the solid curves are the “sweet-spot contours”. We find that $\zeta_s$ is the main contributor to $\zeta_{tot}$, and $\zeta_s$ can be tuned from positive to negative values to cancel out $\zeta$.

**Appendix D: Effects due to RF flux crosstalk**

In the presence of RF flux crosstalk where flux penetrates each of the qubit loops when we drive the coupler, dynamic effects arises that reduce the gate fidelity. For simplicity, we ignore the effects of the pulse’s Gaussian envelope in this section. The Hamiltonian of the system is

$$H_{eff,ct} = -\sum_{\mu=a,b} \left( \frac{\omega_\mu - \omega_d}{2} \sigma^\mu_d + \Omega_{ct,\mu} \cos(\omega_d t) \sigma^\mu_z \right) + J_{ac} \cos(\omega_d t) \sigma^a_x \sigma^b_x$$

where $\Omega_{ct,\mu}$ describes the maximum qubit-flux drive amplitude due to crosstalk. The drive frequency will typically be set to $f_d \approx f_{01} + f_{10}$ for a $\sqrt{\text{SWAP}}$-style gate, or $f_d \approx f_{01} + f_{10}$ for a $\sqrt{\text{SWAP}}$-style gate, with $\omega_d = 2\pi f_d$. In each case, the drives are detuned from single-qubit transition frequencies. This unwanted drive thus dynamically shifts the qubit frequencies rather than inducing Rabi oscillations [43]. To see this, we first move to a frame rotating at the drive frequency. Making the RWA for the single-qubit drive terms while keeping all terms associated with the two-qubit drive, we obtain

$$H_{eff,ct} = -\sum_{\mu=a,b} \left( \frac{\omega_\mu - \omega_d}{2} \sigma^\mu_d + \Omega_{ct,\mu} \cos(\omega_d t) \sigma^\mu_z \right)$$

$$(D2)$$

$$+ J_{ac} \cos(\omega_d t) \left( \sigma^a_x + \Omega_{ct,a} \sigma^a_z \right) \left( \sigma^b_x + \Omega_{ct,b} \sigma^b_z \right),$$

where $\delta\omega_{\mu} = \Omega_{ct,\mu}^2/2(\omega_d - \omega_d)$ and we have returned to the lab frame. Thus, the leading-order effects of ac-flux crosstalk are to renormalize the qubit frequencies as well as modify the effective two-qubit drive term from $XX$ to a combination of $XX, ZX, XZ$ and $ZZ$. To avoid these unwanted dynamic effects, it is thus critical to cancel geometric flux crosstalk. We discuss the quantitative contribution of ac-flux crosstalk to gate infidelities in Appendix J.

**Appendix E: Experimental setup**

The experiment was performed in a Bluefors LD-250 dilution refrigerator with the wiring configured as shown in Fig. 7. The flux and charge inputs are attenuated at the 4K stage and the mixing chamber with standard XMA attenuators, except the final 20dB attenuator on the RF charge line (threaded copper). The DC and RF-flux lines included commercial low-pass filters (Mini-Circuits® ZFBT-4R2GW+) where the coupler bias-tee was modified such that the capacitor is replaced with a short to further lower the high-pass cutoff frequency. The DC and RF-flux signals were combined in a bias-tee (Mini-Circuits® XMA), except the final 20dB attenuator on the DC charge line (threaded copper). The DC and RF-charge lines also had additional low-pass filters with a sharp cutoff (8 GHz) from K&L microwave. Home-made Eccosorb (CR110) IR filters were added on the flux, input and output lines, which further improved the $T_1$ and $T_2$ coherences, and reduced the qubit and resonator temperatures. The device was heat sunk to the base stage of the dilution refrigerator (stabilized at 15 mK) via an OFHC copper post, while surrounded by an inner lead shield thermalized via a welded copper ring. This was additionally surrounded by two cylindrical $\mu$-metal cans.
\section*{Appendix F: Device fabrication}

The device (shown in Fig. 1 in the main text) was fabricated on a 430\,$\mu$m thick C-plane sapphire substrate. The base layer of the device, which includes the majority of the circuit (excluding the Josephson junctions), consists of 200nm of tantalum with features fabricated via optical lithography and HF etch at wafer-scale.

