Practical implications of SFQ-based two-qubit gates

Mohammad Reza Jokar*
Computer Science Department
University of Chicago
Chicago, USA
jokar@uchicago.edu

Richard Rines*
Computer Science Department
University of Chicago
Chicago, USA
richrines@uchicago.edu

Frederic T. Chong
Computer Science Department
University of Chicago
Chicago, USA
chong@cs.uchicago.edu

***Abstract*—Scalability of today's superconducting quantum computers is limited due to the huge costs of generating/routing microwave control pulses per qubit from room temperature. One active research area in both industry and academia is to push the classical controllers to the dilution refrigerator in order to increase the scalability of quantum computers. Superconducting Single Flux Quantum (SFQ) is a classical logic technology with low power consumption and ultra-high speed, and thus is a promising candidate for in-fridge classical controllers with maximized scalability. Prior work has demonstrated high-fidelity SFQ-based single-qubit gates. However, little research has been done on SFQ-based multi-qubit gates, which are necessary to realize SFQ-based universal quantum computing.**

In this paper, we present the first thorough analysis of SFQ-based two-qubit gates. Our observations show that SFQ-based two-qubit gates tend to have high leakage to qubit non-computational subspace, which presents severe design challenges. We show that despite these challenges, we can realize gates with high fidelity by carefully designing optimal control methods and qubit architectures. We develop optimal control methods that suppress leakage, and also investigate various qubit architectures that reduce the leakage. After carefully engineering our SFQ-friendly quantum system, we show that it can achieve similar gate fidelity and gate time to microwave-based quantum systems. The promising results of this paper show that (1) SFQ-based universal quantum computation is both feasible and effective; and (2) SFQ is a promising approach in designing classical controller for quantum machines because it can increase the scalability while preserving gate fidelity and performance.

***Index Terms*—SFQ-based quantum gate, Quantum control, Scalable quantum computer, Cryogenic electronic**

I. Introduction

A great milestone in quantum computing is the recent development of quantum computer prototypes thanks to great efforts in industry and academia. Superconducting quantum computing is one of the most promising technologies to realize a quantum computer, and has been used to realize quantum computer prototypes with $<$100 qubits [1], [2], [5], [7], [12], [22]. These prototypes rely on sending analog microwave signals per qubit from a classical controller at room temperature to the quantum chip inside a dilution refrigerator in order to perform quantum operations. Unfortunately, this scheme introduces severe scalability challenges due to high costs of electronics that are used to generate the microwave signals at room temperature, as well as heat dissipation inside the dilution refrigerator caused by routing the high-bandwidth signals to the quantum chip [10], [13], [16]. Thus, design decisions must be made to address the scalability challenges of today's quantum computer prototypes and realize large-scale quantum computers, which are essential in running many quantum algorithms and also performing quantum error correction.

One active research area in industry and academia is designing in-fridge classical controllers to increase the scalability of quantum machines by generating and routing the control signals locally. Due to maturity of CMOS logic, Cryo-CMOS is one attractive logic technology to build in-fridge controllers. Prior work demonstrated Cryo-CMOS controller prototypes that generate microwave control pulses inside the dilution refrigerator, and can scale to hundreds of qubits given the power budget of dilution refrigerators [23]. Meanwhile, Superconducting Single Flux Quantum (SFQ) is proposed as an alternative logic technology in the literature. SFQ logic is less mature than CMOS but can maximize the scalability of in-fridge controllers due to its very low power consumption and ultra-high speed [10], [13], [14], [16].

SFQ-based controllers can perform quantum operations by generating a train of SFQ pulses (instead of microwave control pulses) inside the dilution refrigerator and applying them directly to the qubits [13], [14]. Prior work demonstrated high-fidelity single-qubit gates with low leakage to the non-computational subspace using SFQ pulses [10], [13]. Prior work also demonstrated SQF-based two-qubit gates considering a quantum system model that takes into account only the first two energy levels of the qubits (i.e., qubit computational subspace) [3]. However, there is a lack of a detailed analysis in the literature on high-fidelity SFQ-based two-qubit gates with low leakage to the non-computational subspace. Thus, a key unanswered question is: are SFQ-based two-qubit gates with high fidelity and low leakage feasible and effective? In this paper, we present the first thorough study on SFQ-based two-qubit gates, and demonstrate that we can realize them with high fidelity and low leakage by carefully designing our quantum optimal control method, and qubit architecture.

We first demonstrate that it is essential to take higher energy levels of the qubits into consideration in our optimal control method. Similar study has been done in the literature on SFQ-based single-qubit gates [14] where the authors show that taking into account the first three lowest energy levels (i.e., the qubit computational subspace and one higher energy level)

*These two authors contributed equally.

in the optimal control method is sufficient to find high-fidelity gates with low leakage to the high energy levels. In this paper, we show that SFQ-based two-qubit gates have much higher tendency to leak to higher energy levels, thus it is challenging to find high-fidelity gates even if we take into account up to five energy levels in our optimal control method. Thus, we must take further steps by developing SFQ-based optimal control methods to suppress leakage and investigating different qubit architectures and configurations that reduce leakage.

We first study transmon qubit devices with Ω_x control fields which are widely used in both SFQ-based and microwave-based systems [3], [11], [13]. We show that we can realize high-fidelity SFQ-based two-qubit gates with low leakage to higher energy levels after carefully designing our optimal control method and tuning the qubit configurations. We then investigate two possible extensions of this design in order to reduce the SFQ-based two-qubit gate time while keeping the leakage low: (1) the addition of σ_z SFQ control pulses implemented via frequency-tunable split-transmon devices; and (2) the use of SFQ control pulses in combination with high-anharmonicity fluxonium qubits.

Finally, we compare our SFQ-friendly quantum system with microwave-based quantum systems, and show that we can achieve similar gate fidelity and gate time using SFQ. This shows that SFQ is a promising approach to implement classical controllers for quantum computers because it can maximize the scalability of quantum computers due to the unique characteristics of SFQ logic, while delivering similar fidelity and performance to that of state-of-the-art microwave-based systems.

To summarize, our key contributions are as follows:

- We present the first study of SFQ-based two-qubit gates that takes into consideration the leakage to higher energy levels.
- We identify and discuss the main challenge in realizing high-fidelity SFQ-based two-qubit gates, which is high leakage to non-computational qubit subspace.
- We develop optimal control methods that suppress the population of higher energy levels during the execution of the two-qubit gates.
- We study various qubit architectures and configurations in an attempt to engineer a quantum system with low leakage.
- We engineer an SFQ-friendly quantum system, and show that it can achieve similar gate fidelity as that of microwave-based system – a promising result.

The rest of the paper is organized as follows. Sec. II presents a background on qubit architecture and configurations, quantum optimal control and SFQ-based gates, followed by the a discussion on the motivation of our paper. Sec. III presents our methodology and the results of our detailed study on SFQ-based two-qubit gates, followed by a comparison with microwave-based two-qubit gates. Finally, Sec. IV concludes the paper.

II. Background and Motivation

Here we provide details of the physical systems we are targeting in order to distill the basic toolbox of quantum operations available to us for implementing high-performance SFQ-based quantum gates. We motivate our analysis by describing the challenges of implementing high-fidelity gates on realistic quantum systems, the existing strategies for overcoming them on systems with analog control, and prior work on SFQ-based gates aiming to do the same.

