Fast, High-Fidelity Conditional-Phase Gate Exploiting Leakage Interference in Weakly Anharmonic Superconducting Qubits

M. A. Rol\textsuperscript{1,2}, F. Battistel\textsuperscript{1}, F. K. Malinowski\textsuperscript{1,2}, C. C. Bultink\textsuperscript{1,2}, B. M. Tarasinski\textsuperscript{1,2}, R. Vollmer\textsuperscript{1,2}, N. Haider\textsuperscript{1,3}, N. Muthusubramanian\textsuperscript{1,2}, A. Bruno\textsuperscript{1,2}, B. M. Terhal\textsuperscript{1,4}, and L. DiCarlo\textsuperscript{1,2}

\textsuperscript{1}QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

\textsuperscript{2}Kavli Institute of Nanoscience, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

\textsuperscript{3}Netherlands Organisation for Applied Scientific Research (TNO), P.O. Box 96864, 2509 JG The Hague, The Netherlands

\textsuperscript{4}JARA Institute for Quantum Information, Forschungszentrum Juelich, D-52425 Juelich, Germany

(Received 26 March 2019; published 18 September 2019)

Conditional-phase (cz) gates in transmons can be realized by flux pulsing computational states towards resonance with noncomputational ones. We present a 40 ns cz gate based on a bipolar flux pulse suppressing leakage (0.1\%) by interference and approaching the speed limit set by exchange coupling. This pulse harnesses a built-in echo to enhance fidelity (99.1\%) and is robust to long-timescale distortion in the flux-control line, ensuring repeatability. Numerical simulations matching experiment show that fidelity is limited by high-frequency dephasing and leakage by short-timescale distortion.

DOI: 10.1103/PhysRevLett.123.120502

A steady increase in qubit counts \cite{1–4} and operation fidelities \cite{5–9} allows quantum computing platforms using monolithic superconducting quantum hardware to target outstanding challenges such as quantum advantage \cite{10–12}, quantum error correction (QEC) \cite{13–17}, and quantum fault tolerance (QFT) \cite{18,19}. All of these pursuits require two-qubit gates with fidelities exceeding 99\%, fueling active research.

There are three main types of two-qubit gates in use for transmon qubits, all of which harness exchange interactions between computational states (|ii\rangle, i, j \in \{0, 1\}) or between computational and non-computational states (i or j \geq 2), mediated by a coupling bus or capacitor. Cross-resonance gates \cite{8,20} exploit the exchange interaction between |01\rangle and |10\rangle using microwave-frequency transversal drives. Parametric gates \cite{7,21} employ radio-frequency longitudinal drives, specifically flux pulses modulating the qubit frequency, to generate sidebands of resonance between |01\rangle and |10\rangle for iSWAP or between |11\rangle and |02\rangle or |20\rangle for conditional phase (cz). The oldest approach \cite{22,23} uses baseband flux pulses to tune |11\rangle into near resonance with |02\rangle to realize cz. Either because they explicitly use noncomputational states, or because of frequency crowding and the weak transmon anharmonicity, the three approaches are vulnerable to leakage of information from the computational subspace. Leakage is very problematic in applications such as QEC, complicating the design of error decoders and/or demanding operational overhead to generate seepage \cite{24–28}, generally reducing the error thresholds for QFT. This threat has motivated the design of fast-adiabatic pulses \cite{29} to mitigate leakage and architectural choices in qubit frequency and coupler arrangements \cite{30} to explicitly avoid it. Surprisingly, many recent demonstrations \cite{7,8,31} of two-qubit gates place emphasis on reaching or approaching 99\% fidelity without separately quantifying leakage.

Although baseband flux pulsing produces the fastest two-qubit gates to date (30–45 ns), two challenges have kept it from becoming the de facto two-qubit gating method. First, because the pulse displaces one qubit 0.5–1 GHz below its flux-symmetry point, i.e., the sweet spot, the sensitivity to flux noise increases dephasing and impacts fidelity. The second challenge is nonatomicity. If uncompensated, distortions in the flux-control lines originating from limited waveform-generator bandwidth, high-pass bias tees, low-pass filters, impedance mismatches, on-chip response, etc., can make the action of a pulse depend on the history of flux pulses applied. To date, predistortion corrections have been calculated in advance, requiring prior knowledge of the timing of all the flux-pulse-based operations required by the quantum circuit, and significant waveform memory. This standard practice is incompatible with real-time determination and execution of operations, as is required for control flow and feedback in a fully programmable quantum computer \cite{32,33}.