We perform standard TAMI cleaning for annealed sapphire substrates, followed with nanostrip etching at 50C for 10 minutes, and sulfuric acid etching at 140C for 10 minutes. We subsequently deposit 200nm tantalum in an AJA ATC 2200 sputtering tool at 800C. A 2000nm thick layer of AZ 1518 was used as the (positive) photoresist, and the large features were written using a Heidelberg MLA 150 Direct Writer. After 60 seconds of development with MIF 300 and a 10 minute oven bake at 120C, we perform 20 seconds of HF etching using Ta etchant 1:1:1 (Transene Tantalum Etchant 111).

The junction mask was fabricated via electron-beam lithography with a bi-layer resist (MMA-PMMA) comprising of MMA EL11 and 495PMMA A6 spun at 4000RPM. The e-beam lithography was performed on a 100kV Raith EBPG5000 Plus E-Beam Writer. All Josephson junctions were made with the Dolan bridge technique and etched for 2 minutes using a 3:1 DI:IPA solution at 6C. Aluminum was subsequently evaporated onto the chip in a Plassys electron beam evaporator using double angle evaporation (\(\pm 21^\circ\)), first depositing a 70nm Al layer, performing static oxidation, and subsequently depositing a 90nm Al layer. The wafer was then diced into 7\,$\times\,$7 mm chips, mounted on a printed circuit board, and subsequently wire-bonded.

\section*{Appendix G: Readout and active reset}

We utilize the active feedback capabilities of the Xilinx RFSoC FPGA board with the QICK open source control codebase. We perform simultaneous dispersive readout, using readout lengths of 10.4 and 18.2\,$\mu$s respectively. The readout data will then be processed on the
FIG. 8. Histogram of readout data after active reset initialized the qubit in either ground or excited state.

FPGA board and compared to a previously calibrated threshold, and the board will conditionally play a single qubit $\pi$ pulse to initialize the qubit to its ground state.

We first take 10000 readouts with the qubit uninitialized at the sweet spot. Because the qubit’s Boltzmann temperature ($< 2.8\text{mK}$) is lower than the temperature of the environment, we expect to see the qubit in thermal equilibrium. Indeed, we see approximately a 50-50 population split between the ground and excited states. Using a double Gaussian function, we fit the data to determine the demarcation line that best classifies the qubit state for use in readout and active reset.

We measure our reset fidelity by initializing the qubit in $|g\rangle$ and $|e\rangle$ using active reset, and then take another measurement right after the reset to get histogram shown in Fig. 8. We extracted post-selected active reset fidelities of 93% and 95% respectively for each qubit, and it is very close to the readout fidelities due to the fast feedback time and high fidelity single qubit $\pi$ pulse. By fitting the thermal population of the qubit ground and excited states, we can also measure the qubit temperatures to be $\sim 50\text{mK}$.

Appendix H: Flux crosstalk calibrations

The DC and RF flux crosstalk of the system are significant and require calibration to mitigate unwanted effects that degrade gate fidelity. We measure the DC flux crosstalk by performing 2D flux sweeps for different pairs of flux lines while probing the resonator. This reveals sharp flux-dependent features that originate from bringing qubit energy levels on resonance with the readout resonator. The slope of these lines is the DC flux crosstalk of our system.