A. Physical system

The evolution of a quantum systems are governed by their Hamiltonian. For universal quantum computation, we require that the system provide (1) well-defined *qubits*, or separable two-level quantum subsystems which can be independently initialized and measured; (2) a mechanism for generating entanglement between these qubits; and (3) a method for precisely controlling the system's evolution [4] For the purposes of this investigation, we consider pairs of statically-coupled superconducting qubits, with the overall system Hamiltonian,

$$\hat{H} = \sum_q \hat{H}_q + \sum_q \hat{H}_{q,d}(t) + \hat{H}_{qq}, \tag{1}$$

where $\hat{H}_q$ are the static Hamiltonian of each qubit device, $\hat{H}_{q,d}$ result from the time-dependent control signals applied to each qubit, and $\hat{H}_{qq}$ is contributed by the inter-qubit coupling (and therefore is responsible for entanglement generation). In the following, we express these terms for various hardware configurations in terms of conjugate flux and charge number quantum operators $\hat{\phi}$ and $\hat{n}$, where $[\hat{n}, \hat{\phi}] = i$.

1) Transmons: The superconducting transmon qubit comprises a Josephson junction (JJ) shunted to ground with a capacitor in order to minimize its sensitivity to charge noise [8], [20]. The transmon Hamiltonian can be written,

$$\hat{H}_q = 4E_C\hat{n}^2 - E_J\cos\hat{\phi}, \tag{2}$$

where $E_C = e^2/2C_q$ indicates the capacitive energy (with C_q including both the shunt capacitance and that of the JJ) and $E_J = I_c\Phi_0/2\pi$ is the Josephson energy of a transmon with critical current I_c.

The spectrum of the single-transmon system is found be diagonalizing Eq. (2). For $E_C \ll E_J$ the transmon Hamiltonian can be expanded in the SHO Fock state basis,

$$\hat{H}_q \approx \omega_{01}\hat{a}^\dagger\hat{a} + \frac{\alpha}{2}\hat{a}^\dagger\hat{a}(\hat{a}^\dagger\hat{a} - 1), \tag{3}$$

where $\omega_{01} = \sqrt{8E_JE_C} - E_C$ is the qubit's oscillation frequency (that is, the energy gap between the ground and first excited state), $\alpha = \omega_{12} - \omega_{01} = -E_C$ is its anharmonicity, and we have made the substitution,

$$\hat{n} = \sqrt{E_J/32E_C}\big(\hat{a} + \hat{a}^\dagger\big), \quad \hat{\phi} = i\sqrt{2E_C/E_J}\big(\hat{a} - \hat{a}^\dagger\big), \tag{4}$$

using the standard creation (annihilation) operators $\hat{a}^\dagger$ ($\hat{a}$). Typical transmon qubits are configured with oscillation frequencies $\omega_{01}/2\pi$ between 3 and 6 GHz and anharmonicity

$\alpha/2\pi$ between 100 and 300 MHz [9]. The nonzero anharmonicity makes it possible to isolate and address the system's $\{|0\rangle, |1\rangle\}$ subspace, providing the required well-defined two-level qubit.

2) Frequency-tunable transmons: The single-JJ transmon's oscillation frequency is fixed by its hardware components. We can instead construct a *frequency-tunable* transmon by splitting its single JJ into a pair of parallel junctions (dc-SQUID) and driving an external magnetic flux φ_e through the enclosed loop. In this case the junction energy E_J in Eq. (2) is replaced with the flux-dependent effective energy [9],

$$E'_J = \sqrt{(E_{J1} + E_{J2})^2 \cos^2 \varphi_e + |E_{J1} - E_{J2}|^2 \sin^2 \varphi_e}, \quad (5)$$

where $E_{J1,2}$ are the Josephson energies of the respective JJs and φ_e is the applied flux in units of $\pi\Phi_0$. The applied flux can then be used to tune the qubit's oscillation frequency, or equivalently implement z-axis rotations of the qubit.

For multi-qubit systems, flux control has also been employed to implement two-qubit gates by inducing resonant oscillations between multi-qubit states: for example, by bringing the qubit frequencies together, coherent oscillations between the $|01\rangle$ and $|10\rangle$ state will generate the iSWAP (or $\sqrt{\text{iSWAP}}$) gate, whereas a CZ gate can be implemented using the resonance between the $|11\rangle$ and $|02\rangle$ (or $|20\rangle$). The latter case takes advantage of the higher energy levels of the transmon system, allowing the quantum state to temporarily leave the two-level qubit subspace during the execution of the gate. Frequency-tunable transmons enable fast resonant two-qubit operations while decreasing crosstalk by allowing noninteracting qubits to be "parked" at well-separated oscillation frequencies. Frequency tunability comes at the cost of added complexity and sensitivity to magnetic flux noise.

3) Fluxonium: Though the transmon's nonzero anharmonicity makes it possible to target the two-level (qubit) subspace for quantum computation, its weakness relative to the oscillation frequency makes it prone to leakage to higher level states. Alternative qubit technologies such as fluxonium [15] have been shown to increase anharmonicity with minimal cost in terms of noise sensitivity. The fluxonium qubit is constructed similarly to the transmon, but with an additional inductive shunt to ground implemented using an array of Josephson junctions connected in series. The resulting Hamiltonian is [9],

$$\hat{H}_q = 4E_C\hat{n}^2 + E_L\hat{\phi}^2 - E_J \cos\Big(\hat{\phi} + \varphi_e\Big), \quad (6)$$

where $E_L \ll E_J$ is the inductive energy of the junction array and φ_e is an external magnetic flux through the qubit loop.

Fluxonium's sensitivity to flux noise is minimized at $\varphi_e = 0$ and $\varphi_e = \pi$, where symmetry ensures that the energy dependence on φ_e vanishes to first order. In the latter case, the qubit's oscillation frequency ω_{01} is significantly reduced relative to that of the subsequent transition (ω_{12}), resulting in large, positive anharmonicity. It is less trivial to approximate the fluxonium spectrum analytically; instead we diagonalize Eq. (6) numerically to determine the computational basis states and energy spectrum of our system. With typical hardware configurations, fluxonium qubits at $\varphi_e = \pi$ have $\omega_{01} \sim 1$ GHz, while ω_{12} is 2-5 times larger. For remainder of this paper, we assume that fluxonium is operating with a $\varphi_e(t) = \pi$ static bias flux.