In this Letter, we introduce a fast (40 ns), low-leakage (0.1\%), high-fidelity (99.1\%), and repeatable flux-pulse-based cz gate suitable for a full-stack quantum computer executing operations in real time on transmon-based quantum hardware. These attractive characteristics are enabled by a zero-average bipolar flux-pulsing method, nicknamed net zero (NZ), which uses the |11\rangle \leftrightarrow |02\rangle avoided crossing twice. Harnessing the analogy to a Mach-Zehnder interferometer, NZ exploits destructive interference to minimize leakage to |02\rangle while approaching the speed limit set by the exchange coupling in the
two-excitation manifold. The flux symmetry of the transmon Hamiltonian makes the phases acquired by the pulsed qubit first-order insensitive to low-frequency flux noise, increasing fidelity relative to a unipolar pulse. Crucially, the zero-average characteristic makes NZ insensitive to long-timescale distortions remaining in the flux-control line after real-time precompensation, making the CZ gate repeatable.

Detailed numerical simulations supplied with calibrated experimental parameters and direct measurement of short-timescale distortions show an excellent match to experimental parameters and direct measurement of short-timescale distortions. By performing a convolution of the desired signal $\Phi_{\text{target}}(t)$ with $\tilde{h}^{-1}$, the qubit experiences the pulse

$$
\Phi(t) = h\bigcirc V_{\text{AWG}}(t) = h\bigcirc(\tilde{h}^{-1} * \Phi_{\text{target}})(t).
$$

The predistortion corrections are performed using a combination of real-time filters implemented in a Zurich Instruments HDAWG and a short (20 ns) FIR filter implemented off-line.

By eliminating the dc component of the pulse, NZ CZ gates are resilient to long-timescale distortions [45]. Because the transmon Hamiltonian is symmetric with respect to the sweet spot, it is possible to use both positive and negative amplitudes to perform a CZ gate [Fig. 1(b)] while satisfying the zero-average condition

$$
\int_0^{T_{\text{CZ}}} \Phi_{\text{target}}(t')dt' = 0.
$$

If Eq. (3) holds, the dc component is zero and the components in the Fourier transform $\Phi_{\text{target}}(\omega)$ at frequencies $\omega \gtrsim (2\pi/T_{\text{CZ}})$ are suppressed. Writing Eq. (2) in the Fourier domain: $\Phi(\omega) = \mathcal{H}(\omega)\tilde{h}^{-1}(\omega)\Phi_{\text{target}}(\omega)$, it follows that if $\Phi_{\text{target}}(\omega)$ does not contain any components at $\omega < (2\pi/T_{\text{CZ}})$, then $\Phi(\omega)$ does not depend on any components of $\mathcal{H}(\omega)$ at frequencies $\omega < (2\pi/T_{\text{CZ}})$. As a consequence, the required corrections for NZ pulses do not accumulate, eliminating the need for accurate long-timescale distortion corrections and the resulting history-dependent errors [Fig. 1(d)].

To measure the repeatability of CZ gates, the phase $\phi_{01}$ acquired by the pulsed qubit during a CZ gate is measured as a function of the separation time $T_{\text{Sep}}$ between pulses (Fig. 2). Because of the detuning from the sweet spot, a small change in amplitude during the pulse leads to a significant change in frequency. This makes the acquired phase sensitive to distortions. We observe that not correcting distortions leads to significant phase errors ($\sim 80^\circ$). Correcting distortions using a predistortion filter keeps the error small ($<10^\circ$) for the first 500 ns but shows history-dependent behavior for longer timescales. Using NZ pulses in combination with a predistortion filter eliminates all
history dependence. Hence, we conclude that NZ pulses are robust against remaining long-timescale distortions.

We next investigate a built-in echo effect that provides protection against flux noise. Because the derivative of the flux arc is equal and opposite in sign at the positive and negative halves of the NZ pulse, we expect $\phi_{01}$ and $\phi_{2Q}$ to be first-order insensitive to low-frequency flux noise. As a test, we measure the dependence of $\phi_{2Q}$ on an applied dc flux offset for both a unipolar and NZCZ gate [Fig. 3]. As shown in Fig. 3(b), $\phi_{2Q}$ is first-order (second-order) sensitive for a unipolar (NZ) pulse. We have also measured how the dephasing time depends on the detuning for both a square flux pulse and two half-square flux pulses with opposite sign [34]. We find that the dephasing rate is significantly reduced when the opposite-sign flux pulses are used, confirming that NZ pulses have a built-in echo effect.