To minimize RF flux crosstalk, we use compensation pulses that are played at the same time as the original drive pulse. These compensation pulses have the same length as the original pulse, but we tune the phase and amplitude of these pulses to minimize the effect of flux crosstalk. Since the phases of these compensation pulses are independent parameters, we first calibrate them individually. Because the flux crosstalk can be understood as an off-resonant qubit drive, we would observe an AC Stark-shifted qubit frequency when performing a Rabi experiment. We Rabi drive at the $\sqrt{\text{SWAP}}$ frequency, which is the sum of bare qubit frequencies. The Rabi contrast is maximized when the qubit frequency shift is minimized, making this a good metric for measuring RF crosstalk. Therefore, we sweep the phase of the compensation pulse while playing it simultaneously with the drive pulse, as shown in Fig 9(a). The maximum Rabi contrast of the $|00\rangle \leftrightarrow |11\rangle$ oscillation determines the correct compensation pulse phase.

After fixing the correct phases for both compensation channels, we determine the optimal cancellation pulse amplitudes in a similar sweep. As before, we look for the maximum Rabi contrast of the $|00\rangle \leftrightarrow |11\rangle$ oscillation, but in this case, we sweep the amplitude of the compensation pulse, shown in Fig 9(b). Because the cancellation gains from the two qubit channels will affect each other, we did this iteratively to find the optimal amplitudes for both of them.

Due to the linear nature of the flux crosstalk, the amplitude of drive pulse and compensation pulses have a linear relationship, and the phase difference between all...
three pulses is not amplitude dependent. With cancellation phases and amplitudes for one drive pulse calibrated, we can utilize this procedure to find compensation pulses for all $|00\rangle \leftrightarrow |11\rangle$ drive pulses. The two-qubit gates reported in this manuscript are all constructed in this way, with the coupler drive pulse and cancellation pulses on each qubit.

**Appendix I: Two-qubit gate calibrations**

![Graphs](image)

**FIG. 10.** (a) Playing the pulse sequence in Fig. 12 while sweeping the phase of qubit A phase gates ($Z_A$). At $\phi_A = 357^\circ$, the phase of $\phi_{01} + \phi_{10}$ is perfectly cancelled, thus $\phi_{01} + \phi_{10} = 3$. (b) Playing the pulse sequence in Fig. 13 with various copies of the $\sqrt{\text{SWAP}}$ gates, and measure the final state of both qubits. With this sequence, the qubit A should only see a decay process to thermal equilibrium, and qubit B state dependence on the number of gates is shown in Eqs. I3. By fitting the data we can extract the value of $\phi_{01}$ and $\phi_{10}$.

We calibrate the $\sqrt{\text{SWAP}}$ and $\sqrt{\delta\text{SWAP}}$ gates with gate sequences designed to amplify the gate errors. The gate calibration process consists of two steps, the rotation angle calibration and phase calibrations. We can write the Hermitian matrix for a gate generated by a generic $XX$ parametric drive on the coupler as

$$\sqrt{\phi b \text{SWAP}} = \begin{pmatrix}
\cos \theta & 0 & 0 & ie^{i\phi_D} \sin \theta \\
0 & e^{i\phi_{01}} & 0 & 0 \\
0 & 0 & e^{i\phi_{10}} & 0 \\
ie^{i(\phi_{11} - \phi_D)} \sin \theta & 0 & 0 & e^{i\phi_{11}} \cos \theta
\end{pmatrix}$$ (I2)

for $|00\rangle \leftrightarrow |11\rangle$ oscillation and

$$\sqrt{\phi b \text{SWAP}} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & e^{i\phi_{01}} \cos \theta & ie^{i(\phi_{01} + \phi_D)} \sin \theta & 0 \\
0 & ie^{i(\phi_{10} - \phi_D)} \sin \theta & e^{i\phi_{10}} \cos \theta & 0 \\
0 & 0 & 0 & e^{i\phi_{11}} \cos \theta
\end{pmatrix}$$ (I2)

