Local gradient optimization of leakage-suppressing entangling sequences

A. A. Setser and J. P. Kestner

Phys. Rev. A **103**, 012609 — Published 26 January 2021

DOI: 10.1103/PhysRevA.103.012609

Local Gradient Optimization of Leakage-Suppressing Entangling Sequences

A. A. Setser[1] and J. P. Kestner[1]

[1] *Department of Physics, University of Maryland Baltimore County, Baltimore, Maryland 21250, USA*

We use a gradient-based optimization scheme to find single-qubit rotations to be interwoven between timesteps of a noisy logical two-qubit entangling gate in order to suppress the effects of arbitrary logical and leakage noise in the two-qubit gate. We show how the sequence fidelity is affected by imperfections in the single-qubit operations, as well as by various relative strengths of the logical and leakage noise. Our approach is completely general and system-independent, allowing for application to any two-qubit system regardless of the experimental implementation details.

I. INTRODUCTION

Reliable implementation of two-qubit entangling gates is a key step towards creating a useful quantum computer. In order for fault-tolerant quantum computing to be possible, operations on qubits must be performed with error rates less than the "error correction threshold", with the exact value of the threshold varying with the qubit encoding scheme. One of the highest error correction thresholds is around 1%, offered by surface codes [1]. However, it is still desirable to reduce errors as much as possible, in order to reduce the surface code overhead.

The difficulty in physical implementation arises when environmental effects are taken into consideration. Interactions between the system and environment can entangle the two, causing a collapse of the wavefunction, thus destroying the quantum properties of the system. These are called incoherent errors, and cannot be reversed via unitary operations on the system. However, system-environment interactions can also introduce random perturbative effects within the system Hamiltonian, causing the coherent evolution of the qubit to differ from the unperturbed evolution. Such effects are known as coherent errors, and it is possible to dynamically correct these errors via unitary operations on the system [2, 3]. In this work, we focus on such reduction of coherent errors.

For single qubits, coherent errors are often reduced through a wide variety of composite pulse sequences, which typically consist of piecewise constant values of the system parameters chosen such that the errors incurred during each timestep cancel with those of the other timesteps in the evolution, causing the final evolution to be error-free up to a given order in the perturbation [4–6]. For two-qubit entangling gates, that approach is complicated by the larger dimensionality of the Hilbert space. However, the potential benefit is even greater than in the single-qubit case because the errors are generally larger. That is due to the longer evolution times required for two-qubit entangling gates since the qubit-qubit interaction term in the system Hamiltonian is often weak compared to the single-qubit terms.

It is known that high-fidelity single-qubit operations can be used as a resource in two-qubit pulse sequences to suppress the effects of arbitrary coherent noise within the two-qubit logical subspace. These two-qubit pulse sequences are typically found analytically and are restricted to either small numbers of single-qubit rotations, or allow for larger numbers of single-qubit operations at the cost of restricting them to simple π rotations for simplicity [7–12]. However, this ease of experimental implementation typically comes at the cost of reduced noise suppression.

This contrasts with pulse shaping through optimal control methods, where parameters in the system Hamiltonian are varied in time according to numerically generated waveforms, so that a desired outcome is achieved, e.g., a noise-free system evolution. Krotov's method and gradient-ascent-pulse-engineering (GRAPE) are common gradient-based optimal control methods [13–18], while CRAB and DCRAB [19–21] are common gradient-free methods. The efficiency of gradient-based and gradient-free methods depends on the number of optimization parameters, as well as the behavior of the optimization functional. Thus, different methods are chosen based on the specific optimization problem.

Such methods have proven to be successful in a range of more general quantum control problems [22–27], as well as for optimal control of two-qubit gates [28–35]. We have also applied similar techniques previously, where we showed that a GRAPE method can be used to find optimal single-qubit rotations to suppress arbitrary logical noise in a two-qubit gate [36]. It was shown via simulations that the numerically optimized sequences were generally more effective at suppressing arbitrary logical noise than the analytically derived sequences discussed previously, at the cost of being somewhat more difficult to implement experimentally since the single-qubit rotations were not restricted to simple rational multiples of π.

While our previous method focused on suppressing logical noise only, we are also interested in the possibility of suppressing leakage noise, as may be the case in a system like a superconducting qubit with multiple energy levels [37–39]. In this paper, we present a variant of the numerical optimization scheme developed in Ref. [36] that now addresses both leakage and logical errors simultaneously. We optimize single-qubit rotations inserted between applications of a noisy two-qubit entangling gate such that the final sequence performs a logical entangling operation while suppressing all coherent noise. Despite the fact that the inserted rotations are restricted

to act within the logical subspace whereas the error acts in a much larger space, our method is surprisingly effective. We also show that it is relatively unaffected by imperfections in the interwoven single-qubit operations and we demonstrate how the performance changes with varying logical and leakage noise strengths. The modular, system-independent nature of our approach facilitates application to any two-qubit setup, regardless of the details of the Hamiltonian.