The pulse shape is intended to minimize leakage and is described by two parameters [34]. Parameter $\theta_f$ is a measure of the flux at the middle of the unipolar pulse, and at the middle of each half of NZ. States $|j_{11}\rangle$ and $|j_{02}\rangle$ are resonant at $\theta_f = \pi/2$. Parameter $\lambda_2$ tunes the sharpness of the pulse rise and fall. We follow Ref. [37] in defining the leakage ($L_1$) of an operation as the average probability that a random computational state leaks out of the computational subspace.

In order to gain insight into how $\phi_{2Q}$ and $L_1$ depend on the pulse shape, we perform an experiment and compare this to simulations. The conditional oscillation experiment (Fig. 4) consists of a Ramsey-like experiment that allows us to measure $\phi_{2Q}$ and estimate $L_1$. This experiment measures the phase acquired during an (uncalibrated) CZ gate by the target qubit ($q_{\text{targ.}}$) while either leaving the control qubit ($q_{\text{contr.}}$) in the ground state, or adding an excitation to $q_{\text{contr.}}$. The difference between the phase acquired when $q_{\text{contr.}}$ is in $|j_{11}\rangle$ and when $q_{\text{contr.}}$ is in $|j_{02}\rangle$ gives $\phi_{2Q}$. If leakage from $|j_{11}\rangle$ to $|j_{02}\rangle$ occurs, $q_{\text{contr.}}$ is in $|j_{02}\rangle$ when the second $\pi$ pulse is applied, adding, instead of removing, an excitation to $q_{\text{contr.}}$. The leakage probability $L_1$ can be estimated as $L_1 = m/2$, where $m$ is the number of oscillations between the $\pi/2$ pulses.
where \( m \) is the population difference on the control qubit between both variants of the experiment. Because of relaxation effects, \( L_1 \) slightly overestimates \( L_1 \).

The simulations model the system realistically and allow us to extract \( \phi_{20} \), \( L_1 \), and the average gate fidelity \( F \) for a single application of the gate [34]. The pulse is modeled as a trajectory in a two-qutrit Hamiltonian. The noise model accounts for relaxation and dephasing effects as well as the effect of remaining distortions. The latter are measured using the cryoscope technique [35].

Both experiment and simulation show a fringe of low leakage [Figs. 4(b) and 4(d)]. This fringe can be understood as “leakage interference” between \( |11 \rangle \) and \( |02 \rangle \) by analogy to a Mach-Zehnder interferometer [34]. Such analogy has been exploited in a variety of platforms [39–42,46] to demonstrate coherent control of a single qubit by showing Stückelberg oscillations [38] as a consequence of periodic driving of the qubit into an avoided crossing. Here we pulse in-and-out of \( |11 \rangle \leftrightarrow |02 \rangle \) twice to realize low-leakage two-qubit gates. The states \( |11 \rangle \) and \( |02 \rangle \) correspond to two paths of the interferometer. The first part of the NZ pulse (red in Fig. 1) corresponds to the first (imbalanced) beam splitter. In general, after the first beam splitter most of the population remains in \( |11 \rangle \) but part is transferred to \( |02 \rangle \).

Pulsing through the sweetspot (green in Fig. 1) corresponds to the arms of the interferometer. The two paths are detuned by \( \sim 800 \) MHz, causing a phase to be acquired before the paths are recombined at the second half of the NZ pulse (blue in Fig. 1) corresponding to the second beam splitter. The phase difference between the two paths will cause interference that either enhances or suppresses the leakage to \( |02 \rangle \).

