







Solovay-Kitaev algorithm and randomized compilation

Oliver Maupin ¹, Ashlyn D. Burch ^{2,*}, Christopher G. Yale ², Matthew N. H. Chow ^{3,2,†}, Terra Colvin, Jr. ⁴,
Brandon Ruzic,² Melissa C. Reville,² Brian K. McFarland,² Eduardo Ibarra-García-Padilla ², Alejandro Rascon,^{2,3}
Andrew J. Landahl,^{2,3} Susan M. Clark,² and Peter J. Love^{4,5}

¹*Quantum Performance Laboratory, Sandia National Laboratories, Albuquerque, New Mexico 87185 and Livermore, California 94550, USA*

²*Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

³*Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131, USA*

⁴*Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA*

⁵*Computational Science Initiative, Brookhaven National Laboratory, Upton, New York 11973, USA*



(Received 4 April 2025; accepted 13 April 2026; published 4 June 2026)

We analyze the use of the Solovay-Kitaev (SK) algorithm to generate an ensemble of one-qubit rotations over which to perform randomized compilation. We perform simulations to compare the trace distance between the quantum state resulting from an ideal one-qubit R_z rotation and discrete SK decompositions. We find that this simple randomized gate synthesis algorithm can reduce the approximation error of these rotations in the absence of gate errors in simulation by at least a factor of 2 compared to a naive gate synthesis algorithm. We test the technique under the effects of a simple coherent noise model and find that it can mitigate coherent noise. We also run our algorithm on Sandia National Laboratories' QSCOUT trapped-ion device and find that randomization is able to help in the presence of realistic noise sources.

DOI: [10.1103/PhysRevA.113.062418](https://doi.org/10.1103/PhysRevA.113.062418)

I. INTRODUCTION

Quantum computing is currently in the noisy intermediate-scale quantum (NISQ) era of development [1]. Noisy intermediate-scale quantum computers exhibit quantum coherence, scale to hundreds of qubits, achieve error rates of less than 1%, and have enabled many small-scale demonstrations of quantum algorithms, as well as larger-scale experiments that push towards utility [2]. Noisy intermediate-scale quantum machines lack quantum error correction; however a wide range of error mitigation techniques have been developed which aim to extract better results from noisy quantum data [3].

To move beyond the NISQ era requires implementation of quantum error correction (QEC) [4]. Quantum error correction encodes each logical qubit in many physical qubits. If the error rate on the physical qubits is below a threshold (whose value depends on the particular scheme used), then the error rates for all encoded gates on the logical qubits can be lower than those on the physical qubits. This enables continued reduction in error rates at the price of more physical qubits, but without further improvements in physical qubit noise. These techniques are necessary to reach the very large numbers of qubits and gate operations that are required by applications at the quantum advantage scale.

With recent demonstrations of logical qubits in several platforms we are seeing the first demonstrations of post-NISQ quantum technology [5–9]. This new era of development has been termed early fault-tolerant quantum computation (EFTQC) [10]. Like NISQ experiments, these EFTQC demonstrations will inform the development of future large-scale quantum computation. These advancements raise the following question: Can NISQ experiments inform the development of EFTQC devices? In the present paper we address one aspect of this question by considering a key element of QEC, the decomposition of continuous operations into circuits over a discrete gate.

The Solovay-Kitaev (SK) theorem [11–15] shows that it is possible to find such decompositions of continuous rotations into a discrete gate set as is required by QEC. This process is called gate synthesis and is the focus of ongoing research [16–18]. This compilation takes place in the quantum computing stack between the high-level circuit as outlined by a quantum algorithm and the logical program expressed in logical operations, as shown in Figs. 1(a) and 1(b). However, this decomposition can be costly, requiring extra circuit depth compared to physical rotations on NISQ devices. When then could an EFTQC outperform arbitrary rotations on a NISQ device?

Current experimental demonstrations of quantum error correction focus on implementing universal and reliable logical operations that outperform physical operations [5,6,8,9]. For these demonstrations to outperform the high fidelities of single-qubit gates on NISQ devices, it is important to reduce the cost of compilation, as the required increase in circuit depth penalizes the performance of EFTQC devices [5]. In this work we propose a method for decreasing the cost of

*Present address: Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA.

†Present address: HRL Laboratories, Malibu, CA 90265-4797, USA.

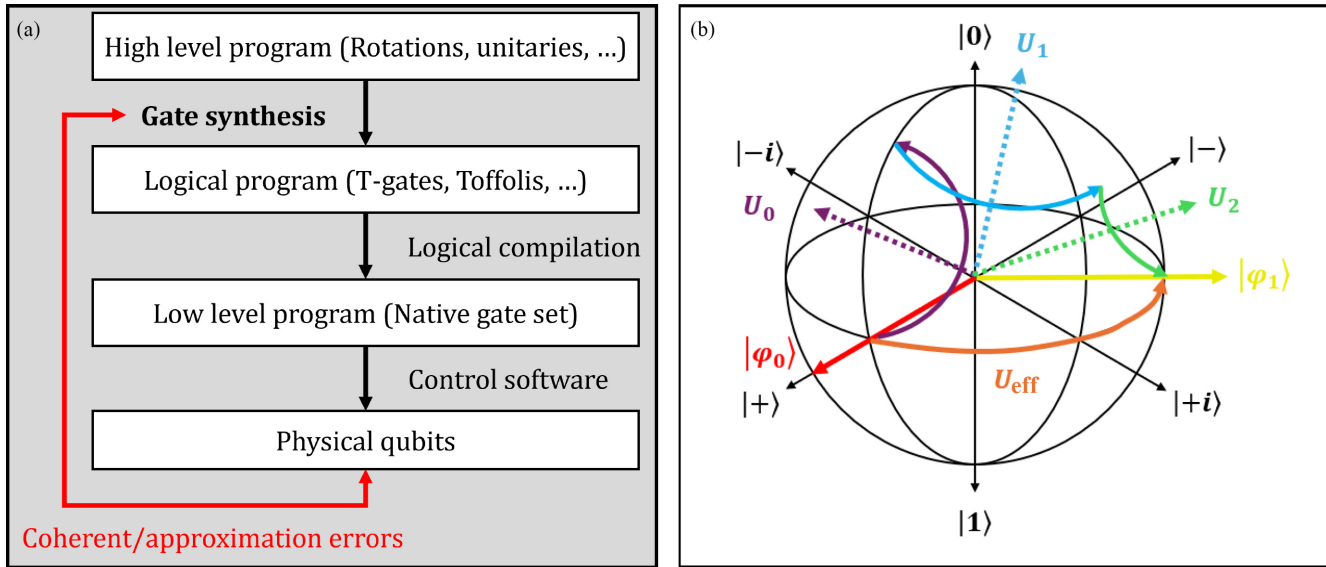


FIG. 1. Overview of gate synthesis. (a) Gate synthesis takes arbitrary operations from a high-level program and compiles them into discrete logical operations acting on logical qubits. This work focuses on the effects of coherent errors and approximation errors on gate synthesis and vice versa. (b) Single-qubit gate synthesis decomposes an arbitrary rotation U_{eff} into a series of discrete operations $U_0, U_1, U_2, \dots \in \mathcal{G}$ from a gate set \mathcal{G} .

single-qubit gate synthesis in the presence of coherent errors and benchmark its effectiveness using a trapped-ion device as a baseline of performance.