II. MODEL AND OPTIMIZATION SCHEME

We model each qubit as a two level system, coupled to a third leakage level of higher energy. This could be, for example, an excited state of a weakly anharmonic superconducting qubit or an excited valley state of a silicon spin qubit. Although more than one leakage level may exist, the population of such levels becomes increasingly unlikely as the energy of the leakage level increases. We therefore consider only a single leakage level. In order for the numerical optimization scheme to be effective for a wide range of noise values, we sample M total noise realizations and require that the optimized single-qubit operations perform well over the average of these realizations [31].

Within a noise realization m, we assume that there exists an evolution operator $U^{(m)}$, such that the system state $|\psi(t)\rangle$ evolves according to $|\psi(T)\rangle = U^{(m)}|\psi(0)\rangle$, where T is the gate duration. The single-qubit evolution operators reside in SU(3), which is generated by the Gell-Mann matrices, λ_i, with $i \in [0, 8]$. The forms of the matrices are shown in Appendix A. We denote the upper 2×2 block of $U^{(m)}$ as the logical subspace. The logical subset of operations is generated by λ_1, λ_2, and λ_3, since they are respectively equal to the Pauli matrices σ_X, σ_Y, and σ_Z within the logical subspace and are zero outside of it. Leakage effects are generated by λ_4 through λ_8, since they contain terms which couple the logical subspace to the leakage subspace.

Following Ref. [36], we choose a time evolution operator composed of a series of N time steps with the form

$$U^{(m)} = \prod_{n=N}^{1} \exp\left[-\frac{i\pi}{N}\lambda_{3,3}\right] \exp\left[-\frac{i}{N}\Delta^{(m)}\right] R_n, \quad (1)$$

where $\lambda_{i,j} = \lambda_i \otimes \lambda_j$ and the R_n's are arbitrary single-qubit rotations. Within each step of the evolution, $\exp[-i\pi\lambda_{3,3}/N]$ is equal to the Nth root of a 2π conditional phase gate within the logical subspace and an identity operation outside of it. We do not make any assumptions as to how this effective logical Ising-like operation is implemented. While a strictly logical σ_{ZZ} Hamiltonian is one possibility, such an operation can also be formed by any strictly logical entangling operation plus the appropriate logical single-qubit operations. The exact Hamiltonian, while relevant in an experimental implementation, is not directly relevant for the theory presented here.

Imperfections in the evolution operator are introduced through the term $\exp\left[-i\Delta^{(m)}/N\right]$, where

$$\Delta^{(m)} = \sum_{ij} \delta_{i,j}^{(m)} \lambda_{i,j} \quad (2)$$

and the $\delta_{i,j}^{(m)}$ are randomly chosen (generally small relative to 2π) numbers. We refer to $\delta_{i,j}^{(m)}$ as the noise coefficients in the evolution operator since they serve as small perturbations multiplying all possible evolution generators $\lambda_{i,j}$. The sum over i and j ensures that both logical and leakage noise coefficients are considered. Logical noise coefficients are those with $i, j \in \{0, 1, 2, 3\}$ (not including the identity-multiplying coefficient $i = j = 0$) which serve as perturbations within the two-qubit logical subspace. Leakage noise coefficients are those with the remaining i, j which serve to couple the logical subspace to terms outside of the logical subspace.

We have assumed that the strength of the noise coefficients scales with the phase of the entangling gate. This is clearly true in the most straightforward case where the entanglement is generated by a Hamiltonian $H = \lambda_{3,3} + \delta H_{\text{drift}}$. Then the leading order effect of the random drift term on the evolution is $\exp(-i\pi\lambda_{3,3}/N)\exp(-i\pi\delta H_{\text{drift}}/N)$. Similarly, if the entanglement is generated by a Hamiltonian like $H = \lambda_{1,1} + \lambda_{2,2} + \delta H_{\text{drift}}$, then the desired building block can be obtained via the sequence $\sigma_{HH}\exp(-i\pi H/2N)\sigma_{XI}\exp(-i\pi H/2N)\sigma_{XI}\sigma_{HH} = \exp(-i\pi\lambda_{3,3}/N)\exp(-i\mathcal{O}(\delta/N))$, where σ_{HH} is shorthand for application of a local Hadamard on each qubit and σ_{XI} is a local X gate on one qubit. In this example, if one works out the $\delta_{i,j}^{(m)}$ corresponding to Eq. (2) one finds they have additional sinusoidal dependences on N, but that is irrelevant since we average over noise realizations. The key point is that there is again an overall scale factor of δ/N. More generally though, when the entangling gate is implemented via some other time-dependent entangling Hamiltonian, possibly with local rotations interspersed, the noise coefficients do not necessarily scale as $1/N$ in this way. Nonetheless, we will proceed with this heuristic assumption.