for $|01\rangle \leftrightarrow |10\rangle$ oscillation, where $\phi_D$ is the phase of the coupler drive, $\phi_{01}$, $\phi_{10}$, and $\phi_{11}$ are phases due to the frequency shift of levels while the drive is on, which have relationship $\phi_{11} = \phi_{01} + \phi_{10} + \phi_{zz}$ [44]. In our system, since all frequency shifts and the $ZZ$ term during the parametric drive are small, we have $\phi_{11} \approx \phi_{01} + \phi_{10}$ and all of $\phi_{01}$, $\phi_{10}$, and $\phi_{11}$ are close to 0. We first calibrate the rotation angle ($\theta$). We fixed the pulse length for $\sqrt{\text{SWAP}}$ at 99000/384 $\approx$ 257.8ns and $\sqrt{\delta\text{SWAP}}$ at 39000/384 $\approx$ 101.6ns. Since our AWG (Xilinx RFSoC) has a 384MHz processor, the pulse lengths are integer multiples of 1000/384ns. We initialize the system in state $|00\rangle$, sweep the pulse amplitude while playing the same pulse consecutively for $4n + 2$ times, and look for the amplitude that gives us the highest fidelity $|11\rangle$ state. With playing the same gate for up to 402 times, we can obtain the value of parameter $\theta$ up to $1 \times 10^{-4}$ in precision.

![Diagrams](image)

**FIG. 11.** (a) Experimental data of state tomography on $|gg\rangle - i|ee\rangle$. (b) State tomography of an ideal $|gg\rangle - i|ee\rangle$ calculated with theory.

With $\theta$ calibrated to be $\pi/4$, gate phases $\phi_{01}$, $\phi_{10}$, $\phi_{11}$ and $\phi_D$ are subsequently calibrated. Since the calibration process is very similar for $\sqrt{\text{SWAP}}$ and $\sqrt{\delta\text{SWAP}}$ gates, here we only use $\sqrt{\delta\text{SWAP}}$ as an example for explanation. We calibrate these phases with the sequence shown in Fig. 12:

where $Z_A$ and $Z_B$ are single qubit phase gates on each qubit with phase $\phi_A$ and $\phi_B$. A unit consists of one $\sqrt{\text{SWAP}}$ gate and two single qubit $Z$ gates, and we noticed that with $4n + 2$ number of units, only when...
$\phi_A + \phi_B = -\phi_{11}$ do we get a bSWAP gate up to some phases. Thus we play a sequence of 321 units while sweeping $\phi_A$ (shown in Fig 10), and measured $\phi_{11}$ to be around -3 degrees. We subsequently play the sequence as shown in Fig 13.

After the calibration, we set qubit A Z gate after the parametric coupler drive $\phi_A = -\phi_{10}$, qubit B Z gate $\phi_B = -\phi_{01}$, and change the drive phase by $-\phi_D$. In this experiment all the single qubit Z gates are virtual, therefore we only need to update the phase of the pulses after a $\sqrt{bSWAP}$ or $\sqrt{SWAP}$ gate for each qubit.

\begin{equation}
\frac{1}{2}(1 + (-1)^n \sin (4n(\phi_{10} - \phi_{11}) + 2(\phi_{10} - \phi_{11}) + \phi_D))
\end{equation}

and with $4n + 2$ units, we have qubit B ground state population,

and qubit A ground state population should always be 0 for all the n values. Fitting this function gives us $\phi_{10} - \phi_{11}$ and $2\phi_{10} - 2\phi_{11} + \phi_D$ (shown in figure 10). With the $\phi_{11}$ value measured from the step above, we can derive $\phi_{10}$, $\phi_{11}$, and $\phi_D$. Since $\phi_{ZZ}$ is very small in our system, we can calculate $\phi_{01}$ as $\phi_{11} - \phi_{10}$, and the result is very close to directly measuring $\phi_{01}$ by switching operations on qubit A and qubit B.

After the calibration, we set qubit A Z gate after the parametric coupler drive $\phi_A = -\phi_{10}$, qubit B Z gate $\phi_B = -\phi_{01}$, and change the drive phase by $-\phi_D$. In this experiment all the single qubit Z gates are virtual, therefore we only need to update the phase of the pulses after a $\sqrt{bSWAP}$ or $\sqrt{SWAP}$ gate for each qubit.