Coherent noise is believed to be corrected or at least mitigated by error-correction codes intended for use with Solovay-Kitaev decompositions [19,20]. However, the use of error-correcting codes does not rule out coherent errors that can occur in early fault-tolerant devices that merely suppress the error rather than eliminating it entirely. This is the case for EFTQC architectures with smaller code sizes [20] or architectures wherein cycles of syndrome extraction are costly and may only be applied sparingly throughout the duration of circuits [21]. In this context, coherent errors at the physical level can manifest as logical errors, impacting performance.

Our method aims to reduce the effects of coherent and approximation errors by using the inherent compilation degeneracy in existing gate synthesis algorithms. Some gate synthesis algorithms attempt to calculate decompositions according to a cost metric, but do not try to perform exact synthesis to find the most accurate decomposition [14,17,22,23]. More specifically, the user chooses some allowed precision ϵ , and the algorithm finds a decomposition that approximates the desired rotation within that precision, according to some distance measure. These algorithms usually minimize the required number of T gates. This allowed precision ϵ creates a degeneracy in the choice of decomposition, as there are many sequences R_i that achieve the same precision. Imposing a limit on the number of gates or the number of T gates reduces the size of this ensemble of approximate rotations but still allows for multiple different approximations of the same rotation of the same quality.

This compilation degeneracy allows for the use of NISQ era twirling techniques such as randomized compilation (RC), a technique that has been used [24–26] to tailor coherent quantum errors into stochastic errors. The idea of randomized

compilation is simple. A logical quantum circuit is duplicated r times, each copy being decomposed into a different random set of operations at the gate level. Given a total sampling budget N , each of these circuits is then run on the device N/r times, and the resulting measurement outcomes are averaged together. This procedure maintains the global circuit structure but randomizes the gate sequence to mitigate the effects of coherent errors, namely, RC via Pauli twirling can tailor coherent error into a Pauli channel, and RC via Clifford twirling can further tailor that error into a purely depolarizing channel.

In much the same way, we can probabilistically sample from an ensemble of approximate rotations generated for each of k arbitrary rotations in our circuit. Choosing one sampled sequence for each rotation produces a single sampled circuit, and enough random instantiations of these sampled circuits will randomize coherent errors when their measurement results are averaged together.

Previous works in gate synthesis [17,27,28] focus on minimizing approximation error incurred during gate synthesis using probabilistic methods. Probabilistic state synthesis uses a weighted mixture of an ensemble of circuits to generate an approximation to a desired quantum state. The mixture over the ensemble results in better average accuracy than any individual circuit. Probabilistic state synthesis is therefore the state equivalent of RC [29], except the error to be minimized includes the synthesis error and not only the hardware error. This is the same notion that underlies multiproduct formulas [30] in Hamiltonian simulation.

In the present paper we explore the effects of coherent and projection noise in the early fault-tolerant regime in addition to synthesis error when using SK sequences, comparing the accuracy and cost of gate synthesis techniques using Sandia National Laboratories’ QSCOUT trapped-ion device. The physical single-qubit gates of trapped-ion systems are significantly better than the two-qubit gates, especially

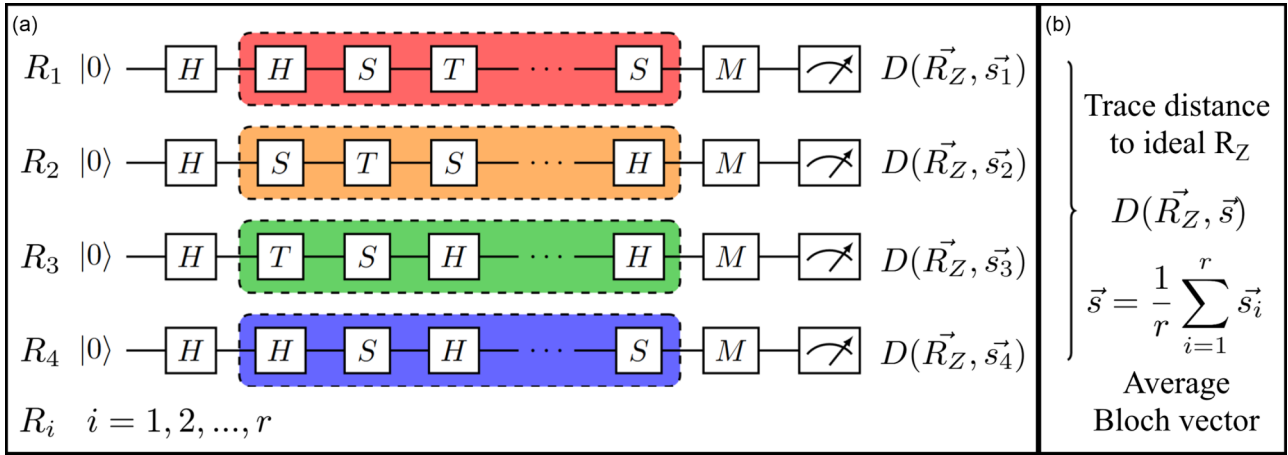


FIG. 2. Schematic of our randomization protocol and analysis. (a) Individual rotation sequences R_i are generated using gridsynth with slight variations in the initial parameters. The accuracy of each sequence can be evaluated by taking the trace distance of the corresponding Bloch vectors \vec{s}_i to the exact R_Z rotation as measured with state tomography by rotating the state before measurement with Clifford operations M . (b) To implement randomization, the measurement results for these sequences are averaged as in randomized compiling to achieve an effective rotation and Bloch vector \vec{s} . The accuracy of this effective rotation can then be evaluated via trace distance, which is less than the average trace distance for an individual sequence.

for trapped-ion architectures which display extremely high single-qubit gate fidelities [31–33]. This makes them a good baseline for comparison with the low error rates of EFTQC systems, as they are able to achieve long sequences of single-qubit gates with high fidelity. This foreshadows future benchmarking efforts on EFTQC devices.

This paper is structured as follows. We begin in Sec. II by detailing the algorithm that we use to generate SK sequences and analyzing the resulting ensemble of rotations. We then discuss the coherent noise model we use in Sec. III. In Sec. IV we show results from a noiseless simulation of the technique, simulation under our simple coherent noise model, and experiment on a NISQ trapped-ion quantum computer. We summarize our findings in Sec. V.

II. SEQUENCE ENSEMBLE

Before we present our results, we first discuss our method for generating distinct SK sequences. We then analyze the characteristics of the resulting ensemble and present our randomization algorithm. The procedure is shown graphically in Fig. 2.

A. Generating sequences

We consider decompositions into the set of $\{H, S, T\}$ gates, as defined by

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}. \quad (1)$$

These gates are sufficient to approximate any continuous one-qubit rotation about the X, Y , or Z axis of the Bloch sphere and are a common choice of gate set. We will refer to a particular decomposition of a rotation as an SK sequence,

$$R = U_k U_{k-1} \cdots U_1 U_0, \quad U_j \in \{H, S, T\}, \quad (2)$$

where the length of the sequence k can vary depending on the rotation R_Z and approximation error. These sequences or decompositions can approximate any rotation $R(\hat{n}, \theta)$ about the axis \hat{n} by an angle θ . It is sufficient to synthesize continuous rotations about one axis, because rotations about any axis can be synthesized from the gate set $\{H, S, T\}$ and continuous rotations around one axis. We focus on approximating the $R_Z(\theta)$ rotation.