Note that the model for the noise does not depend on n and is therefore the same throughout multiple timesteps in the evolution operator $U^{(m)}$. This corresponds to noise which would be generated through a static physical process, i.e., one which is not varying in time. Besides this static constraint, we do not reference any specific physical noise process in our work. Any common physical noise process would likely have most of the $\delta_{i,j}^{(m)}$ equal to 0, and only a few would be nonzero. However, we take all $\delta_{i,j}^{(m)}$ to be nonzero in order to show that our method is effective for the worst case scenario when all of the evolution generators are affected by noise, even though such a case is unlikely for a real system.

The form for the evolution operator essentially consists of splitting a noisy 2π conditional phase gate into N timesteps and inserting arbitrary single-qubit rotations in

between them. We denote such a sequence of operations as a length-N sequence. The single-qubit rotations steer the evolution dynamics, so that the final interaction between qubits in the logical subspace is not limited to the σ_{ZZ} interaction that would be generated in the absence of the single-qubit rotations. We choose a Pauli vector parametrization for the single-qubit operations,

$$\begin{aligned} R_n = &\exp\left[i\left(\alpha_{1,n}\lambda_1 + \beta_{1.n}\lambda_2 + \gamma_{1,n}\lambda_3\right)\right] \\ &\otimes \exp\left[i\left(\alpha_{2,n}\lambda_1 + \beta_{2.n}\lambda_2 + \gamma_{2,n}\lambda_3\right)\right], \end{aligned} \tag{3}$$

where $\alpha_{1,n}\ldots\gamma_{2,n}$ are the free parameters for optimization. While the exact details of the implementation of the single-qubit operations are not directly relevant in our work, we consider the case when the error (both coherent and incoherent) from the single-qubit rotations is negligible compared to the error from the entangling segments. This can be ensured by performing the rotations much faster than the entangling gates or via a dynamically corrected pulse to reduce their error.

The optimal free parameters are the ones that minimize the optimization function we choose. There are two requirements for our functional: reduce the effects of the noise coefficients in the final operation (both in the logical and leakage subspace) and generate a perfect entangler within the logical subspace. A perfect entangler is a gate which can produce a maximally entangled state from an unentangled one [40]. The suppression of the effects of the noise terms can be measured by first comparing the noisy gate $U^{(m)}$ to the gate in the absence of noise. We denote this quantity as the gate fidelity and it is given by

$$F\left(U^{(m)}\right) = \frac{1}{81}\left|\mathrm{tr}\left(O^\dagger U^{(m)}\right)\right|^2, \tag{4}$$

where

$$O = \prod_{n=1}^{N} \exp\left[-\frac{i\pi}{N}\lambda_{3,3}\right] R_n. \tag{5}$$

The gate error due to noise is then defined as

$$\varepsilon\left(U^{(m)}\right) = 1 - F\left(U^{(m)}\right). \tag{6}$$

Note that the target gate includes the single-qubit rotations and is therefore changing over the course of the optimization. Thus, the final operation is not known ahead of time and is arbitrary. However, the final single-qubit operations R_n will be known, and thus O will be known at the end of the optimization. Thus, optimization of the gate error due to noise with respect to the free parameters will produce a known final operation which is robust against noise.

Although this would produce a dynamically corrected operation, it is unlikely to be a perfect entangler within the logical subspace, since the single-qubit rotations will affect the entanglement dynamics. Since $U^{(m)}$ is in SU(9), we quantify the qubit-qubit entanglement it produces by first projecting onto the SU(4) logical subspace in order to obtain a nonunitary effective logical evolution, $U^{(m)}_{\mathrm{proj}}$, then examining the Makhlin invariants of the projected operation, $g_1\left(U^{(m)}_{\mathrm{proj}}\right), g_2\left(U^{(m)}_{\mathrm{proj}}\right)$, and $g_3\left(U^{(m)}_{\mathrm{proj}}\right)$ [41]. Although this definition is strictly valid only when the total leakage vanishes, since the leakage is being minimized via Eqs. (4) and (5), it is an effective way to quantify entangling power within the cost function for the purposes of optimization. Specifically, the distance between a logical two-qubit operation and the nearest perfect entangler can be expressed according to the Makhlin invariants as [42]

$$d\left(U^{(m)}_{\mathrm{proj}}\right) = g_3\sqrt{g_1^2 + g_2^2} - g_1. \tag{7}$$