4) Coupling: We focus on systems with static coupling between qubits, such that the interaction Hamiltonian H_{qq} is constant and uncontrollable (as opposed to, for example, tunable coupling systems [1] which allow the interaction to be switched on and off on-demand but which may complicate the implementation of an SFQ-based controller). For superconducting qubits coupled via a capacitance C_{qq}, the coupling Hamiltonian in Eq. (1) is,

$$\hat{H}_{qq} = g_{qq}\hat{n}_{q_0}\hat{n}_{q_1}, \quad (7)$$

where $g_{qq} = 4e^2C_{qq}/C_{q_0}C_{q_1}$ quantifies the coupling strength. Expressed in Eq. (9) in the energy-basis rest frame of the undriven qubit, the dominant matrix elements of the system Hamiltonian (after the rotating wave approximation) are,

$$\hat{H}^{rf}_{qq}(t) = J\sum_k c^{(0)}_{k-1,k}c^{(1)}_{l,l-1}e^{i(\omega^{(0)}_{k,k-1} - \omega^{(1)}_{l-1,l})t} \times \big(|k,l-1\rangle\langle k-1,l| + h.c..\big), \quad (8)$$

where J is a normalized coupling constant, and $c_{k,k-1} = c_{k-1,k} \approx \sqrt{k}/2$ for transmons whereas for fluxonium must be computed numerically for each qubit by diagonalizing its Hamiltonian (Eq. (6)). Though this coupling cannot be disabled, the *effective* coupling between qubits is inversely proportional to the separation between the two qubits' oscillation frequencies due to destructive interference caused by time-averaging the rotating phases. We can therefore preserve the independence of the qubits by designing the system such that frequencies of coupled qubits are well separated.

5) Drive: The most common architecture for manipulating statically-coupled qubits is to apply microwave control signals directly to the qubits via a coupling capacitor. Given a time-dependent voltage source $V_d(t)$, the microwave drive Hamiltonian is,

$$\hat{H}_{q,d} = V_d(t)\frac{C_d}{C_d + C_q}\hat{n}, \quad (9)$$

where C_d is the capacitance of the coupling capacitor. Expressed in the rest frame of the qubit and assuming a microwave drive $V_d(t) = \Omega_x(t)V_0\cos\omega_d t$ (where $\Omega_x(t)$ is the pulse envelope and $V_0 = \sqrt{32E_C/E_J}(1 + C_d/C_q)$ absorbs the details of the drive hardware),

$$H^{rf}_d(t) = \Omega_x(t)\sum_k c_{k+1,k}e^{i(\omega_{k,k+1} - \omega_d)t}\,|k+1\rangle\langle k| + h.c.. \quad (10)$$

The time-dependent phases in Eq. (10) allow us to selectively drive a given transition while others are suppressed by the time-dependent phase. For example, continuously driving with $\omega_d = \omega_{01}$ will drive Rabi oscillations in the qubit subspace while the qubit's nonzero anharmonicity $\omega_{12} - \omega_{01} = \alpha$ will suppress the $|1\rangle \leftrightarrow |2\rangle$ transition. However, in order to have

a finite gate time, the envelope $\Omega_x(t)$ must itself contain Fourier components which can diminish this suppression by overlapping with higher-order transitions, especially given the transmon's relatively small anharmonicity. Analytical pulse shaping models such as the DRAG scheme [17] are therefore employed to precisely minimize the overlapping frequency components in the pulse shape.

Using the cross-resonance interaction [18], [19] it is also possible to induce two-qubit entangling operations with a precisely detuned control signal applied to one qubit, making fixed-frequency transmons and microwave control sufficient for universal quantum computation. Successful quantum computer prototypes have been developed using this control mechanism alone [2]. However, the speed of cross-resonance gates is proportional to the effective coupling between the qubits, creating a tradeoff between gate time and crosstalk.

B. Microwave optimal control

In practice, the broad control schemes outlined in Sec. II-A5 are insufficient for high-precision quantum gates. Analytical leakage suppression schemes are especially challenging for multi-qubit systems due to the exponentially increasing complexity of the energy spectrum and the contributions of the coupler. This complexity is especially prevalent for cross-resonance gates, in which some transition between multi-qubit states is intentionally driven while others must be suppressed. Further, it is often desirable to allow the system to evolve outside the two-level subspace during the execution of the gate as it provides more possible paths for realizing complicated operations within a short gate time. (c.f. the frequency-tunable CZ implementation described above). Typical microwave systems therefore employ search-based optimal-control strategies such as the ubiquitous gradient ascent pulse engineering (GRAPE) tool [11] to generate pulse waveforms which implement quantum gates with high fidelity and low leakage.

C. Fidelity functions

We quantify the performance of learned gates using two variants of average gate fidelity. In its general form, the average gate fidelity of a quantum operation $\mathcal{E}$ relative to a unitary target gate T is defined,

$$\overline{F}(U,T) = \int d\psi \, \langle\psi|T^\dagger \mathcal{E}(\psi) T|\psi\rangle \,, \tag{11}$$

where the average is over the normalized Haar distribution of quantum states. Because we are interested only in how the gate affects *qubits*, we would like our fidelity metric to (1) be agnostic to the behavior of the gate when applied to states outside the qubit subspace, and (2) penalize gate-induced leakage from within the computational subspace. We therefore define,

$$\mathcal{E}(\psi) = \Pi U \Pi \, |\psi\rangle\langle\psi| \, \Pi U^\dagger \Pi \tag{12}$$

$$\Pi = (|0\rangle\langle 0| + |1\rangle\langle 1|)^{\otimes n}, \tag{13}$$

where U is the simulated (unitary) evolution including higher level states and Π projects it into the qubit subspace. The average in Eq. (11) is then taken over just states $|\psi\rangle$ states in the two-level subspace $SU(2^n)$, so that a *subspace-averaged gate fidelity* can be calculated [6],

$$\overline{F_1}(U,T) = \frac{\mathrm{tr}\left(\Pi U \Pi U^\dagger \Pi\right) + \mathrm{tr}\left(T \Pi U^\dagger \Pi\right)^2}{2^{2n} + 2^n}, \tag{14}$$

Because of the constrained control set available with SFQ control, we would further like to broaden our search target as much as possible. In practice, single-qubit Z rotations can often be commuted through subsequent gates or implemented virtually. We therefore define a second *Z-independent gate fidelity* metric, which is independent of trailing Z-rotations:

$$\overline{F_2}(U,T) = \sup_{\vec{\alpha}} \overline{F_1}\left(Z_{\vec{\alpha}} \tilde{U}, T\right), \tag{15}$$

$$Z_{\vec{\alpha}} = Z(\alpha_1) \otimes \cdots \otimes Z(\alpha_n). \tag{16}$$

Finally, we can explicitly quantify leakage by computing the probability of measuring a state outside the qubit subspace after applying the gate to a state initially within that space. Again averaging over the uniform distribution of all possible two-level input states, the *average leakage* can be calculated,

$$\overline{L}(U) = 1 - \frac{\mathrm{tr}\left(\Pi U \Pi U^\dagger \Pi\right)}{2^n}. \tag{17}$$

From Eqs. (14), (15) and (17) one can show that $\overline{F_1}(U,T) \leq \overline{F_2}(U,T) \leq 1 - \overline{L}(U)$, so that as desired our fidelity metrics are upper-bound by the degree of leakage.

D. SFQ control

It has been proposed that quantum gates be implemented by applying SFQ pulses to the qubit directly in place of analog microwave control signals. The gate implementation is then described by a binary pulse train as shown in Fig. 1, where in each cycle of the SFQ clock a pulse is either applied or not applied to the qubit.