The algorithm we use to generate our SK sequences is known as gridsynth [23]. It decomposes $R_Z(\theta)$ rotations into the $\{H, S, T\}$ gate set, minimizing the number of T gates. Nontransversal gates are the most challenging to implement in QEC, and T gates cannot be implemented transversally for many error-correcting codes. This includes codes such as the surface code [34,35], the Steane code [36], the Shor code [37], or two-dimensional (2D) color codes [38], and known methods for implementing them nontransversally such as magic state distillation [39–41] or cultivation [42] are costly. The gridsynth algorithm is able to obtain solutions with T counts $t + O(\log[\log(1/\epsilon)])$, where t is the T count of the second most optimal solution. We refer the reader to [23] for more details, but we do note a few aspects of interest.

The gridsynth algorithm takes as input a random seed, the desired rotation angle θ , the binary precision $\epsilon = 2^{-b}$ in realizing that rotation, and an “effort” integer to determine how much time the algorithm should spend finding an optimal solution. We may vary these parameters in order to generate multiple unique sequences for the same rotation. By varying the desired rotation angle θ by a value smaller than ϵ , we can find multiple different sequences R_i , where $i = 1, 2, \dots, r$, that satisfy our precision constraint. These sequences may have slight variations in T counts but are close enough to be grouped together in an equivalent ensemble. The gate count and T count for these sequences increase linearly as a function of binary precision b , as can be seen in Fig. 3.

We note that in order for a precision of $\epsilon = 2^{-b}$ to affect the results of an experiment implementing a single rotation,

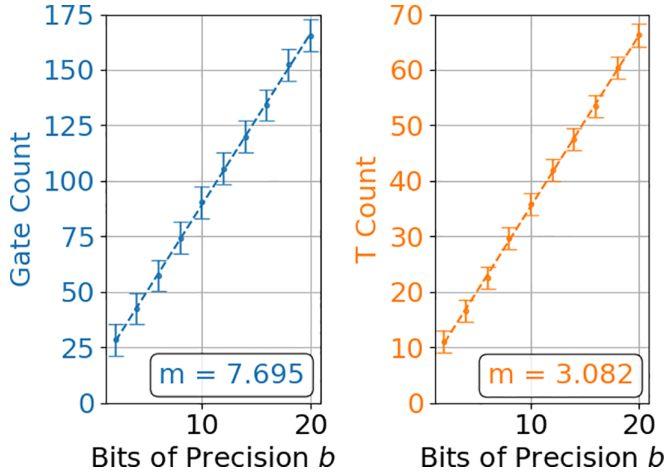


FIG. 3. Average gate count (left) and average T -gate count (right) for sequences generated by gridsynth as a function of the bits of precision b , where $\epsilon = 2^{-b}$. Error bars are the standard deviation over an ensemble of 100 sequences. Fitted slopes are 7.695 for the gate count and 3.082 for the T -gate count.

it is necessary to also have a sufficient sampling budget. This requires a number of samples

$$n \sim O(\epsilon^{-2}) = 4^b, \quad (3)$$

which grows very quickly with the desired precision. This shows that for any fixed number of samples there is an implied maximum value b_{\max} of b beyond which further increases in sequence precision are not distinguishable from sequences obtained with precision b_{\max} . Further discussion of sampling noise can be found at the end of Sec. IV A.

B. SKARC algorithm

We now outline our randomization procedure Solovay-Kitaev and randomized compilation (SKARC) for a given quantum circuit C containing different R_Z gates and a total sampling budget N .

For each instance of an R_Z gate we wish to compile in C , we generate an ensemble of unique gate sequences R_i , $i = 1, 2, \dots, r$, using gridsynth as described in Sec. II. These sequences can then be randomly selected with or without replacement in place of each corresponding R_Z rotation in C . It is possible to randomize over all possible combinations of compiled R_Z gates, but usually unnecessary to see the benefits of randomization, so oftentimes the total number of randomized circuits can be truncated, as will be explored in this work in Sec. IV A. We focus on a circuit containing a single R_Z gate, but our results are easily extended to circuits with multiple rotations. Once all circuits have been compiled this way, each randomly compiled circuit is measured N/r times, dividing the total sampling budget between the compilations. The resulting N measurement outcomes are then combined into one distribution for analysis, effectively averaging the measurement results over r different decompositions of the R_Z gate.

We study the performance of our procedure via state tomography. We create and execute circuits for each SK sequence R_i and then measure each of these circuits in the X , Y ,

and Z bases. This results in $3r$ different sets of measurement outcomes that can be used to reconstruct r Bloch vectors \vec{s}_i , one for each decomposition:

$$\vec{s}_i = \langle X \rangle \hat{i} + \langle Y \rangle \hat{j} + \langle Z \rangle \hat{k}. \quad (4)$$

An ensemble of m such vectors determines an average vector. This can be done either by combining the measurement results from each sequence together and calculating a new Bloch vector as described above or by simply taking the average of the vectors obtained from each sequence. We define this mean vector for a particular choice of m sequences indexed by $\{i_1, i_2, \dots, i_m\}$ as

$$\vec{s}_{\{i_1, i_2, \dots, i_m\}} = \frac{1}{m} (\vec{s}_{i_1} + \vec{s}_{i_2} + \dots + \vec{s}_{i_m}). \quad (5)$$

In order to quantify the performance of SKARC, we consider the trace distance between the target quantum state and the mean quantum state from averaging. As these decompositions are for single-qubit rotations, the trace distance $D(\vec{r}, \vec{s})$ is equivalent to half the Euclidean distance between the vectors on the Bloch sphere, where

$$D(\vec{r}, \vec{s}) = \frac{|\vec{r} - \vec{s}|}{2}. \quad (6)$$

A smaller trace distance $D(\vec{R}_Z, \vec{s}) < D(\vec{R}_Z, \vec{s}_i)$ indicates that we have better realized the target rotation. Here \vec{R}_Z is the vector corresponding to the target R_Z rotation, \vec{s} is the mean vector for the entire ensemble, and \vec{s}_i is the vector for a single sequence chosen at random or optimized for its cost.

Generating an ensemble of r sequences is the first step in employing averaging over sequences. We will use noiseless simulations to evaluate the effect of randomization on the accuracy of the results, as in [17,27,28]. In simulations with coherent noise added, and in experiment, we are evaluating the effect of randomization as a form of randomized compilation to mitigate coherent noise.

Across the ensemble of r total sequences generated by gridsynth, some of them are going to be more robust to certain types of noise than others. We can evaluate the variation between different ensembles of size $m < r$ by sampling subsets of our overall ensemble of size m . It is intractable to consider all possible combinations for $r = 100$ unique sequences, so instead we randomly choose at most $q = 1000$ different ensembles of size $1 \leq m \leq 100$ for simulation and $1 \leq m \leq 20$ for experiment. We sample these subensembles randomly with replacement and calculate the mean trace distance as

$$D(m) = \frac{1}{q} \sum_{j=0}^q D(\vec{s}_{\{i_1, i_2, \dots, i_m\}_j}, \vec{R}_Z), \quad (7)$$

where \vec{R}_Z is the Bloch vector resulting from a noiseless continuous R_Z rotation.