This distance measure can take on negative values for certain operations, which can be problematic for the optimization routine [36]. To check when these problematic operations occur, we calculate the quantity

$$s\left(U^{(m)}_{\mathrm{proj}}\right) = \pi - \cos^{-1}(z_1) - \cos^{-1}(z_3) \tag{8}$$

from the ordered roots (z_1, z_2, z_3) of the equation [43]

$$z^3 - g_3 z^2 + \left(4\sqrt{g_1^2 + g_2^2} - 1\right) z + (g_3 - 4g_1) = 0, \tag{9}$$

and note that the actual distance metric that should be minimized in order to realize a logical perfect entangler is

$$\mathcal{D}\left(U^{(m)}_{\mathrm{proj}}\right) = \begin{cases} d & d > 0 \text{ and } s > 0 \\ -d & d < 0 \text{ and } s < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

This is a true metric, in the sense that it is positive for non-perfect entanglers and equal to zero for perfect entanglers.

The total functional for optimization is the sum of the gate error due to noise and the distance to the nearest perfect entangler averaged over all noise realizations,

$$J = \frac{1}{M}\sum_{m=1}^{M} \varepsilon\left(U^{(m)}\right) + \mathcal{D}\left(U^{(m)}_{\mathrm{proj}}\right). \tag{11}$$

Minimization of the functional serves to produce a perfect entangler within the logical subspace that is robust against logical and leakage noise.

We use the L-BFGS-B gradient-based minimization algorithm [44], which is implemented within the SciPy optimization package [45]. We choose a gradient-based minimization algorithm, since we have found that it is more efficient than gradient-free algorithms for the numbers of optimization parameters we consider [46]. The L-BFGS-B algorithm also offers an increase in convergence speed through the estimation of the Hessian of the functional. SciPy's implementation of the algorithm also allows the

gradient of the functional to be estimated numerically, so we do not need to calculate the analytic gradient of J.

Since the L-BFGS-B algorithm is a *local* search method, the convergence of the routine is highly dependent on the initial "guess" parameters we choose at the start of the minimization. In order to choose an effective initial guess, if the greatest divisor of N is d, we repeat the solution for the length d sequence N/d times to use as the guess for the length N sequence. This ensures that longer-length sequences will be constructed from shorter-length sequences which have already been optimized according to our criteria. For prime length sequences, we set the guess for the free parameters to be all zeroes, so that the single-qubit operations are initialized as identity operations.

III. RESULTS

When evaluating the success of the optimization routine, we separately track our two criteria: suppressing the effects of noise on the final operation and generating a logical perfect entangler. In order to determine the noise present in the final operations, we calculate the final gate error due to noise via Eq. (6). In order to determine the entanglement dynamics of the final operations, we consider the entanglement fidelity of an SU(4) operation, U, which is given by

$$F_{\mathrm{PE}}(U) = \frac{1}{16}\left|\mathrm{tr}\left(V^{\dagger}U\right)\right|^{2}, \tag{12}$$

where V is the closest perfect entangler to U. Eq. 12 thus serves as a measure of overlap between U and the nearest perfect entangler, which is an effective measure of the entanglement capabilities of U. Since U is in SU(4), it can be written in terms of a Cartan decomposition as,

$$U = k_1 A k_2, \tag{13}$$

where k_1, and k_2 are purely local operations and $A = \exp\left[-i/2\left(c_1\sigma_{XX} + c_2\sigma_{YY} + c_3\sigma_{ZZ}\right)\right]$ is a purely nonlocal operation in SU(4) $\setminus$ SU(2) $\otimes$ SU(2) [47, 48]. c_1, c_2, and c_3 are known as the Weyl chamber coordinates of U [43].

Ref. 42 shows that this fidelity can be simplified to

$$F_{\mathrm{PE}}(U) = \begin{cases} \cos^4\left(\frac{c_1+c_2-\frac{\pi}{2}}{4}\right) & c_1+c_2 \le \frac{\pi}{2} \\ \cos^4\left(\frac{c_2+c_3-\frac{\pi}{2}}{4}\right) & c_2+c_3 \ge \frac{\pi}{2} \\ \cos^4\left(\frac{c_1-c_2-\frac{\pi}{2}}{4}\right) & c_1-c_2 \ge \frac{\pi}{2} \\ 1 & \text{otherwise.} \end{cases} \tag{14}$$

Note that Eqs. (12) and (14) are squared, unlike in Ref. 42. This is done to be consistent with our definition of gate fidelity, Eq. (4), which is also squared.