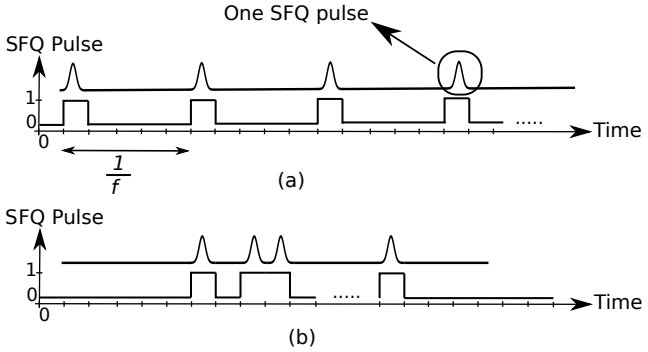


Fig. 1: Bit representation of SFQ pulse trains. (a) coherent pulses are applied to the qubit (1 pulse per qubit oscillation period) to perform rotations around the y axis. (b) a bitstream found by genetic algorithm to perform arbitrary unitary. Bitstreams are processed one bit at a time; if the bit is "0", no pulse is applied to the qubit, and if the bit is "1", one SFQ pulse is applied to the qubit.

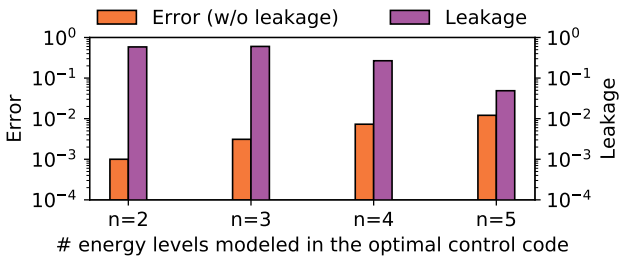


Fig. 2: Error and leakage of the best SFQ-based CZ gate found by the genetic algorithm on transmon qubit devices with Ω_x control fields. The error number reported in this plot does not take into consideration the leakage to higher energy level and is solely calculated by measuring the overlap between the operated unitary and the target unitary. Thus, low error does not necessarily translate to low leakage.

A single SFQ pulse is as a rapid Gaussian voltage waveform,

$$V_d(t) = \frac{\Phi_0}{\sqrt{2\pi\tau^2}} e^{-t^2/2\tau^2}, \tag{18}$$

with a total area of exactly $\int V_d(t) = \Phi_0$ and a typical pulse width of $\tau = 0.25$ ps. Approximating $V_d(t) \approx \Phi_0 \delta(t)$ and considering Eq. (9) in the energy-basis rest frame of the transmon, we expect a single pulse at time t_0 to implement the instantaneous gate,

$$U_x^{rf} = \exp\left\{-i\delta_\theta \sum_k e^{i\omega_{k,k+1} t_0} \sqrt{k}\, |k\rangle\langle k-1| + h.c..\right\}, \tag{19}$$

where $\delta_\theta \approx \Phi_0/V_0$ is the *tip angle*, indicating the rotation angle produced by a single pulse in the qubit subspace. The tip angle is typically in the range of 10^{-3} to 10^{-1} radians [3], [13], and is directly configurable via choices of qubit and coupling hardware; in our analysis we find that this configuration is extremely important for achieving high-performance SFQ gates.

In order to expand our narrow SFQ control toolset, we also consider a SFQ-based σ_z operation for frequency-tunable transmons. In this case, rather applying pulses to the qubit via a capacitive coupler, we assume that they are inductively coupled to the split transmon's dc-SQUID loop. Again approximating a single pulse as a delta function, the resulting gate is then simply,

$$U_z = \sum_k e^{ik\delta_z} |k\rangle\langle k|, \tag{20}$$

where δ_z is the z-axis tip angle determined by the hardware configuration (where in this case we may require a minimal amount of additional hardware to broaden the pulse shape in order to achieve a non-negligible phase kick). Though on a single-qubit system σ_z control would not be sufficient for universal quantum control, it turns out to be remarkably effective for realizing two-qubit gates when combined with the free evolution due to the static coupler.

Unlike the microwave drive, with SFQ pulses we cannot simply select a drive frequency in order to selectively drive a given transition while off-resonant transitions are suppressed by the rotating phase in Eq. (19). Instead, we are limited to selecting discrete clock cycles t in which to apply $U_x(t)$. If we constrain our system to the qubit subspace ($k \leq 1$), Eq. (19) is simply a unitary rotation $R_{\omega t}(\delta_\theta)$ by angle δ_θ about a time-dependent axis on the xy-plane. In this case the problem is reduced to one of single-qubit gate composition (taking as basis gates the set of lab-frame single-clock-cycle unitaries generated by applying pulses to each possible subset of qubits), for which many analytic and search methods have been studied. Empirically, in both prior work and our own examination it appears that pulse trains implementing high-fidelity, low leakage single-qubit gates are still readily discoverable when we model the system with additional energy states [13], [14], [16]. This is perhaps unsurprising observing Eq. (19): though each pulse may result in some population transfer out of the qubit subspace, the simple energy spectrum of a single qubit near its ground state makes it reasonable to expect symmetries to exist in which pairs or small groups of pulses will generate destructive interference in the non-qubit subspace (in fact, such symmetries we employed explicitly as part of the search algorithm outlined in [13]).

E. Prior work on SFQ-based gates and the motivation of this paper

There has been detailed analysis of SFQ-based single-qubit gates in the literature [13], [14], [16]. Prior work has studied the SFQ-based coherent control of qubits, and demonstrated that we can perform rotations around the X or Y axis by repeatedly applying SFQ pulses every qubit oscillation period [10], [16]. Prior work also reveal that the quantum gates based on coherent SFQ pulses suffer from leakage to higher energy levels, thus they utilize genetic algorithm to find better SFQ-based gates with low leakage and also short gate time [14], [16]. They show that taking into account three lowest energy levels of the qubit in their model is sufficient to realize low-leakage gates using SFQ pulses.

In [16], the authors envision the possibility of performing SFQ-based two-qubit gates. In [3], the authors implement a quantum optimal control version of the AlphaZero learning algorithm [21] to optimize the quantum dynamics, and use SFQ-based optimal control as a benchmark in their study. The authors show that they can find SFQ pulse trains to do $\sqrt{ZX}$ gate with high fidelity. However, their model of a two-qubit quantum system does not take into consideration the leakage out of the computational subspace.

Fig. 2 shows the importance of taking higher energy levels into consideration when learning SFQ pulses to perform two-qubit gates. In each case, we report the error of the best SFQ-based two-qubit gate we find with a genetic algorithm when modeling the quantum system using n energy levels. We then simulate the learned bitstream using a model that allows for evolution to higher energy levels, and report the leakage (Eq. (17)) of the resulting gate.

TABLE I: The parameters used in the genetic algorithm.

Population size	70
Selection size	60
Mutation probability	0.001
Maximum number of iterations	200,000
Target fidelity	0.999

We can easily find SFQ-based two-qubit gates with 0.999 fidelity with $n = 2$ (as shown in prior work as well [3]). However, we find that the learned SFQ pulse train results in a gate with high leakage when allowed to evolve out of the two-level subspace. This is an expected result—prior work has shown that we need to consider $n = 3$ to find low-leakage single-qubit gates [14]. What is more surprising is that as shown in Fig. 2, the genetic algorithm cannot find SFQ-based two-qubit gates with low leakage even with $n = 5$. For each n, the evolution resulting from a learned sequence continues to longer be contained within the n-level subspace in which the sequence was learned when higher energy levels are added to the model. This modified evolution will result in a gate with both poor accuracy and high leakage. Thus, unlike the single-qubit gate case, taking the higher energy level into consideration alone is not sufficient.