This average trace distance across 1000 possible combinations of sequences for a given ensemble size is then used to determine the efficacy of SKARC. In Sec. IV we compare the average error with and without randomization as a function of the size of the ensemble that we average over. We should expect that as the size of the ensemble of distinct sequences increases, our error metrics will converge to some finite value. Moreover, by studying the characteristics of the ensemble of rotations implemented by our ensemble of circuits, we can learn more about the limits of SKARC.

III. COHERENT NOISE MODEL

The coherent noise model we use in the remainder of this work is to multiplicatively overrotate each rotation comprising a H gate in our circuit by a fraction $1 + \delta$, for some error strength δ . This gate-dependent coherent noise model is similar to those used in [43,44] and the original RC study [29], though the latter uses an additive overrotation error rather than a multiplicative one. We only model simple single-qubit coherent errors, as this work is focused only on single-qubit rotations. We only include H gates because on the QSCOUT ion trap the S and T gates are implemented with virtual R_Z rotations and incur no error. The H gates are implemented as a product of gates $X\sqrt{Y}$ and each of these gates is overrotated by δ so that $X(\delta) = R_x[\pi(1 + \delta)]$ and $\sqrt{Y}(\delta) = R_y[\frac{\pi}{2}(1 + \delta)]$.

We choose values of δ that vary in the range $[0, 10^{-2}]$. If δ were much larger than this, then it would be very difficult to resolve the desired ϵ , even for a small number of bits of precision. On a real quantum device, the strength of this coherent overrotation will depend on the error rate of the device in question and how much error-correcting codes suppress errors.

The effect of this noise model on an individual SK sequence can be modeled as an erroneous single-qubit rotation, that is, if the noiseless approximate rotation is $R_i(\hat{n}, \theta)$, then the noisy rotation will be $\tilde{R}_i = R_i\mathcal{E}$, where $\mathcal{E} = R(\hat{b}, \phi)$ is a rotation of the state away from the noiseless result about an axis \hat{b} by an angle ϕ . The axis and angle of this error will themselves be functions of the original axis \hat{n} , the original angle θ , and the strength of the noise model δ . However, the functional form of this error in terms of these parameters is nontrivial and sequence dependent. As a result, each sequence will experience this noise model in a unique manner. Averaging over many sequences will result in an ensemble of outputs converting the coherent noise to stochastic noise.

IV. RESULTS

In this section we present the results from two different wavefunction simulations using Sandia National Laboratories' JaqalPaq emulator [45], as well as an experiment on the QSCOUT device at Sandia National Laboratories [31], a trapped-ion NISQ computer. Noiseless simulation results are discussed in Sec. IV A. Results of simulations including the coherent noise model defined in Sec. III are discussed in Sec. IV B. Experimental results are discussed in Sec. IV C. Each of these sections will reference and compare data shown in Figs. 4–6, which we will first describe here.

The simulated data come from wave-function simulations of the circuits that are equivalent to an infinite number of samples. In total, $r = 100$ unique sequences were executed in simulation. The experimental data on the QSCOUT device were obtained from 24 000 samples per measurement basis and executed $r = 20$ sequences. The circuit is comprised of an H gate to prepare the $|+\rangle$ state, followed by an $R_Z(\theta = 1)$ rotation, followed by measurement. We choose the $|+\rangle$ state as our input state, as it is perpendicular to the axis of rotation on the Bloch sphere, and we choose the angle $\theta = 1$ as it is not a clean fraction of π and therefore sufficiently far away from the unitaries generated by the gridsynth algorithm so as not to have a trivial decomposition even for many bits of precision.

To visualize the distribution of Bloch vectors originating from different sequences, we plot them on a 2D projection of the Bloch sphere. Figure 4 shows data for precision $b = 4$. The Bloch vector corresponding to the target rotation $R_Z(\theta = 1)$ is used as the normal vector for the 2D projection and is shown as a black dot in the center of the target. The state vectors for each of the r SK sequences \vec{s}_i are shown as colored crosses, with the mean of these state vectors \vec{s} shown as a circle. Experimental results are shown in red.

The triangle markers highlight the SK sequence with the smallest number of H gates, which is the least affected by noise in our noise model. For all precisions tested except $b = 7$ and 10, the sequence with the fewest H gates is also the sequence with the fewest T gates, which is the normal output of the gridsynth algorithm and the sequence with the least theoretical error for many error-correcting codes. For precisions $b = 7$ and 10, the sequence with the fewest H gates has the second-fewest T gates instead. We note that the Ross-Selinger algorithm does not attempt to minimize the number of H gates used in its compilation, so exact comparisons between experimental NISQ results for the least noisy sequence and simulated EFTQC results for the least noisy sequence are imperfect, even if they are identical among generated sequences.

Figure 5 shows simulated and experimental trace distances as a function of the size of the ensemble of sequences $m \leq r$ that we average over. The combinations of sequences are drawn randomly, with the total number of possible combinations of sequences capped at 1000. For the simulated data, the precision of the sequences ranges from $b = 2$ to $b = 20$ in even increments from top to bottom. For the experimental data, the precision of the sequences is in the range $b = [4, 5, 6, 7]$ from top to bottom for the solid lines. The dashed horizontal lines in the experimental data represent the trace distance for the sequence with the fewest H gates and least expected error, colored in the same order as the mean trace distance data.

Figure 6 compares trace distances to the target rotation with and without SKARC for simulated and experimental data as a function of the bits of precision b . The solid lines with circles correspond to trace distances for mean Bloch vectors over the entire ensemble. The dashed lines correspond to trace distances for the SK sequence with the fewest H gates and thus least error. Here the error bars are calculated as the variance in the each of the X , Y , and Z components of the Bloch vectors across the ensemble of sequences. This error is then propagated through to the trace distance calculation.

A. Noiseless simulation

We first consider the noiseless case of $\delta = 0$. Figure 4(a) shows the distribution of Bloch vectors realized by the ensemble of SK sequences. The mean of the ensemble lies close to the target rotation in the center. This is the expected result, as the approximation error inherent to each individual sequence is randomly distributed and thus will cancel out in aggregate. Moreover, this mean lies closer to the center than the SK sequence output by gridsynth with the least theoretical error highlighted with a triangle. This shows that use of randomization over SK sequences can improve the accuracy of the aggregate result.