The error associated with the distance to the nearest perfect entangler, averaged over noise realizations, is then

$$\varepsilon_{\mathrm{PE}} = \frac{1}{M}\sum_{m=1}^{M} 1 - F_{\mathrm{PE}}\left(U_{\mathrm{proj}}^{(m)}\right), \tag{15}$$

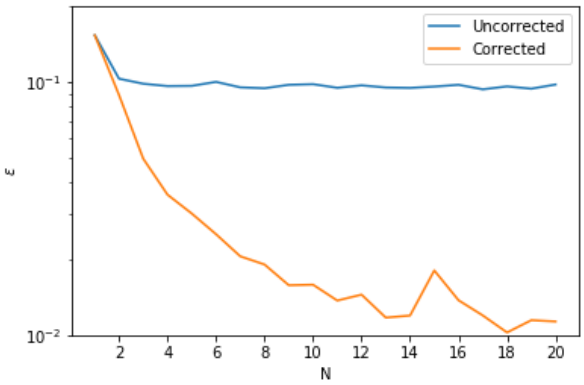

FIG. 1. Gate error due to noise, Eq. (6), in relation to sequence length for a logical σ_{ZZ} interaction, assuming access to perfect single-qubit rotations.

where $U_{\mathrm{proj}}^{(m)}$ is the projection of our final optimized SU(9) operation onto the logical subspace. The projection is necessary since Eq. (14) requires a logical 4×4 operation. Since this quantity is between 0 and 1, it is a more direct measure of the final entanglement capabilities of the operations, compared to the metric $\mathcal{D}$ that was used in the optimization. Note that this entanglement error is completely independent from the earlier defined gate error due to noise $\varepsilon(U^{(m)})$. The entanglement error is meant to quantify the distance between a gate and a perfect entangler, while $\varepsilon(U^{(m)})$ quantifies the effects of noise on the gate.

The noise coefficients $\delta_{i,j}^{(m)}$ are drawn randomly from a normal distribution with a standard deviation of $\sigma_{\mathrm{nonlocal}} = 0.065$. This value is chosen so that when no rotations are inserted, the gate error due to noise is around 10%, which is a realistic situation [49]. We find that $M = 100$ is enough to ensure that our results are robust against a general noise realization, i.e., the optimized solutions obtained by running the routine over different sets of 100 noise realizations do not change significantly.

The results of the optimization are shown in Fig. 1. While there are two separate quantities that track the optimization performance, Eqs. (15) and (6), Fig. 1 shows only the gate error due to noise, ε, and not the perfect entangler error $\varepsilon_{\mathrm{PE}}$. This is due to the fact that $\varepsilon_{\mathrm{PE}}$ is equal to 0 for all sequences with $N > 2$ and thus does not need to be shown. As seen in the plot, the gate error due to noise initially decreases rapidly with increasing N, but the gains diminish as the sequence grows longer. We do not know what causes this saturation at large N, but it is useful that order-of-magnitude improvements in the gate error due to noise can be obtained already with $N \sim 10$. A supplementary data file is available which contains the solutions for all sequence lengths for this case and all further cases we consider [50].

We also consider the effects of imperfections in the

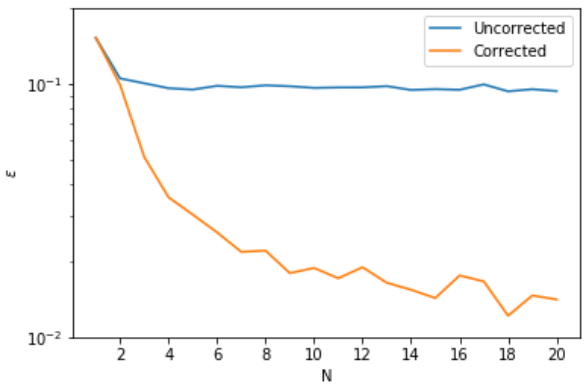

FIG. 2. Gate error due to noise, Eq. (6), in relation to sequence length for a logical σ_{ZZ} interaction, for the case of noisy single-qubit rotations