In this paper, we characterize the requirements of realizing high-fidelity SFQ-based two-qubit gates. We develop quantum optimal control methods, and also investigate various qubit architectures and configurations in an attempt to engineer an SFQ-friendly quantum system that can perform high-fidelity SFQ-based two-qubit gates.

III. Detailed study of SFQ-based two-qubit gates

In this section we first discuss our methodology, followed by the results of our study on SFQ-based two-qubit gates under various qubit architectures and configurations. Then, we compare our results with that of microwave-based quantum systems.

A. Methodology

We model SFQ-based quantum operations by numerically integrating the relevant system Hamiltonian (Eq. (1)) over a single SFQ clock cycle for each possible combination of input pulses. The learning algorithm then searches for pulse streams corresponding to optimal sequences of these basis operations. In order to avoid sequences which would spill into higher levels if made available (as described in Sec. II-E), we generate each unitary evolution using extra energy levels, and then project out the extra levels after each pulse in the sequence. The resulting non-unitarity of the evolution is then quantified by our fidelity metrics as additional leakage, forcing the algorithm to prioritize sequences which are constrained to the given number of energy levels. Though in principle it is still possible for destructive interference to slightly degrade the gate fidelity in a model with models, this is tightly bound by the overall amount of leakage measured in the non-unitary evolution.

We use a variant of the genetic algorithm used in prior work [3] to find a train of SFQ pulses to perform quantum gates. The parameters of the genetic algorithm is summarized in Table I. The genetic algorithm starts with a population of random SFQ pulse trains, and in each iteration, a number of parent pulse trains from the population are selected for generating new pulse trains based on a crossover function. Finally, if the fidelity is improved in the new SFQ pulse trains, they are replaced with the worst SFQ pulse trains in the population.

We use a variant of the gradient ascent pulse engineering (Grape) code used in [11] to find microwave pulses to perform quantum gates. We use the cost functions presented in [11] in order to suppress the occupation of forbidden states. Similar to the SFQ case, we set the target gate fidelity to 0.999.

B. Entangling SFQ-based two-qubit gates on transmon qubit devices

In this section, we present the results of our analysis on transmon qubit devices. Similar to [11], we use qubit frequencies of $\omega_{01}^{(0)}/2\pi = 3.9$ and $\omega_{01}^{(1)}/2\pi = 3.5$ GHz, anharmonicity of $\alpha/2\pi = -225$ MHz, and $n = 5$ in our study on transmons. We report the results for coupling strength of $J/2\pi = 0.05$ in our main results and then perform a sensitivity analysis on the coupling strength. Note that we show the results for transmon with Ω_x control fields and trasmon with Ω_z control fields separately in order to study the effectiveness of each control field on realizing entangling two-qubit gates.

1) CZ gate on transmon qubits with Ω_x control fields: As discussed in Sec. II, we need to take into consideration the non-computational qubit subspace in our optimal control method in order to reduce the leakage to higher energy levels. In addition, we also need to carefully design the optimal control method and qubit configuration in order to optimize the quantum system for SFQ gates. Fig. 3 shows the error and leakage of the best SFQ pulse train found using genetic algorithm to perform a CZ gate. The leakage to higher energy levels is suppressed by the physical model employed in our optimal control method (as described in Sec. III-A). The error numbers reported in this plot take leakage to higher energy levels into consideration, thus, low error translates into low leakage as shown in Fig. 3. We run the genetic algorithm with the two fidelity functions described in Sec. II-C (denoting *subspace-averaged gate fidelity* as *fid1* and *Z-independent gate fidelity* as *fid2*), two tip angles of 0.003 and 0.03 (similar to the numbers reported in the literature [3], [13]), and three gate times of 10 ns, 20 ns, and 40 ns.

Fig. 3(a) and 3(b) show the error and leakage results of transmon devices with Ω_x control fields, respectively. Our results show that it is hard to find high-fidelity CZ gates while suppressing the leakage to higher energy levels using the 0.03 tip angle with either fidelity function. By decreasing the tip angle to 0.003, we are able to realize high-fidelity gates with 40 ns gate time. Decreasing the tip angle means the amount of energy deposited into the qubit with each SFQ pulse decreases, thus, the required gate time to perform high-fidelity quantum operations increases. Although we are not able to realize fast

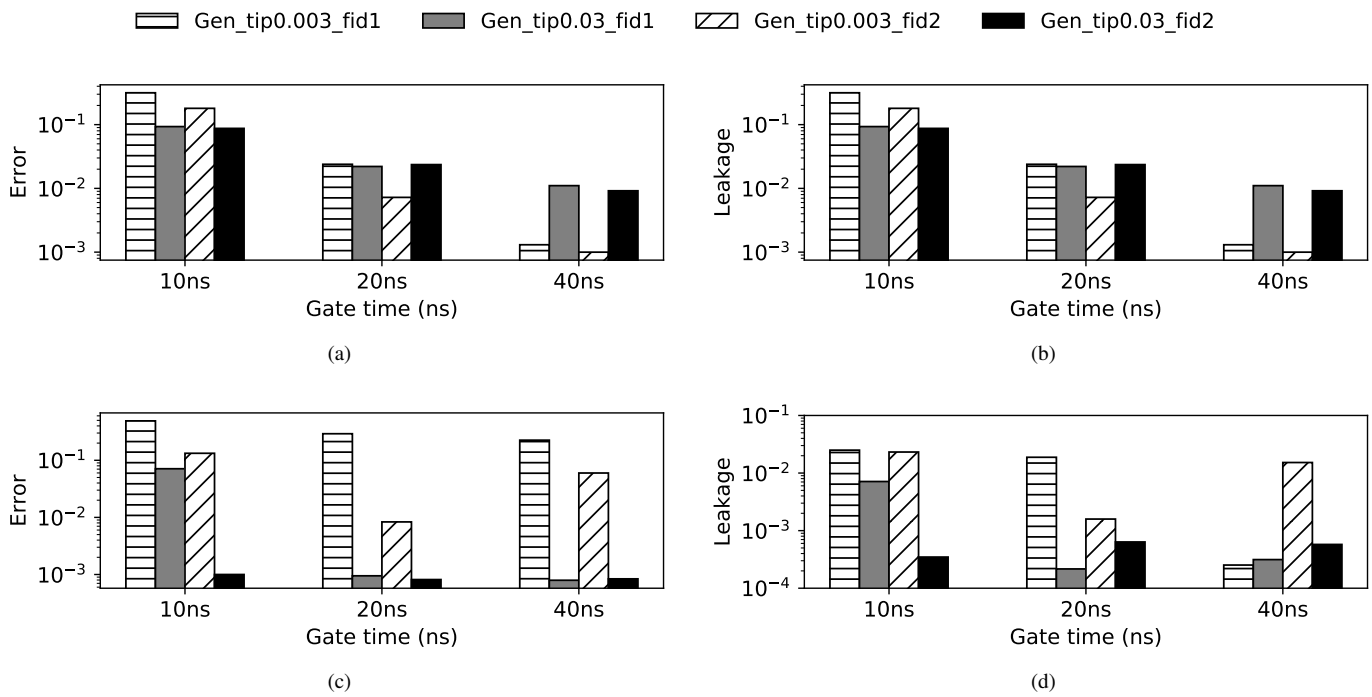


Fig. 3: Error and leakage of the best SFQ-based CZ gate found by the genetic algorithm for transmon qubit devices with Ω_x control fields (plots a and b), and transmon qubit devices with Ω_z control fields (plots c and d). Error is calculated as $1 - fidelity$, and leakage is calculated as the population of non-computational energy levels (averaged over different input states) at the end of the gate. Two different tip angles and two fidelity functions are used in our optimal control method (see Sec. II for the details of our fidelity functions). We run the simulations with $n = 5$ energy levels, and population of the higher energy levels is suppressed in our optimal control method.