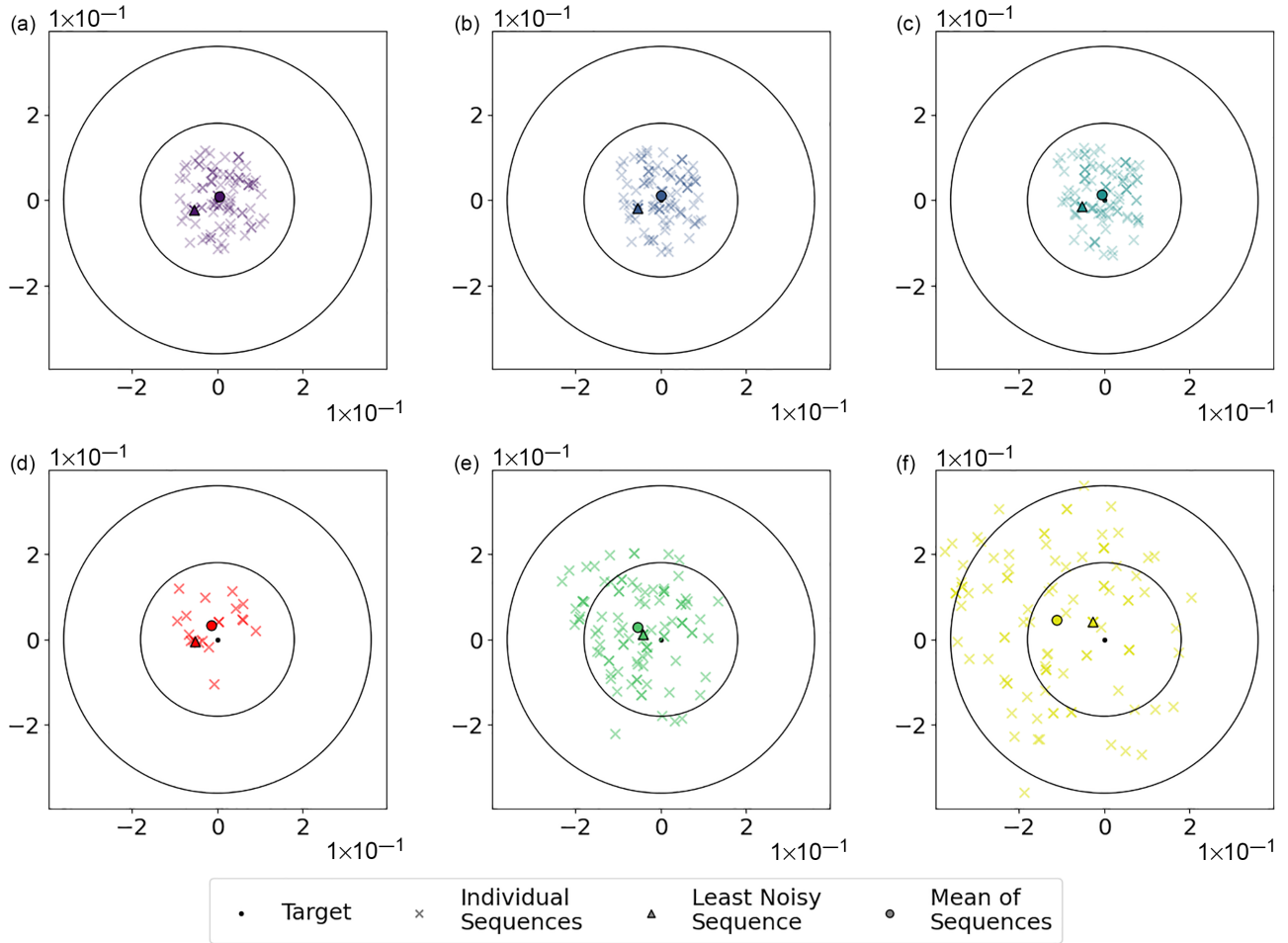


FIG. 4. Comparison of results from noiseless simulation, a coherent noise model, and experiment for precision $b = 4$. Plots are ordered left to right, top to bottom based on trace distance of the mean Bloch vector to the target rotation. (a)–(c), (e), and (f) Two-dimensional projections of Bloch sphere vectors as a function of the strength of the coherent noise model δ . There is a bias in the vectors as a function of the noise, with larger δ shifting left of center due to the cumulative effects of the coherent noise manifesting as a single biasing coherent rotation. A negative δ would instead shift to the right. (d) The projection of experimental data is shown in red, with the mean Bloch vector having a similar trace distance compared to the sequence with the least theoretical error. The trace distance for the mean Bloch vector falls between $\delta = 0.001$ and $\delta = 0.005$. For coherent noise strengths $\delta \leq 0.001$ and experiment, randomization yields a lower trace distance than the sequence with the least theoretical error.

The average trace distance as a function of the ensemble size m for a larger set of noiseless data for $b = [2, 20]$ is shown in Fig. 5(a). We see that the mean trace distance decreases as the ensemble size increases, indicating that randomizing over more SK sequences leads to quantum states that are closer to those of the target R_Z rotation. We note that taking into account the lowest value of the error bars in Fig. 5(a), the smallest possible trace distance would be achieved by averaging over less than the full circuit ensemble, which is expected. There will generally exist some particular combination of sequences that when averaged together will achieve a trace distance that is less than the trace distance for the average of the full ensemble. These sorts of algorithms are the basis of other probabilistic gate synthesis methods [17,28], that is, deliberately picking two sequences that are equally and oppositely spaced from the true rotation will give a very accurate result, but it requires more work to implement this sort of weighting of different unitaries.

The noiseless results in Figs. 6(a) and 6(b) indicate that randomization improves accuracy. The trace distance for the mean Bloch vectors are about $\simeq 0.15$ of the trace distances for the optimal Bloch vectors averaged over varying precisions. This improvement is consistent as the precision is increased and is similar to increasing the bits of precision by 2–3. Using Fig. 3, we see this is equivalent to a reduction of ~ 5 –10 in the required number of T gates needed to achieve the desired precision and a reduction in total gate count of ~ 15 –25.

We also examined the relationship between the number of samples used in simulation and the accuracy of the rotation. For an approximation error of $\epsilon = 2^{-b}$, we would expect to need $N = \epsilon^{-2} = 4^b$ samples to match. Using fewer samples than this means that imprecision due to sampling noise will overshadow any gains made by using more precise SK sequences. Using more samples than this means the opposite, that imprecision due to the approximation error of the sequence will overshadow any gains made by using more samples.