**Appendix J: Error budgeting**

Various sources contribute to single- and two-qubit gate infidelities, including qubit decay, heating, dephasing, carrier-envelope variations, beyond RWA effects and RF flux crosstalk. In this appendix we quantify the contribution of each error source to the infidelity via QuTiP simulations and analytical estimates (where possible).

Simulations suggest that leakage to non-computational states is negligible, attributable to the significant detuning between relevant drive frequencies and undesired transition frequencies. We can thus safely truncate the Hilbert space to the computational subspace. In the following we focus on the specific example of error-budgeting two-qubit gates, with similar results holding for single-qubit gates. In the presence of a coupler-flux drive, the lab-frame Hamiltonian is

where $A$ is the amplitude of the drive, $f(t)$ is the Gaussian envelope and $\theta_{CE}$ is the carrier-envelope phase.

**1. Beyond the RWA**

Both qubits in our sample have Larmor periods of about 20 ns, only a slightly smaller timescale than the single- and two-qubit gates that we implement. Thus, it is possible that fidelities could begin to be limited by effects arising from counter-rotating terms. To obtain a semi-analytic estimate of the associated fidelity reduction, we utilize a Magnus expansion [46–48] following [27]. Noting that the Hamiltonian only couples pairs of states $|00\rangle \leftrightarrow |11\rangle$ and $|01\rangle \leftrightarrow |10\rangle$, we write the Hamiltonian (J3) as a direct sum $H_{2q}(t) = H_{-}(t) \oplus H_{+}(t)$, where

\begin{equation}
H_{\pm}(t) = -\frac{\omega_{q}}{2}\sigma_{z}^{\pm} + A(t) \cos(\omega_{dt} + \theta_{CE})\sigma_{x}^{\pm},
\end{equation}

We have defined $\omega_{q} = \omega_{a} \pm \omega_{b}$ and the Pauli matrices as e.g. $\Sigma_{x}^{+} = |00\rangle\langle11| + H.c.$ and $\Sigma_{z}^{-} = |01\rangle\langle10| + H.c.$ For simplicity we set $\theta_{CE} = 0$, and postpone a discussion of carrier-envelope variations to Sec. J 2. To isolate the effects of the drive, we move into a rotating frame, yielding

\begin{equation}
H'_{\pm}(t) = A(t) \cos(\omega_{dt})[\cos(\omega_{dt})\Sigma_{x}^{\pm} + \sin(\omega_{dt})\Sigma_{y}^{\pm}],
\end{equation}

Utilizing the first two terms in the Magnus expansion, the propagator is $U_{2q}(t) = \exp[\int_{0}^{t} H'_{\pm}(s) ds]$. The Magnus expansion for $H'_{\pm}(t)$ should be truncated as soon as the discrepancy between the numerical solution and the analytical solution becomes significant.
TABLE II. Error budget for single- and two-qubit gates. Analytic estimates of decoherence rates can be found in Ref. [45]. We use gate lengths $\tau = 83.3, 65.1, 101.6, 257.8$ ns for the gates $\sqrt{X_0}$, $\sqrt{X_5}$, $\sqrt{SWAP}$ and $\sqrt{SWAP}$, respectively. The unaccounted error is attributed to insufficient calibration and RF flux crosstalk.