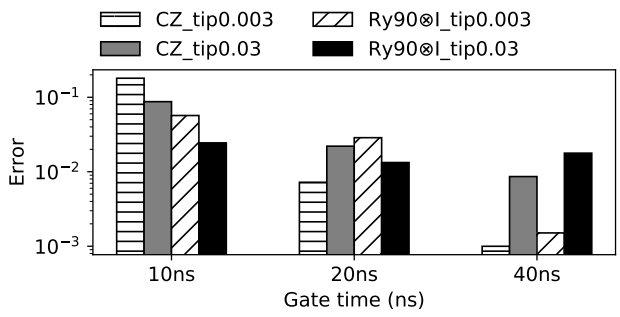


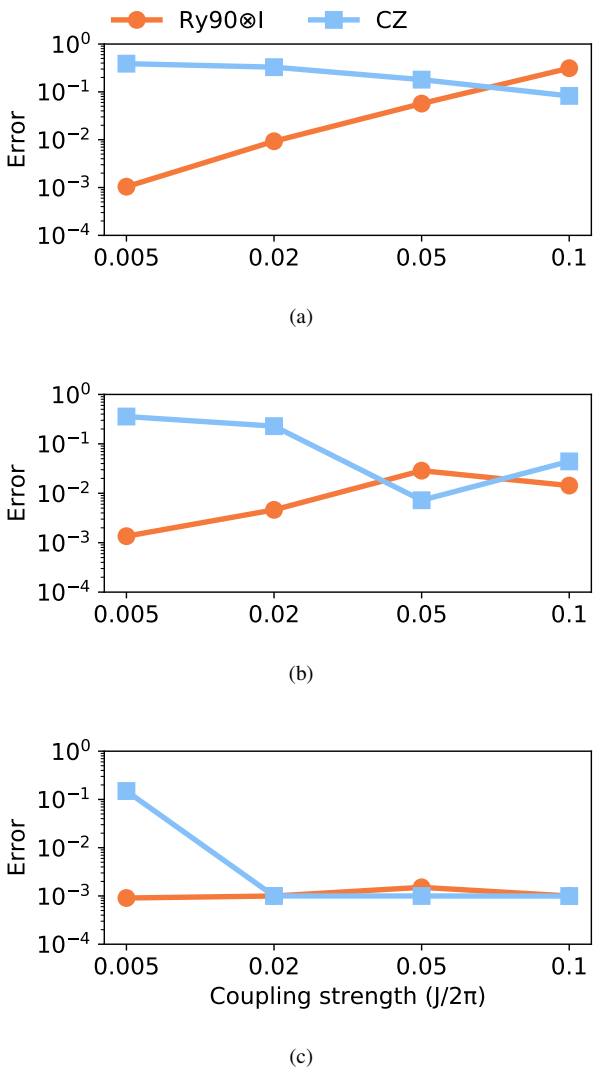
Fig. 4: Error of the best SFQ-based CZ gate (entangling two-qubit gate) and Ry90⊗I gate (non-entangling two-qubit gate) found by the genetic algorithm for transmon qubit devices with only Ω_x control fields.

gates with a low tip angle, we can perform gates with low leakage (with higher gate times) which is desirable.

In general *fid2* results in better SFQ-based pulse trains than *fid1*, indicating that the broader target provided by *fid2* is indeed more friendly to the highly-constrained nature of SFQ-based gate implementation.

2) CZ gate on transmon qubits with Ω_z control fields: Fig. 3(c) and 3(d) show the error and leakage results of transmon devices with Ω_z control fields, respectively. Here, we apply a Ω_z control field only to qubit2 (which is sufficient to realize high-fidelity CZ gates). We observe a significant reduction in the amount of leakage to higher energy levels in the case of transmon devices with Ω_z control fields compared to that of transmon devices with Ω_x control fields. Since the leakage is low in this case, we can afford to use higher tip angles in order to perform fast gates. Our results show that we can realize high-fidelity CZ gates with $<$0.001 error and $<$0.001 leakage with 0.03 tip angle. We need to use *fid2* in the 10 ns gate case, however, if we increase the gate time, we can realize high-fidelity CZ gates with both *fid1* and *fid2*.

Our findings show that Ω_z control field with 0.003 tip angle not sufficient to realize high-fidelity CZ gates. Although we are not able to realize gates with low error with 0.003 tip angle, the gates that we find have low leakage in some cases (which is still not desirable as we care about both error and leakage). Note that although low error translates to low leakage because we take into consideration the leakage to higher energy levels in calculating the error values, the opposite is not necessarily true (for example, identity gate has high error if we calculate its overlap with CZ gate, but it has low leakage to higher energy levels).



(a)

(b)

(c)

Fig. 5: Sensitivity analysis on qubit coupling strength in transmon system with Ω_x control fields. The results are shown for 0.003 tip angle and 10 ns (plot a), 20 ns (plot b), and 40 ns (plot c) gate times.

C. Realizing both entangling and non-entangling SFQ-based two-qubit gates on transmon devices

So far, we demonstrated that we can realize high-fidelity CZ gates with low leakage and short gate time using transmon qubit devices with Ω_z control fields, which is a promising result. However, it is essential to ensure that we can also realize high-fidelity one-qubit gates in our two-qubit quantum system (i.e., non-entangling two-qubit gates). Next, we study the requirements of a system that can perform both entangling and non-entangling SFQ-based two-qubit gates. Transmon system with only Ω_z control fields is suitable to realize high-fidelity entangling two-qubit gates, but it does not provide enough control to perform arbitrary single-qubit gates, which as expected leads to very low fidelity non-entangling two-qubit gates ($> 10^{-1}$ error). Thus, we need more than just Ω_z control fields to realize both entangling and non-entangling two-qubit gates. Next, we investigate two transmon systems as possible candidates.

1) Transmon with Ω_x control fields: Prior work demonstrated SFQ-based single-qubit gates with <20 ns gate time on transmon devices with only Ω_x control fields [13], [14]. In addition, we showed earlier that we can use transmon devices with Ω_x control fields to perform high-fidelity CZ gates with 40 ns gate time and low tip angle. A natural question arises: can we engineer a transmon system with Ω_x control fields that can perform both entangling and non-entangling SFQ-based two-qubit gates with high fidelity? Fig. 4 shows the error results of the best SFQ-based CZ gate (entangling two-qubit gate) and Ry90⊗I gate (non-entangling two-qubit gate) found on transmon system with Ω_x control fields using genetic algorithm. Our results show that we can realize both CZ and Ry90⊗I gates with high fidelity with 0.003 tip angle and 40 ns gate time.