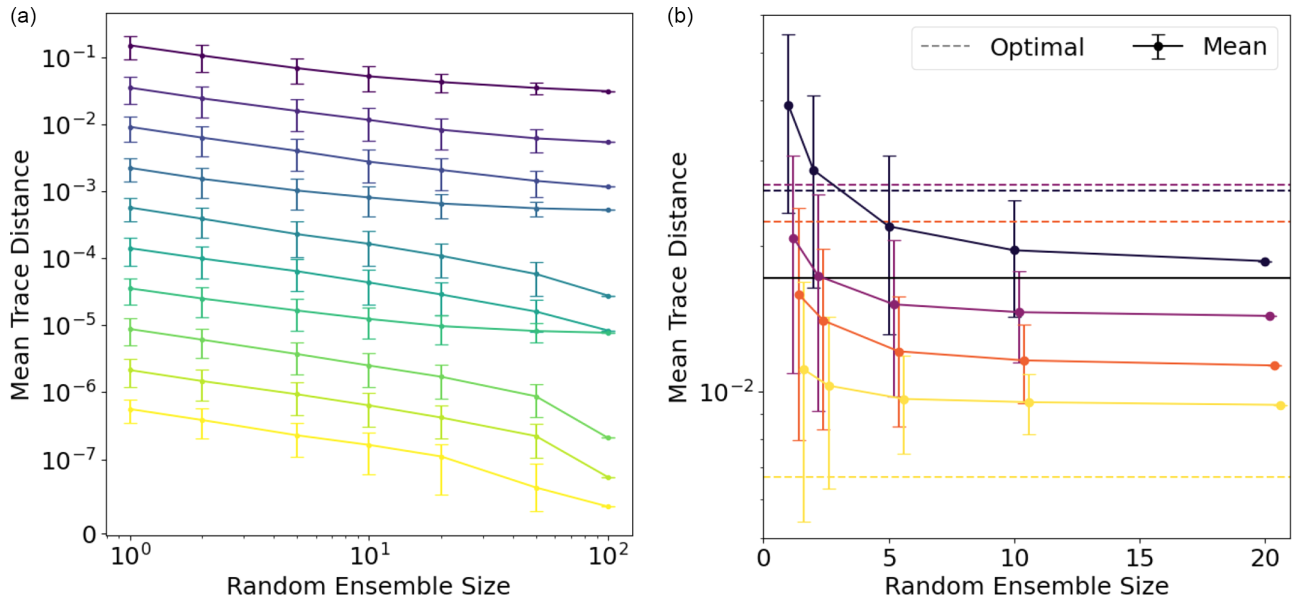


FIG. 5. (a) Simulation results of the mean trace distance across possible combinations of sequences for a given random ensemble size $m \leq 100$. Precision increases from $b = 2$ to $b = 20$ in even increments from top to bottom. Only noiseless simulated data are shown. Accuracy increases with both precision and the size of the ensemble used for averaging. Error bars are taken as the standard deviation of the trace distances for all 1000 randomly chosen combinations of different sequences. (b) Experimental results of the mean trace distance. Precision increases in the range $b = [4, 5, 6, 7]$ from top to bottom. Accuracy again increases with both precision and the size of the ensemble used for averaging. Error bars are calculated as for the simulated data but for $m \leq 20$. Better accuracy can be obtained from choosing particular combinations of sequences as noted by the large error bars for smaller ensemble sizes. Trace distances for sequences with the fewest H gates are shown with dashed lines.

We see the necessity of both a large number of samples and many bits of precision in Figs. 7(a) and 7(b). The red line on these contour plots indicates the $N = 4^b$ number of samples required for the theoretical best precision. We see that this lines up nicely with the shape of the contour itself. Improvements in precision (i.e., lower trace distance) can only be achieved by both increasing the number of samples and the precision of the sequence. This is also true for the randomized case. However, the randomized data generally

show smaller trace distances than the nonrandomized data, as expected. Using randomization, we can achieve better precisions without needing to increase the number of samples.

B. Simulation with coherent noise

The question remains, do the improvements we see in the noiseless case hold in the presence of coherent noise? Single-qubit coherent errors manifest as unitary rotations on

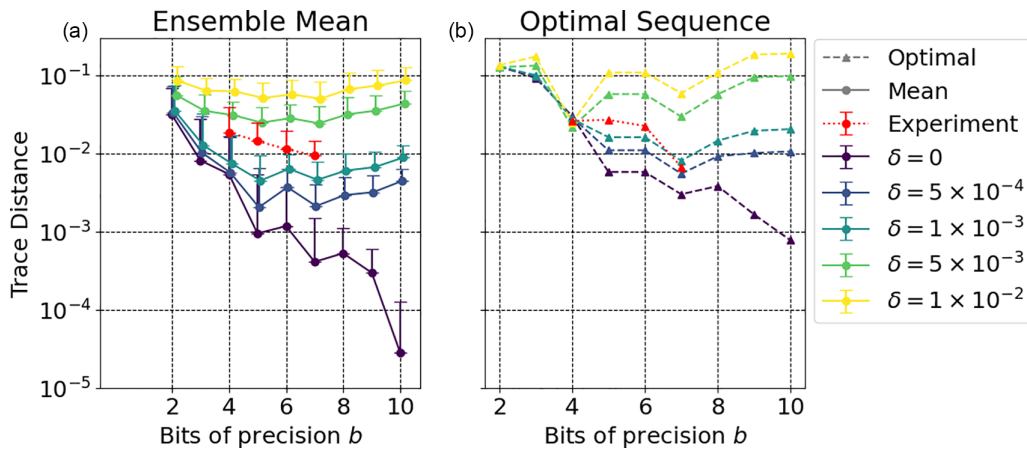


FIG. 6. Trace distance to the target rotation as a function of the bits of precision b . (a) Trace distance for the mean Bloch vector of the full ensemble of random sequences shown with solid lines and circles. The error bars are calculated from the variance in the components of the Bloch vectors across all sequences in the ensemble. Simulated data for different error strengths δ are shown with circles. Experimental results are shown with red dotted lines and circles. (b) Trace distance for the sequence in the ensemble with the fewest H gates and least expected error shown with dashed lines and triangles. There are no error bars for these data, as there is only one sequence. Experimental results are shown with red dotted lines and triangles.

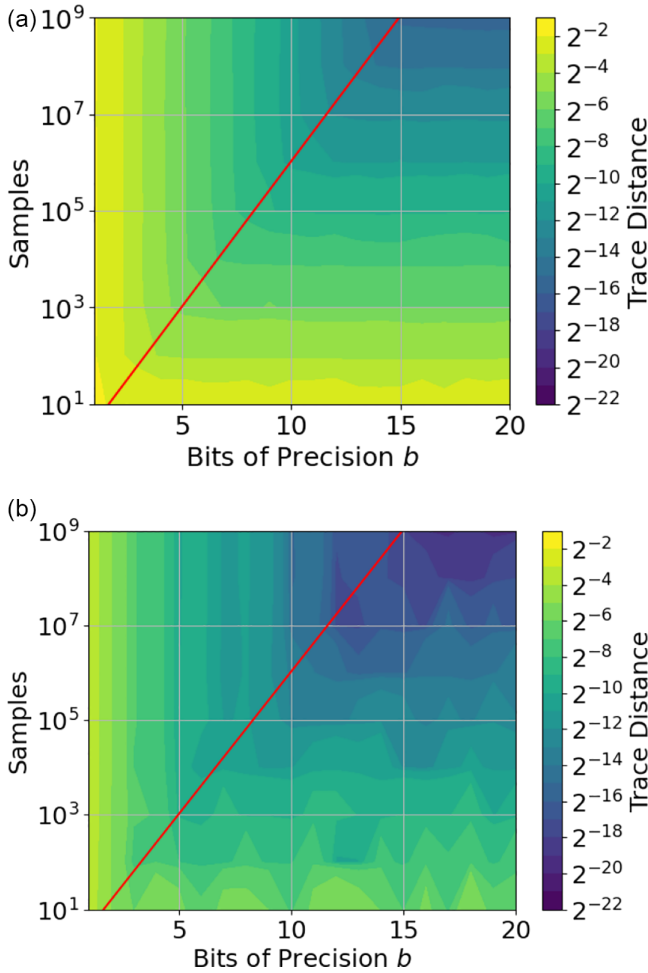


FIG. 7. Trace distance to an ideal R_Z rotation as a function of the number of samples and bits of precision. Results are shown (a) without averaging over random sequences and (b) with randomization over the maximum ensemble size. The red line indicates the number of samples needed to match the approximation error $\epsilon = 2^{-b}$ of an SK sequence. Using more samples above this line leads to diminishing returns in accuracy. Using a higher-precision sequence with fewer samples to the right of this line also leads to diminishing returns.

the Bloch sphere, and so in practice should be corrected in a manner similar to the synthesis error seen in the preceding section. In order to answer this question, we repeat our analysis with the simple model of coherent noise described in Sec. III.

Figures 4(b), 4(c), 4(e), and 4(f) show that each individual sequence experiences the coherent error differently, as evidenced by the varied distributions of sequence vectors \vec{s}_i as a function of the coherent noise parameter δ . However, the change is not totally random. The stronger the error strength is, the more the ensemble shifts to the left, manifesting as a single coherent rotation on the surface of the Bloch sphere. A negative value of δ instead shifts the ensemble to the right.