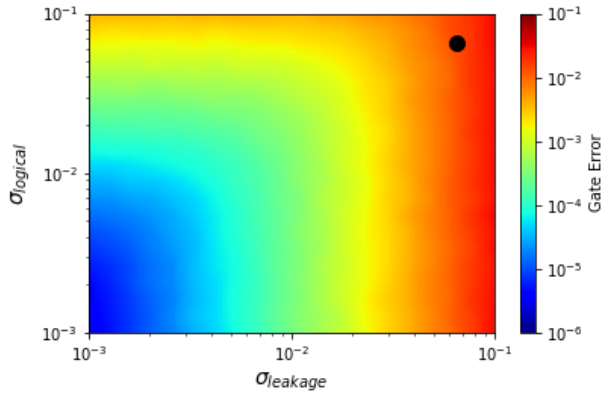

FIG. 3. Gate error due to noise for the $N = 16$ solution obtained in the absence of local noise, in relation to varying standard deviations for the logical and leakage noise. The solution obtained was optimized at $\sigma_{\text{logical}} = \sigma_{\text{leakage}} = 0.065$ (marked on plot).

local operations on the performance of the optimization routine. Like the two-qubit operations, we assume that the local operations have both logical and leakage noise. Noise within the logical subspace is modeled by introducing perturbations into the control parameters $\eta_i \in \{\alpha_{1,i}, \ldots, \gamma_{2,i}\}$ according to

$$\eta_i \to \eta_i' = \eta_i(1 + \delta_\eta), \tag{16}$$

where δ_η is a noise coefficient drawn randomly from a normal distribution with a standard deviation of σ_{local}.

The leakage noise is introduced by multiplying the local operations R_n by the factor

$$\prod_{k=4}^{8} \exp\left(i\sqrt{\alpha_{1,n}^2 + \beta_{1,n}^2 + \gamma_{1,n}^2}\,\delta_k \lambda_k\right) \otimes \exp\left(i\sqrt{\alpha_{2,n}^2 + \beta_{2,n}^2 + \gamma_{2,n}^2}\,\delta_k' \lambda_k\right), \tag{17}$$

where the δ_k and δ_k' are noise coefficients drawn randomly from a normal distribution, also taken to have a standard deviation of σ_{local} for simplicity. The choice for the form of the leakage noise ensures that the noise introduced is proportional to the magnitude of the logical rotations being performed. This is a realistic situation, since larger rotations generally correspond to longer gate times and thus introduce more noise. The standard deviation for the distribution of the local noise coefficients is taken to be $\sigma_{\text{local}} = 0.002$, so that the local rotations have a fidelity of approximately 99.9% when calculated according to

$$F_R = \frac{1}{81}\left|\text{tr}\left(R^\dagger\left(\alpha_1', \ldots, \gamma_2'\right) R\left(\alpha_1, \ldots, \gamma_2\right)\right)\right|^2, \tag{18}$$

and averaged over 1000 sets of error coefficients and 1000 sets of angles drawn randomly from a uniform distribution ranging from -2π to 2π.

The results of this optimization are shown in Fig. 2. We again only show the gate error due to noise, ε, and not the perfect entangler error, ε_{PE}, since all optimized results with $N > 2$ have $\varepsilon_{\text{PE}} = 0$. The scaling in this case is similar although slightly worse than the noise-free local rotation case, achieving a maximum gate fidelity of 98.8% compared to 99.0%. Thus, imperfections in the local operations have only marginal effects at the small sequence lengths we consider. As the sequence length increases, previous work has shown that the effects of single-qubit noise can accumulate to the point where gate errors due to noise begin to increase with increasing N [36]. Furthermore, the solutions obtained in the presence of local noise will perform just as well if the noise is nonlocal only, whereas solutions obtained in the absence of local noise will not significantly decrease the gate error if local noise is introduced.

While the optimization producing Fig. 1 was performed assuming equal strengths for the logical and leakage noise (i.e., that the $\delta_{i,j}$s of Eq. (2) are all drawn from the same distribution), we wish to see how the solutions hold as we separately vary these noise strengths. Figure 3 shows the results of the $N = 16$ solution obtained in the absence of local noise, in relation to varying standard deviations for the logical and leakage noise. The solution is taken from Fig. 1 case, which was optimized with $\sigma_{\text{logical}} = \sigma_{\text{leakage}} = 0.065$. From Fig. 3, we see that the optimized solution is more sensitive to leakage noise than logical noise. This is reasonable, since the terms in Eq. (2) which generate logical noise have $i, j \in \{0, 1, 2, 3\}$ (not including the identity term, $i = j = 0$), while the rest of the terms generate leakage noise. This gives 15 logical noise generators and 65 leakage noise generators, making the overall sequence more susceptible to leakage noise.