Given the good correspondence between experiment and simulation (Fig. 4), we can use simulations to explore the parameter space \( (\theta_f, \lambda_2, T_{20}) \) to find the shortest \( T_{20} \) enabling a high-fidelity, low-leakage CZ gate. The minimum CZ gate duration is fundamentally limited by the coupling strength \( J_2 \) as the time required to acquire 180° of conditional phase at the avoided crossing: \( T_{20} \geq (\pi/J_2) = 25 \) ns. We find a \( T_{20} = 28 \) ns NZ pulse using leakage interference to achieve low leakage. The use of interference is demonstrated by the fact that the corresponding half pulse displays high leakage [34]. We append \( T_{10} = 12 \) ns flux pulses on both qubits to correct the single-qubit phases, making the total duration of the phase-corrected CZ gate \( T_{CZ} = 40 \) ns. We ensure that these phase-correction pulses satisfy Eq. (3) and have a sufficiently low amplitude to not affect \( \phi_{20} \) and \( L_1 \) significantly.

We characterize the performance of the CZ gate using an interleaved randomized benchmarking protocol [5,44] with modifications that allow us to quantify leakage [34,37,43]. The randomized benchmarking sequences are based on 300 random seeds. For each seed, every data point is measured 104 times. We measure an average gate fidelity \( F = 99.10\% \pm 0.16\% \) and leakage \( L_1 = 0.10\% \pm 0.07\% \) for the NZ pulse with \( T_{CZ} = 40 \) ns [Figs. 5(a) and 5(b)]. We could not perform similar measurements for the unipolar pulse since this gate is not repeatable, as demonstrated in Fig. 2.

It is possible to investigate the limits to the performance of the NZ CZ using simulation and compare to the unipolar CZ, even though this is not possible in experiment since the unipolar CZ lacks the required characteristic of being repeatable. We simulate these gates for a range of different error models [Figs. 5(c) and 5(d)]. For each to optimize over \( \theta_f \) and \( \lambda_2 \) to find the lowest \( \epsilon \) and the corresponding \( L_1 \). A first observation is that the infidelity \( (1 - F) \) of the NZ gate does not significantly increase when the low-frequency flux-noise components are included, whereas this does affect the unipolar pulse. It appears that the difference in \( \epsilon \) between the unipolar and NZ pulses for the full model can be attributed completely to this effect. This observation is consistent with the echo effect demonstrated in Fig. 3. Looking at the \( L_1 \) error budgets, \( L_1 \) is limited by short-timescale distortions. This is understandable as minimizing \( L_1 \) requires the pulse to follow a precise trajectory. Distortions also increase \( \epsilon \) through \( L_1 \) [34]. The simulations also indicate that dephasing causes leakage. This can be understood as dephasing effectively corresponds to an
uncertainty in the energy levels. The simulated $L_1$ is larger than the measured $L_1$. This could be explained in two ways, either the distortions are less severe than our estimate, or the simulations, only concerned with a single application of the gate, do not take into account all the relevant effects. Specifically, because the population in the leakage subspace does not completely decohere, this population can seep back into the computational subspace due to an interference effect (similar to that in the NZ pulse itself) at subsequent applications of the gate. Because the first CZ gate cannot benefit from this coherence, the simulations, which only deal with a single CZ gate, slightly overestimate the effective leakage.

In summary, we have demonstrated a flux-based CZ gate for transmon qubits that is fast, low-leakage, high-fidelity, and repeatable. The gate is realized using a bipolar net-zero flux pulse that harnesses leakage interference to achieve speed while maintaining low leakage. The NZ pulse exploits the flux symmetry of the pulsed transmon to build in an echo effect on its single-qubit phase and the conditional phase, increasing fidelity relative to a unipolar pulse. Finally, the action of the NZ pulse is robust to long-timescale distortions in the flux-control line remaining after real-time precompensation, enabling the repeatability of the CZ gate. These features make the realized NZ CZ gate immediately useful in high-circuit-depth applications of a full-stack quantum computer in which a controller issues operations to execute on the quantum hardware in real time. For example, current work in our group uses NZ CZ gates to stabilize two-qubit entanglement by multi-round indirect parity measurements [16]. Future work will incorporate NZ CZ gates into our scheme [30] to realize a surface-code-based logical qubit [18] with monolithic transmon-cQED quantum hardware.

We thank J. Helsen, T. O’Brien, X. Bonet-Monroig, N. Haandaek, Y. Salathe, and V. Ostroukh for useful discussions. This research is supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. Army Research Office Grant No. W911NF-16-1-0071. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. F. B. and B. M. T. are supported by ERC Grant EQEC No. 682726.


[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.123.120502 for detailed information on experimental protocols and simulations performed in this work, which includes Refs. [29,30,35–44].