<table>
<thead>
<tr>
<th>error source</th>
<th>$\sqrt{SWAP}$</th>
<th>$\sqrt{bSWAP}$</th>
<th>$\sqrt{X_0}$</th>
<th>$\sqrt{X_5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decay</td>
<td>$4.6 \times 10^{-4}$</td>
<td>$1.8 \times 10^{-4}$</td>
<td>$1.5 \times 10^{-4}$</td>
<td>$1.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>heating</td>
<td>$4.6 \times 10^{-4}$</td>
<td>$1.8 \times 10^{-4}$</td>
<td>$1.5 \times 10^{-4}$</td>
<td>$1.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>dephasing</td>
<td>$3.0 \times 10^{-4}$</td>
<td>$1.2 \times 10^{-4}$</td>
<td>$1.0 \times 10^{-4}$</td>
<td>$7.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>beyond RWA</td>
<td>$1.4 \times 10^{-5}$</td>
<td>$1.1 \times 10^{-5}$</td>
<td>$5.2 \times 10^{-6}$</td>
<td>$7.5 \times 10^{-7}$</td>
</tr>
<tr>
<td>carrier envelope phase</td>
<td>$5.8 \times 10^{-5}$</td>
<td>$(1.4 \times 10^{-5})$</td>
<td>$(0.5 \times 10^{-5})$</td>
<td>$(0.5 \times 10^{-5})$</td>
</tr>
<tr>
<td>higher order drive terms</td>
<td>$(8.5 \times 10^{-5})$</td>
<td>$(2.3 \times 10^{-5})$</td>
<td>$(1.4 \times 10^{-5})$</td>
<td>$(2.2 \times 10^{-5})$</td>
</tr>
</tbody>
</table>

The gate time $\tau$ is taken to be $\tau = \sqrt{2/\pi}/(A \text{erf}(\sqrt{2}))$, appropriate to obtain a $\sqrt{SWAP}$ or $\sqrt{bSWAP}$ in the RWA with Gaussian envelope $f(t) = \exp(-0.5t^2/(\sigma^2))$ and width $\sigma = \tau/4$.

We obtain for the first terms in the Magnus expansion

$$\Delta_{1,\pm}(\tau) + \Delta_{1,\pm}(\tau) = -i \int_0^\tau \mathrm{d}t \frac{H(t)}{\tau} \int_0^\tau \mathrm{d}t' \frac{H(t')}{\tau} \{ t, t' \} = \left[ \frac{-i}{\tau} \int_0^\tau \mathrm{d}t J(t) \{ \omega(t), \omega(t) \} \right] \cos(\omega(t) \tau) \cos(\omega(t) \tau).$$

These terms describe the familiar Bloch-Siegert shift, where the resonance frequencies of both qubits are shifted by the coupler drive.

As an aside, we comment that we have just derived the propagator to second order for single-qubit gates with a drive on $\sigma_x$ (returning to usual Pauli-matrix notation for clarity), $U_{1,q}(t) = \exp(-i \frac{1}{2} Q_x(t_1) - i \frac{1}{2} Q_y(t))$, dropping the subscripts on $\alpha$ and $\beta$ and redefining $\alpha$ and $\beta$ with an additional factor of 2 for later convenience.

Returning to two-qubit gates, we now specialize to the case of $\sqrt{SWAP}$ with $\omega_{d} = \omega_{a} + \omega_{b}$. The propagator in this case is

$$U_{2,q}(\tau) = \exp \left( -i \frac{1}{2} \frac{\pi}{4} + \delta \right) \left[ \sigma_x^a \sigma_y^b + \sigma_y^a \sigma_x^b \right] -i \frac{\alpha_+}{2} \left[ \sigma_x^a \sigma_y^b - \sigma_y^a \sigma_x^b \right] -i \frac{\beta_+}{2} \left[ \sigma_x^a \sigma_y^b - \sigma_y^a \sigma_x^b \right].$$

where we have isolated the deviation of $\alpha_+$ from the ideal value of $\pi/4$. $\alpha_- = \frac{\pi}{4} + \delta$. We may now read off the Bloch-Siegert shifts $\omega_{BS,a}$, given by $\omega_{BS,a} = (\beta_+ - \beta_-)/\tau$ and $\omega_{BS,b} = (\beta_+ - \beta_-)/\tau$. With the propagator in hand, we may now compute the gate fidelity using the standard formula [49]

$$F = \frac{\text{Tr} [U(\tau) U^\dagger(\tau)] + \left| \text{Tr} [U^\dagger U(\tau)] \right|^2}{d(d+1)},$$

where $d$ is the dimension of the Hilbert space and $U_T$ is the target unitary. Expanding $\text{Tr} [U^\dagger U(\tau)]$ about $(\delta, \alpha_+, \beta_-, \beta_+) = (0, 0, 0, 0)$ and retaining only leading-order terms, we obtain

$$F = 1 - \frac{2}{5} (\alpha_+^2 + \delta^2 + \beta_+^2 + \frac{8}{\pi^2} \beta_-^2),$$