So far we have only considered a two-qubit $\lambda_{3,3}$ interaction in Eq. (1), i.e, a σ_{ZZ} interaction in the logical subspace. However, the performance of the optimization

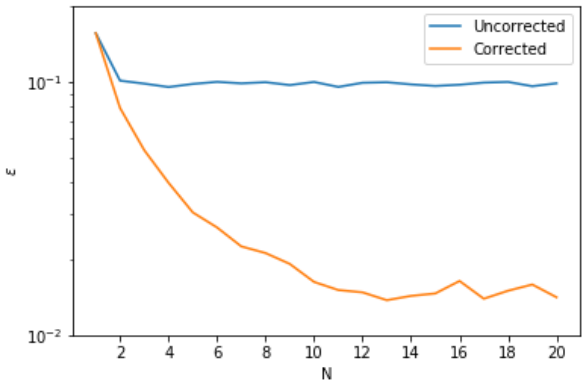

FIG. 4. Gate error due to noise, Eq. (6), in relation to sequence length for a logical $\sigma_{XX} + \sigma_{YY}$ interaction, for the case of noisy single-qubit rotations with error-free logical σ_Z rotations

routine is not limited strictly to this form for the interaction. We can also consider a logical $\sigma_{XX} + \sigma_{YY}$ interaction, which is relevant for many superconducting qubit setups [51]. This is reflected by changing $\lambda_{3,3} \rightarrow \lambda_{1,1} + \lambda_{2,2}$ in Eq. (1). In these types of systems, one can also typically perform "virtual Z gates," in which local logical σ_Z rotations can be performed instantaneously in software by changing the reference phase of the microwave pulses that drive single-qubit rotations [52]. Allowing these noise-free logical σ_Z rotations in our optimization corresponds to setting $\delta_{\gamma_1} = \delta_{\gamma_2} = 0$. (We have already observed that the optimization is not strongly affected by small local noise values, so accounting for virtual gating actually doesn't make a big difference, but we do so just to show that it is not difficult to incorporate such considerations.) We again take $\sigma_{\text{nonlocal}} = 0.065$ and $\sigma_{\text{local}} = 0.002$. The results of this optimization are shown in Fig. 4. The performance of the optimization routine is similar to the logical σ_{ZZ} interaction case, achieving a maximum gate fidelity of 98.6%. As with the previous cases, all sequences with $N > 2$ have $\varepsilon_{\text{PE}} = 0$. Thus, the optimization routine is not significantly affected by changing the logical two-qubit interaction from σ_{ZZ} to $\sigma_{XX} + \sigma_{YY}$.

IV. CONCLUSION

We have shown that two-qubit logical entangling gates with fidelities around 90% can be used in conjunction with logical single-qubit operations to construct two-qubit entangling gates with errors of around 1%. The logical single-qubit operations are interwoven between timesteps of the entangling operation and effectively suppress the effects of both arbitrary logical and leakage noise present in the entangling gates.

We have shown that our numerical optimization is effective even when imperfections in the single-qubit operations are considered. In addition, we have shown that our method is effective both for a logical two-qubit σ_{ZZ} interaction and a $\sigma_{XX} + \sigma_{YY}$ interaction. For the σ_{ZZ} interaction case, we have shown how the optimized solutions depend individually on the strengths of the logical and leakage noise present. The modular nature of this approach allows for application to any two-qubit system, regardless of the Hamiltonian.

ACKNOWLEDGMENTS

This research was sponsored by the Army Research Office (ARO), and was accomplished under Grant Number W911NF-17-1-0287.

Appendix A: Gell-Mann Matrices

For completeness, the Gell-Mann matrices are presented here. They are given by

$$\lambda_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

[1] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A **83**, 020302 (2011).
[2] H. K. Cummins, G. Llewellyn, and J. A. Jones, Phys. Rev. A **67**, 042308 (2003).
[3] J. A. Jones, Prog. Nucl. Magn. Res. Spec. **59**, 91 (2011).
[4] W. Yang, Z.-Y. Wang, and L. Ren-Bao, Frontiers of Physics in China **6**, 2 (2010).
[5] J. T. Merrill and K. R. Brown, "Progress in compensating pulse sequences for quantum computation," in *Quantum Information and Computation for Chemistry* (Wiley,