The performance of SKARC in the presence of coherent error is shown in Fig. 6. The trace distance to the target rotation for simulation data falls with increasing bits of precision for nonzero δ between $b = 2$ and $b = 7$, before rising again for nonzero coherent noise. This increase shows that noise

has a limiting effect on precision. Too much noise and it is impossible to realize highly precise rotations, regardless of whether randomization is used.

This increase in error happens for two reasons. Primarily, the coherent error model shifts the overall distribution away from the target rotation. This effect compounds with longer sequences as they have more coherent error overall; hence, the error increases for more bits of precision. This can be seen for both the ensemble mean and the sequence with the least theoretical error. The second reason is that, even when the coherent noise is tailored by randomization, it has an effect as depolarizing noise. The mean vector of the ensemble has a magnitude slightly less than that of any individual sequence, and for large enough error strengths this will additionally increase error as a function of sequence length. For smaller coherent errors $\delta \sim 1 \times 10^{-4}$, this increase in trace distance is delayed until around $b = 9$ bits of precision.

Comparing the simulated data in Figs. 6(a) and 6(b), the trace distances for the mean Bloch vector are smaller than the trace distance of the sequence with the fewest H gates and least expected error. The exceptions are for precision $b = 4$ and for some coherent noise strengths $b = 7$. We interpret this as the variability of the performance of the optimal sequence being larger than that of the mean, and therefore the optimal sequence sometimes outperforms the average, particularly when taking the overall shift of the ensemble into account.

C. Experimental results

We performed experiments on the QSCOUT device at Sandia National Laboratories [31], a trapped-ion NISQ computer. As the experiments are performed on a single ion and the depths of the circuits are considerable, driving the qubit transition of 12.6428 MHz directly with globally acting microwaves (with seconds-long coherence) is preferable over laser-based Raman gates. We do so with an external microwave horn in which the microwaves are controlled by custom control hardware system that allows for arbitrary phase, amplitude, and duration control [31]. We used $N = 24\,000 \sim 2^{14} \sim 4^7$ samples for each basis measurement due to time constraints, corresponding to a sampling error of $\epsilon = 1/\sqrt{N} \sim 0.006$. This means that for higher precisions than $b = 7$, finite sampling error dominates as opposed to the approximation error, as was discussed in Sec. IV A. As such, we restrict our experimental data to the range $b = [4, 5, 6, 7]$.

The Bloch vectors for each of the sequences R_i are shown in Fig. 4(d) for $b = 4$ as red crosses, with the mean shown as a red circle. For precision $b = 4$, the spread of the experimental data is consistent with a value of δ between 1×10^{-3} and 5×10^{-3} in simulations. The Bloch vector corresponding to the sequence with the fewest H gates and least expected error is shown with a red triangle as part of the ensemble.

Figure 5(b) shows the trace distance of the mean of different ensemble sizes m across different bits of precision b . The trace distance decreases with higher precision and as the size of the random ensemble increases. The error bars in Fig. 5(b) are calculated as the standard error on the mean of the distribution of trace distances for each possible combination of sequences for a given ensemble size. For most precisions, the sequence with the fewest H gates and least theoretical error is

further away than the mean sequence from averaging over the whole ensemble (the data points on the right). However, this is not the case for precision $b = 7$, where this shorter sequence is much more accurate than the average.

In Fig. 6(b) we give the average trace distance for both simulations including coherent noise and for experiment. In Fig. 6(b) we compare the trace distance of the sequence with the fewest H gates. Again we find that the experimental data correspond to an error rate of $\delta \sim 2 \times 10^{-3}$. We see larger trace distances for the sequence with the fewest H gates when compared to averaging across different precision, though the experimental data point at $b = 7$ is lower than expected, as was discussed previously. Moreover, this error rate is in agreement with (if slightly larger than) the estimated error in single-qubit gates from other experimental benchmarks of the QSCOUT device [31].

We see a reduction in the trace distance using SKARC to compute the mean vector compared to the sequence with the fewest H gates, except for the case of $b = 7$. Even in the presence of sampling noise and NISQ-era device noise, averaging the Bloch vectors of the ensemble renders a result that is closer to the target rotation, with improvement similar to that seen for the coherent noise model.

V. CONCLUSION

We found that it is possible to reduce the effects of a wide variety of physical errors and approximation errors inherent to Solovay-Kitaev decompositions of continuous rotations by averaging over an ensemble of distinct gate sequences. Multiple similar-yet-distinct sequences are generated using the gridsynth algorithm by varying the target rotation as described in Sec. II A. When acting on a single-qubit quantum state, these sequences result in an ensemble of Bloch vectors that are clustered around the target rotation state, as discussed in Sec. II B. By averaging the results of multiple circuits corresponding to each sequence, we were able to obtain quantum states that are closer to the ideal rotation as measured by the trace distance between the states.

We presented a noiseless simulation in Sec. IV A and analyzed the ensemble of sequences that results. Across multiple desired approximation precisions $\epsilon = 2^{-b}$, we found that the average of the full ensemble of sequences yields a Bloch vector that outperforms the sequence with the fewest H gates and least expected error. These results mirror those of other more sophisticated probabilistic gate synthesis algorithms such as convex hull finding [17,28]. We studied the effects of sampling noise on our method and found the expected trend: The number of samples needed scales exponentially with the desired precision of the sequence, and undersampling the circuit will lead to errors regardless of randomization. Similarly, oversampling the circuit will be useless if the precision of the rotation is not increased in turn.

Improvement using randomization is maintained under a coherent noise model. In Sec. IV B we explored the effects of a simple multiplicative coherent noise model on the rotations comprising H gates in our circuit, mirroring performance on QSCOUT, a NISQ trapped-ion computer. Under such a model, SKARC can mitigate coherent errors and reduce the trace distance to the target rotation, though a coherent bias remains and limits the effectiveness of the method at realizing precise rotations with large coherent error. For EFTQC architectures with small code sizes or where error correction is not performed after every single-qubit gate, this technique may improve accuracy and/or reduce costs.

Finally, we studied the technique on the QSCOUT device in Sec. IV C. We found that randomization can improve the accuracy of rotations on the device, though as with the coherent noise model, the resulting mean Bloch vector is still biased. The device performance is most closely replicated by a coherent error strength of $\delta \sim 2 \times 10^{-3}$, or a 0.2% overrotation, which matches with experimental benchmarks. Additionally, the number of samples required to realize high precision remains an obstacle in the near term.

These results provide a NISQ baseline for single-qubit gate synthesis algorithms which we can compare against EFTQC devices in the future. Further research should be carried out for a wider variety of noise models that better capture the effects of error-correcting codes. Additionally, it would be of interest to benchmark all or a portion of this technique on an error-corrected qubit to see if the results hold under that paradigm. Methods such as SKARC that rest at the intersection of error mitigation and error correction will become more important as the EFTQC era approaches.

ACKNOWLEDGMENTS

This project was funded by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research Quantum Testbed Program. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the U.S. Government.

DATA AVAILABILITY

The data that support the findings of this article are openly available [46].

-
- [1] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 [2] J. W. Z. Lau, K. H. Lim, H. Shrotriya, and L. C. Kwek, NISQ computing: Where are we and where do we go? *AAPPS Bull.* **32**, 27 (2022).