These terms describe the familiar Bloch-Siegert shift, where the resonance frequencies of both qubits are shifted by the coupler drive.
in the case of $\sqrt{iSWAP}$ and
\[ F = 1 - \frac{2}{5}(\alpha_+^2 + \beta_+^2 + \frac{8}{\pi^2} \beta_+^2), \tag{J13} \]
in the case of $\sqrt{6SWAP}$ (with $\delta$ now defined as $\alpha_+ = \frac{\pi}{8}$). Evaluating the integrals $\alpha_\pm, \beta_\pm$ numerically, we obtain $1 - F = 1.4 \times 10^{-4}, 1.1 \times 10^{-5}$, in the cases of $\sqrt{SWAP}, \sqrt{6SWAP}$, respectively. For the Bloch-Siegert shifts, we obtain $|\omega_{BS,a}|/2\pi = 9$ kHz, $|\omega_{BS,b}|/2\pi = 20$ kHz in the case of $\sqrt{SWAP}$ and $|\omega_{BS,a}|/2\pi = 11$ kHz, $|\omega_{BS,b}|/2\pi = 2$ kHz in the case of $\sqrt{6SWAP}$.

We compare the fidelity results with estimates obtained by numerical simulation of Eq. (J3) using QuTiP [50]. These simulations yield infidelity contributions of $8.4 \times 10^{-5}, 1.3 \times 10^{-5}$ for $\sqrt{SWAP}$ and $\sqrt{6SWAP}$, respectively. In both cases, the analytically-estimated values are within about a factor of two of the numerically-obtained infidelities.

In the case of single-qubit gates with $U_F = \sqrt{X_\mu}$, we obtain for the fidelity to leading order
\[ F = 1 - \frac{4\beta_\mu^2}{3\pi^2} + \delta^2 \frac{1}{6}. \tag{J14} \]
where $\delta = \alpha - \frac{\pi}{8}$. We obtain infidelity estimates of $7.8 \times 10^{-5}, 7.5 \times 10^{-5}$ for $\sqrt{X_a}$ and $\sqrt{X_b}$, respectively. Performing numerical simulations (with the single-qubit version of Eq. (J3)) we obtain infidelities of $5.6 \times 10^{-5}$ in both cases. These results are slightly lower than the analytic estimates, but of the same order of magnitude.

2. Carrier envelope variations

The phase $\theta_{CE}$ is constantly updated during the experiment. As the logical states are defined in the rotating frame, the phase of the pulse should be matched with the dynamical phase difference of the states being swapped. Besides, we make use of the phase $\theta_{CE}$ to realize virtual $Z$ gates. Because the pulse length is only 4 (11) drive periods in the case of the $\sqrt{SWAP}$ (\sqrt{6SWAP}) gate, the energy carried by the pulse is generally a function of $\theta_{CE}$ (to be contrasted with the more standard case when the pulse length is much larger than the drive period, and there is no such dependence on $\theta_{CE}$). Additionally, the effects of the counter-rotating terms can depend on $\theta_{CE}$.

To quantify the contribution of this source of error, we numerically sweep the phase of the pulse and calculate the gate infidelity as a function of $\theta_{CE}$. The average of the increase in infidelity due to phase variation is what is taken as the contribution from carrier envelope phase in Table II.