2014).
[6] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2005).
[7] Y. Tomita, J. T. Merrill, and K. R. Brown, New Journal of Physics **12**, 015002 (2010).
[8] I. Cohen, A. Rotem, and A. Retzker, Phys. Rev. A **93**, 032340 (2016).
[9] J. A. Jones, Phys. Rev. A **67**, 012317 (2003).
[10] T. Ichikawa, U. Güngördü, M. Bando, Y. Kondo, and M. Nakahara, Phys. Rev. A **87**, 022323 (2013).
[11] F. A. Calderon-Vargas and J. P. Kestner, Phys. Rev. Lett. **118**, 150502 (2017).
[12] C. D. Hill, Phys. Rev. Lett. **98**, 180501 (2007).
[13] A. Konnov and V. F. Krotov, Autom. Rem. Contr. **60** (1999).
[14] S. E. Sklarz and D. J. Tannor, Phys. Rev. A **66**, 053619 (2002).
[15] J. P. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003).
[16] D. M. Reich, M. Ndong, and C. P. Koch, J. Chem. Phys. **136**, 104103 (2012).
[17] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magnet. Res. **172**, 296 (2005).
[18] P. de Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov, JMR **212**, 412 (2011).
[19] N. Rach, M. M. Muller, T. Calarco, and S. Montangero, Phys. Rev. A **92**, 062343 (2015).
[20] T. Caneva, T. Calarco, and S. Montangero, Phys. Rev. A **84**, 022326 (2011).
[21] J. A. Nelder and R. Mead, Comput. J. **7**, 308 (1965).
[22] C. Brif, R. Chakrabarti, and H. Rabitz, New Journal of Physics **12**, 075008 (2010).
[23] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. A **69**, 042309 (2004).
[24] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, Eur. Phys. J. D **69**, 279 (2015).
[25] C. P. Koch, J. Phys.: Condens. Matter **28**, 213001 (2016).
[26] H. Zhang and H. Rabitz, Phys. Rev. A **49**, 2241 (1994).
[27] R. L. Kosut, M. D. Grace, and C. Brif, Phys. Rev. A **88**, 052326 (2013).
[28] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. Lett. **91**, 027903 (2003).
[29] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. Lett. **93**, 020502 (2004).
[30] M. M. Muller, D. M. Reich, M. Murphy, H. Yuan, J. Vala, K. B. Whaley, T. Calarco, and C. P. Koch, Phys. Rev. A **84**, 042315 (2011).
[31] M. H. Goerz, E. J. Halperin, J. M. Aytac, C. P. Koch, and K. B. Whaley, Phys. Rev. A **90**, 032329 (2014).
[32] M. H. Goerz, G. Gualdi, D. M. Reich, C. P. Koch, F. Motzoi, K. B. Whaley, J. Vala, M. M. Müller, S. Montangero, and T. Calarco, Phys. Rev. A **91**, 062307 (2015).
[33] D. Dong, C. Chen, B. Qi, I. R. Petersen, and F. Nori, Sci. Rep. **5**, 7873 (2015).
[34] C.-H. Huang and H.-S. Goan, Phys. Rev. A **95**, 062325 (2017).
[35] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, Phys. Rev. A **72**, 042331 (2005).
[36] A. A. Setser, M. H. Goerz, and J. P. Kestner, Phys. Rev. A **97**, 062339 (2018).
[37] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Applied Physics Reviews **6**, 021318 (2019).
[38] G. Wendin, Reports on Progress in Physics **80**, 106001 (2017).
[39] M. J. Peterer, S. J. Bader, X. Jin, F. Yan, A. Kamal, T. J. Gudmundsen, P. J. Leek, T. P. Orlando, W. D. Oliver, and S. Gustavsson, Phys. Rev. Lett. **114**, 010501 (2015).
[40] S. Balakrishnan and R. Sankaranarayanan, Phys. Rev. A **82**, 034301 (2010).
[41] Y. Makhlin, Quantum Information Processing **1**, 243 (2002).
[42] P. Watts, J. Vala, M. M. Müller, T. Calarco, K. B. Whaley, D. M. Reich, M. H. Goerz, and C. P. Koch, Phys. Rev. A **91**, 062306 (2015).
[43] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. A **67**, 042313 (2003).
[44] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM J. Sci. Comput. **16**, 1190 (1995).
[45] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," `http://www.scipy.org/`.
[46] M. H. Goerz, K. B. Whaley, and C. P. Koch, EPJ Quantum Technology **2**, 21 (2015).
[47] N. Khaneja, R. Brockett, and S. J. Glaser, Phys. Rev. A **63**, 032308 (2001).
[48] B. Kraus and J. I. Cirac, Phys. Rev. A **63**, 062309 (2001).
[49] J. M. Nichol, L. A. Orona, S. P. Harvey, S. Fallahi, G. C. Gardner, M. J. Manfra, and A. Yacoby, NPJ Quantum Information **3**, 3 (2017).
[50] "https://github.com/armansetser/leakage-pulse-sequence-solutions,".
[51] D. C. McKay, S. Filipp, A. Mezzacapo, E. Magesan, J. M. Chow, and J. M. Gambetta, Phys. Rev. Applied **6**, 064007 (2016).
[52] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **96**, 022330 (2017).