One interesting observation is that the gate time of Ry90⊗I gate is longer than the gate time of the SFQ-based single qubit gates reported in prior work [13], [14]. In general, it is more challenging to realize precise single-qubit gates in a two-qubit system compared to the one-qubit systems because of crosstalk with the neighbor qubit. We can reduce crosstalk and achieve faster single-qubit gate times by reducing the coupling strength (J), however this will in turn complicate the realization of two-qubit entangling gates. Fig. 5 shows a sensitivity analysis on the coupling strength. Our results show that in almost all the configurations, the fidelity of the entangling gates increases by increasing the coupling strength, and the fidelity of non-entangling gates decreases by increasing the coupling strength.

2) Transmon with both Ω_x and Ω_z control fields: One possible configuration is to dedicate both Ω_x and Ω_z control fields to the transmon qubit devices, such that we can realize short CZ gates using the Ω_z control fields. However, we note that this control comes at the cost of hardware complexity and heightened sensitivity to magnetic flux noise.

D. SFQ-based two-qubit gates on fluxonium qubit devices

In this section, we investigate fluxonium qubit devices as a possible candidate to realize both SFQ-based entangling and non-entangling gates with low leakage and short gate time. Our model for the fluxonium devices assumes qubit1 (qubit2) is configured with $E_J = 5.5$ (5.7), $E_C = 1.5$ (1.2) and $E_L = 1.0$, and a static $\varphi_e = \pi$ external flux. Fig. 6 shows the results of our study on fluxonium devices with Ω_x control fields.

Fig. 6(a) and 6(b) show the error and leakage results of an SFQ-based CZ gate, respectively. Our results show that we can realize high-fidelity CZ gates with a gate time of 20 ns thanks to the low leakage of fluxonium devices. Similar to the case of transmon with Ω_x control fields, better results are achieved with lower tip angle. Fig. 6(c) and 6(d) show the error and leakage results of an SFQ-based Ry90⊗I gate, respectively. Our results show that the genetic algorithm can find high-fidelity gates with 20 ns gate time.

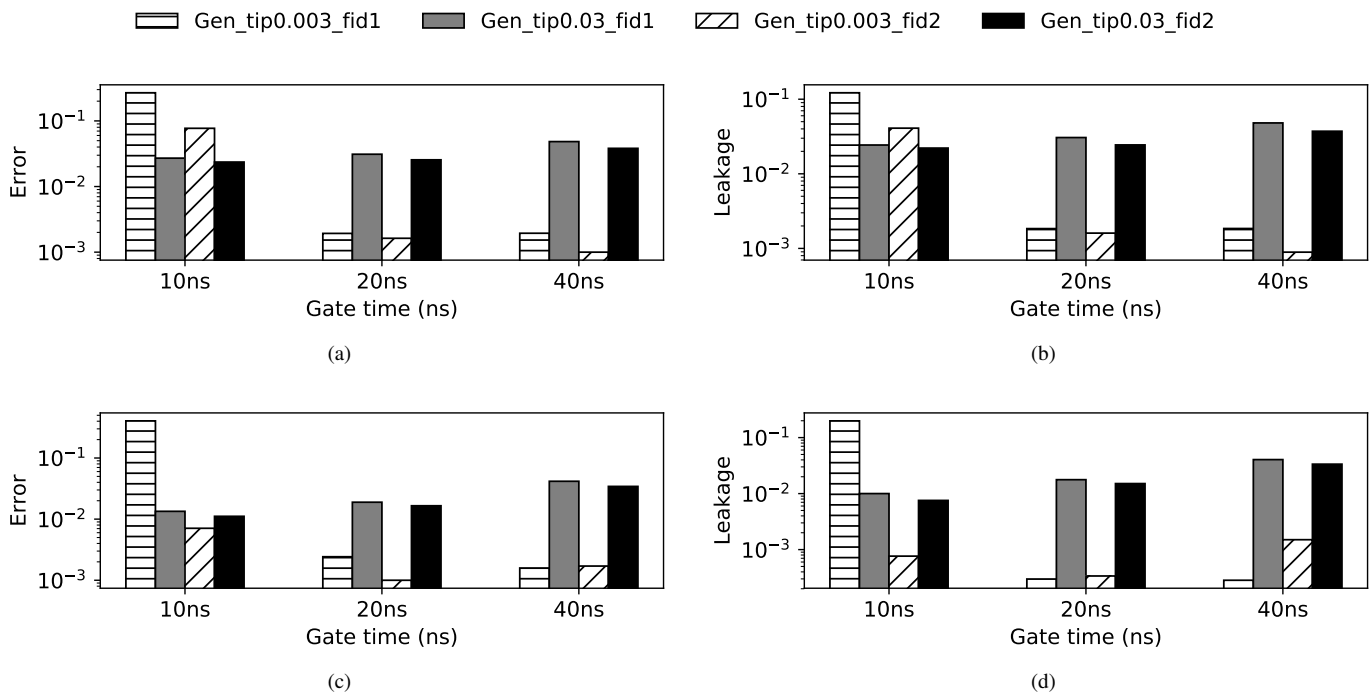


Fig. 6: Error and leakage of the best SFQ-based CZ gate (plots a and b) and Ry90⊗I gate (plots c and d) found by the genetic algorithm for fluxonium qubit devices with Ω_x control fields.

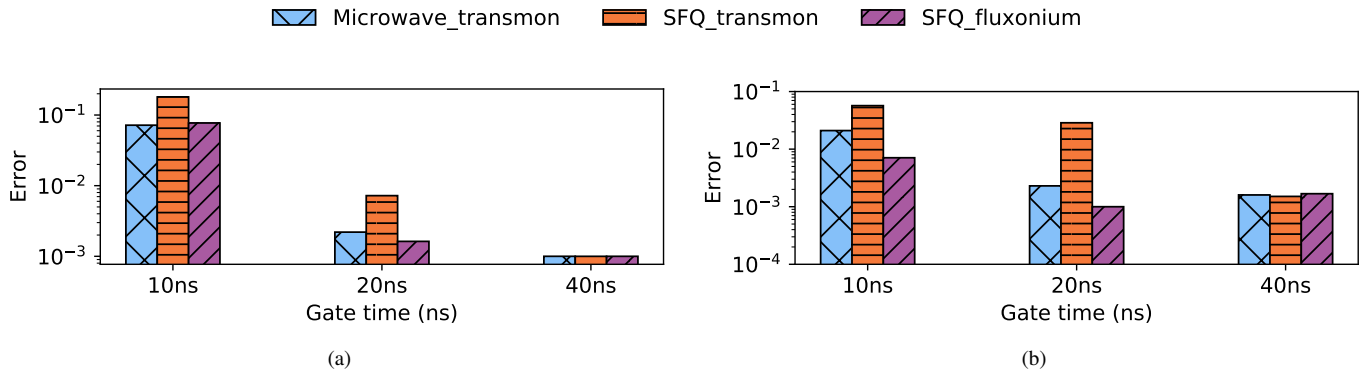


Fig. 7: Error comparison between microwave-based gates obtained using Grape code and SFQ-based gates obtained using genetic algorithm. The results are reported for CZ gate (plot a) and Ry90⊗I gate (plot b).