- [3] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, *Rev. Mod. Phys.* **95**, 045005 (2023).
 [4] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, 2013).

- [5] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, *et al.*, Fault-tolerant control of an error-corrected qubit, *Nature (London)* **598**, 281 (2021).
- [6] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, *Nature (London)* **614**, 676 (2023).
- [7] Google Quantum AI and Collaborators, Quantum error correction below the surface code threshold, *Nature (London)* **638**, 920 (2025).
- [8] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, *et al.*, Logical quantum processor based on reconfigurable atom arrays, *Nature (London)* **626**, 58 (2024).
- [9] I. Pogorelov, F. Butt, L. Postler, C. D. Marciniak, P. Schindler, M. Müller, and T. Monz, Experimental fault-tolerant code switching, *Nat. Phys.* **21** 298 (2025).
- [10] A. Katabarwa, K. Gratsea, A. Caesura, and P. D. Johnson, Early fault-tolerant quantum computing, *PRX Quantum* **5**, 020101 (2024).
- [11] A. Y. Kitaev, Quantum computations: Algorithms and error correction, *Russ. Math. Surv.* **52**, 1191 (1997).
- [12] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, *Electron. Colloq. Comput. Complex.* **TR96** (1994).
- [13] A. Y. Kitaev, A. H. Shen, and M. N. Vyalys, *Classical and Quantum Computation* (American Mathematical Society, Providence, 2002).
- [14] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, *Quantum Inf. Comput.* **6**, 81 (2006).
- [15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 2nd ed. (Cambridge University Press, Cambridge, 2010).
- [16] B. Koczor, Sparse probabilistic synthesis of quantum operations, *PRX Quantum* **5**, 040352 (2024).
- [17] S. Akibue, G. Kato, and S. Tani, Probabilistic unitary synthesis with optimal accuracy, *ACM Trans. Quantum Comput.* **5**, 1 (2024).
- [18] N. Yoshioka, S. Akibue, H. Morisaki, K. Tsubouchi, and Y. Suzuki, Error crafting in mixed quantum gate synthesis, *Quantum Phys.* **11**, 95 (2025).
- [19] D. Greenbaum and Z. Dutton, Modeling coherent errors in quantum error correction, *Quantum Sci. Technol.* **3**, 015007 (2016).
- [20] S. Bravyi, M. Englbrecht, R. König, and N. Peard, Correcting coherent errors with surface codes, *npj Quantum Inf.* **4**, 55 (2017).
- [21] N. Berthusen, S. J. S. Tan, E. Huang, and D. Gottesman, Adaptive syndrome extraction, *PRX Quantum* **6**, 030307 (2025).
- [22] K. Matsumoto and K. Amano, Representation of quantum circuits with Clifford and $\pi/8$ gates, [arXiv:0806.3834](https://arxiv.org/abs/0806.3834).
- [23] N. J. Ross and P. Selinger, Optimal ancilla-free Clifford+T approximation of z-rotations, *Quantum Inf. Comput.* **16**, 901 (2016).
- [24] M. Ware, G. Ribeill, D. Ristè, C. A. Ryan, B. Johnson, and M. P. da Silva, Experimental Pauli-frame randomization on a superconducting qubit, *Phys. Rev. A* **103**, 042604 (2021).
- [25] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor, *Phys. Rev. X* **11**, 041039 (2021).
- [26] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature (London)* **618**, 500 (2023).
- [27] M. B. Hastings, Turning gate synthesis errors into incoherent errors, *Quantum Inf. Comput.* **17**, 488 (2016).
- [28] E. Campbell, Shorter gate sequences for quantum computing by mixing unitaries, *Phys. Rev. A* **95**, 042306 (2017).
- [29] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, *Phys. Rev. A* **94**, 052325 (2016).
- [30] S. Zhuk, N. F. Robertson, and S. Bravyi, Trotter error bounds and dynamic multi-product formulas for Hamiltonian simulation, *Phys. Rev. Res.* **6**, 033309 (2024).
- [31] S. M. Clark, D. Lobser, M. C. Revelle, C. G. Yale, D. Bossert, A. D. Burch, M. N. Chow, C. W. Hogle, M. Ivory, J. Pehr, B. Salzbrenner, D. Stick, W. Sweatt, J. M. Wilson, E. Winrow, and P. Maunz, Engineering the quantum scientific computing open user testbed, *IEEE Trans. Quantum Eng.* **2**, 1 (2021).
- [32] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, J. G. Bohnet, N. C. Brown, N. Q. Burdick, W. C. Burton, S. L. Campbell, J. P. Campora, C. Carron, J. Chambers, J. W. Chan, Y. H. Chen, *et al.*, A race-track trapped-ion quantum processor, *Phys. Rev. X* **13**, 041052 (2023).
- [33] IonQ Staff, *IonQ Aria: Practical performance* (IonQ, College Park, 2024), <https://ionq.com/resources/ionq-aria-practical-performance>.
- [34] S. B. Bravyi and A. Y. Kitaev, Quantum codes on a lattice with boundary, [arXiv:quant-ph/9811052](https://arxiv.org/abs/quant-ph/9811052).
- [35] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A* **86**, 032324 (2012).
- [36] A. M. Steane, Error correcting codes in quantum theory, *Phys. Rev. Lett.* **77**, 793 (1996).
- [37] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).
- [38] A. J. Landahl, J. T. Anderson, and P. R. Rice, Fault-tolerant quantum computing with color codes, [arXiv:1108.5738](https://arxiv.org/abs/1108.5738).
- [39] E. Knill, R. Laflamme, and W. Zurek, Threshold accuracy for quantum computation, [arXiv:quant-ph/9610011](https://arxiv.org/abs/quant-ph/9610011).
- [40] E. Knill, R. Laflamme, and W. Zurek, Resilient quantum computation: Error models and thresholds, *Proc. R. Soc. London A* **454**, 365 (1998).
- [41] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, *Nature (London)* **549**, 172 (2017).
- [42] C. Gidney, N. Shutty, and C. Jones, Magic state cultivation: Growing T states as cheap as CNOT gates, [arXiv:2409.17595](https://arxiv.org/abs/2409.17595).
- [43] B. Zhang, S. Majumder, P. H. Leung, S. Crain, Y. Wang, C. Fang, D. M. Debroy, J. Kim, and K. R. Brown, Hidden inverses: Coherent error cancellation at the circuit level, *Phys. Rev. Appl.* **17**, 034074 (2022).
- [44] T. Müller, T. Stollenwerk, D. Headley, M. Epping, and F. K. Wilhelm, Coherent and non-unitary errors in

- ZZ*-generated gates, *Quantum Sci. Technol.* **10**, 015058 (2025).
- [45] A. J. Landahl, D. S. Lobser, B. C. A. Morrison, K. M. Rudinger, A. E. Russo, J. W. Van Der Wall, and P. Maunz, Jaqal, the quantum assembly language for QSCOUT, [arXiv:2003.09382](https://arxiv.org/abs/2003.09382).
- [46] O. Widzowski Maupin, A. Burch, C. Yale, M. Chow, T. Colvin, B. Ruzic, M. Reville, B. McFarland, E. Ibarra Garcia Padilla, A. Rascon, A. Landahl, S. Clark, *et al.*, Data and code for “Solovay-Kitaev Algorithm and Randomized Compilation” Zenodo (2026), <https://doi.org/10.5281/zenodo.19855220>.