3. Higher-order drive terms

As discussed in Ref. [27], the form of the drive operator in Eq. (J1) is an approximation. In reality, the drive operator associated with $\Phi_{ext,c}$ is not perfectly $XX$ but includes also nonzero $IX, XF$ and other components. Similarly for the drive operators associated with $\Phi_{ext,a}, \Phi_{ext,b}$, the drive operators contain nonzero $ZX, XZ$ and other terms. Including these additional drive terms in the simulations contributes infidelities on the order of $10^{-5}$, see Table II.

4. Decoherence

The gate fidelity is reduced due to decay, heating and dephasing effects. In the following we specialize to the case of the $\sqrt{SWAP}$ gate, taking $\omega_d = |\omega_b - \omega_a|$. The discussion follows similarly for the case of $\sqrt{6SWAP}$, with $\omega_d = \omega_b + \omega_a$. The Hamiltonian after performing the RWA is
\[ H_{RWA}^2 = \frac{A}{4}(\sigma_x^a \sigma_x^b + \sigma_y^a \sigma_y^b). \tag{J15} \]

We perform numerical simulations of the Lindblad master equation
\[ \frac{d\rho(t)}{dt} = -i[H_{RWA}^2, \rho(t)] + \sum_k \Gamma_k D(L_k)\rho(t) \tag{J16} \]
where
\[ D(L)\rho = L\rho L^\dagger - \frac{1}{2}(L^\dagger L, \rho), \]
is the dissipator and $\rho(t)$ is the system density matrix. We include collapse operators $\sigma_a^\mu, \sigma_a^\mu, \sigma_b^\mu$ and associated decay rates $\Gamma_1/2, \Gamma_1/2, \Gamma_2/2$, respectively ($\mu = \{a, b\}$). The factors of 1/2 in the decay and heating rates are due to the equal contributions of decay and heating to $T_1$, while the factor of 1/2 in the dephasing rate ensures that coherences decay at the appropriate rate [45]. By turning on each decoherence channel one at a time, we numerically estimate the contribution of each, see Table II. We then proceed to turn on all decoherence channels to estimate the overall fidelity reduction due to decoherence. For both the $\sqrt{6SWAP}$ and $\sqrt{SWAP}$ gates, decoherence makes up about half of the measured infidelity.

Analytical estimates of the first-order contribution of decoherence to the infidelity are given in Ref. [45], which we reproduce here for completeness
\[ F \approx 1 - \frac{d}{2(d+1)} \tau \sum_{\mu}(\Gamma_1^\mu + \Gamma_2^\mu), \tag{J17} \]
where $\tau$ is the duration of the gate and $d$ is the dimension of the Hilbert space. Infidelity estimates using this formula for single- and two-qubit gates can be found in Table II. We find excellent agreement between the numerical results and analytical formulas.
5. RF flux crosstalk

As discussed in Appendix D, flux crosstalk can contribute to dynamic frequency shifts and unwanted entanglement. Defining the maximum fluxes due to crosstalk $\xi_\mu$ through the qubit loops, the qubit drive amplitudes are $\Omega_{\pm,\mu} = \xi_\mu E_{L,\mu}(0)|\phi_\mu|1\rangle$, $\mu = a, b$. To estimate the infidelity associated with such crosstalk, we substitute Eq. (D1) for Eq. (J1), including the Gaussian envelope $f_i$.

$$\xi_\mu/2\pi = 10^{-3},$$

the infidelities rise to $1.1 \times 10^{-3}, 0.069$ for $\sqrt{3}\text{SWAP}$ and $\sqrt{3}\text{SWAP}$, respectively. These infidelities would be the leading source of error, necessitating the careful flux-crosstalk cancellation we undertook in our experiment.

6. Overall error budget

With all of the error channels included, the infidelities of single qubit gates, $\sqrt{3}\text{SWAP}$, and $\sqrt{3}\text{SWAP}$ gates are reduced to 99.97%, 99.87%, and 99.95% respectively. Our numerical simulations suggest that gate fidelity is predominantly constrained by decoherence. The remaining error is attributed to insufficient calibration and RF flux crosstalk.