The fluxonium results show the feasibility and effectiveness of both SFQ-based entangling and non-entangling two-qubit gates with both short gate time and low error and leakage using only Ω_x control fields.

E. Comparison with microwave-based gates

Finally, we compare our results with that of microwave-based gates obtained from the Grape code [11]. Fig. 7 shows the error results for CZ gate and Ry90⊗I gate for three designs: (1) microwave-based design with transmon devices; (2) SFQ-based design with transmon devices; (3) SFQ-based design with fluxonium devices. We learn the SFQ pulses with both *subspace-averaged gate fidelity* and *Z-independent gate fidelity* functions and report the best number. In the microwave case, *subspace-averaged gate fidelity* is sufficient to realize high-fidelity gates.

The results reported in Fig. 7 show that with 10 ns and 40 ns gate times, the three designs have similar performance in both entangling and non-entangling gates (all the design can realize high-fidelity gates with 40 ns gate time). However. with the 20 ns gate time, we observe that high fidelity results are achieved only for microwave-based design and SFQ-based design with fluxonium.

The comparison results show that we can perform high-fidelity SFQ-based gates with similar gate time and gate fidelity to that of microwave-based system. Thus, SFQ is a

promising approach to implement classical controllers as they can deliver quantum computers with both high scalability and high fidelity.

IV. Conclusion

Superconducting Single Flux Quantum (SFQ) is a classical logic technology which is proposed in the literature to implement in-fridge classical controllers in order to maximize the scalability of quantum computers. In this paper, we demonstrate the first thorough analysis of SFQ-based two-qubit gates – a key remaining step in realizing SFQ-based universal quantum computing. Our results show despite the severe challenges of realizing SFQ-based two-qubit gates, they are both feasible and effective if we carefully design our quantum optimal control method and qubit architecture. We characterize the requirements of such gates, and carefully engineer SFQ-friendly quantum systems that can perform both high-fidelity two-qubit gates and single-qubit gates. More importantly, we demonstrate that the fidelity and gate time of these gates are comparable to that of microwave-based gate – these results show that SFQ approach can potentially not only increase the scalability of quantum machines but also maintain the fidelity and effectiveness of quantum gates, thus SFQ is a promising approach to implement classical controllers for quantum machines.

References

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Liakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," *Nature (London)*, vol. 574, no. 7779, 10 2019.

[2] M. Brink, J. M. Chow, J. Hertzberg, E. Magesan, and S. Rosenblatt, "Device challenges for near term superconducting quantum processors: frequency collisions," in *2018 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2018, pp. 6–1.

[3] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, "Global optimization of quantum dynamics with alphazero deep exploration," *npj Quantum Information*, vol. 6, no. 1, p. 6, Jan 2020. [Online]. Available: https://doi.org/10.1038/s41534-019-0241-0

[4] D. P. DiVincenzo, "The Physical Implementation of Quantum Computation," *Fortschritte der Physik*, vol. 48, no. 9-11, pp. 771–783, Sep. 2000. [Online]. Available: http://arxiv.org/abs/quant-ph/0002077

[5] X. Fu, M. A. Rol, C. C. Bultink, J. van Someren, N. Khammassi, I. Ashraf, R. F. L. Vermeulen, J. C. de Sterke, W. J. Vlothuizen, R. N. Schouten, C. G. Almudever, L. DiCarlo, and K. Bertels, "An experimental microarchitecture for a superconducting quantum processor," in *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*, ser. MICRO-50 '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 813–825. [Online]. Available: https://doi.org/10.1145/3123939.3123952

[6] J. Ghosh, "A note on the measures of process fidelity for non-unitary quantum operations," *arXiv e-prints*, p. arXiv:1111.2478, Nov. 2011.

[7] J. Kelly, "A preview of bristlecone, google's new quantum processor," *Google Research Blog*, vol. 5, 2018.

[8] 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*, vol. 76, p. 042319, Oct 2007. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.76.042319

[9] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," *Applied Physics Reviews*, vol. 6, no. 2, p. 021318, 2019.

[10] E. Leonard, M. A. Beck, J. Nelson, B. Christensen, T. Thorbeck, C. Howington, A. Opremcak, I. Pechenezhskiy, K. Dodge, N. Dupuis, M. Hutchings, J. Ku, F. Schlenker, J. Suttle, C. Wilen, S. Zhu, M. Vavilov, B. Plourde, and R. McDermott, "Digital coherent control of a superconducting qubit," *Phys. Rev. Applied*, vol. 11, p. 014009, Jan 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.11.014009

[11] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, "Speedup for quantum optimal control from automatic differentiation based on graphics processing units," *Physical Review A*, vol. 95, no. 4, p. 042318, 2017.

[12] G. Li, Y. Ding, and Y. Xie, "Towards efficient superconducting quantum processor architecture design," in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2020, pp. 1031–1045.

[13] K. Li, R. McDermott, and M. G. Vavilov, "Scalable hardware-efficient qubit control with single flux quantum pulse sequences," *arXiv preprint arXiv:1902.02911*, 2019.

[14] P. J. Liebermann and F. K. Wilhelm, "Optimal qubit control using single-flux quantum pulses," *Physical Review Applied*, vol. 6, no. 2, p. 024022, 2016.

[15] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, "Fluxonium: Single cooper-pair circuit free of charge offsets," *Science*, vol. 326, no. 5949, pp. 113–116, 2009. [Online]. Available: https://science.sciencemag.org/content/326/5949/113

[16] R. McDermott, M. Vavilov, B. Plourde, F. Wilhelm, P. Liebermann, O. Mukhanov, and T. Ohki, "Quantum–classical interface based on single flux quantum digital logic," *Quantum science and technology*, vol. 3, no. 2, p. 024004, 2018.

[17] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, "Simple pulses for elimination of leakage in weakly nonlinear qubits," *Phys. Rev. Lett.*, vol. 103, p. 110501, Sep 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.103.110501

[18] G. S. Paraoanu, "Microwave-induced coupling of superconducting qubits," *Phys. Rev. B*, vol. 74, p. 140504, Oct 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.74.140504

[19] C. Rigetti and M. Devoret, "Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies," *Phys. Rev. B*, vol. 81, p. 134507, Apr 2010. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.81.134507

[20] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Suppressing charge noise decoherence in superconducting charge qubits," *Phys. Rev. B*, vol. 77, p. 180502, May 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.77.180502

[21] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, "Mastering the game of go without human knowledge," *nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[22] M. Steffen, D. P. DiVincenzo, J. M. Chow, T. N. Theis, and M. B. Ketchen, "Quantum computing: An ibm perspective," *IBM Journal of Research and Development*, vol. 55, no. 5, pp. 13–1, 2011.

[23] J. P. G. Van Dijk, B. Patra, S. Subramanian, X. Xue, N. Samkharadze, A. Corna, C. Jeon, F. Sheikh, E. Juarez-Hernandez, B. P. Esparza, H. Rampurawala, B. R. Carlton, S. Ravikumar, C. Nieva, S. Kim, H. J. Lee, A. Sammak, G. Scappucci, M. Veldhorst, L. M. K. Vandersypen, E. Charbon, S. Pellerano, M. Babaie, and F. Sebastiano, "A scalable cryo-cmos controller for the wideband frequency-multiplexed control of spin qubits and transmons," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 11, pp. 2930–2946, 2020.