Fidelity-enhanced variational quantum optimal control

R. J. P. T. de Keijzer,[1,2,*] L. Y. Visser,[2,3] O. Tse,[2,3] and S. J. J. M. F. Kokkelmans[1,2]
[1]*Department of Applied Physics, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*
[2]*Eindhoven Hendrik Casimir Institute, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*
[3]*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

(Received 31 January 2025; accepted 16 May 2025; published 28 May 2025)

Creating robust quantum operations is a major challenge in the current noisy intermediate-scale quantum computing era. Recently, the importance of noise-resilient control methods has become more pronounced in the field. Ordinarily, noisy quantum systems are described by the Lindblad equation. However, minimizing noise susceptibility using this equation has proven challenging because of its irreversibility. In this study, we propose a method for creating robust pulses based on the stochastic Schrödinger equation. This equation describes individual noise realizations under any colored noise process, contrary to the Lindblad equation, which describes mean system behavior under white noise. Using stochastic optimal control techniques, our method, fidelity-enhanced variational quantum optimal control, is able to construct higher-fidelity paths than its nonstochastic counterpart. By accounting for both environmental noise sources as well as noise sources inherent to the control system, highly significant increases in fidelity are noted for both single- and multiple-qubit state preparations.

DOI: 10.1103/PhysRevA.111.052625

I. INTRODUCTION

In the noisy intermediate-scale quantum (NISQ) era, quantum computers operate with a limited number of qubits that are highly susceptible to noise [1], stemming from sources like stimulated emission, measurement errors, or control noise [2–4]. To advance beyond this stage in terms of noise resilience and scalability, significant improvements in both hardware and control software are essential. One promising approach on the software level is to create quantum operations that are least susceptible to noise by refining the control function, or pulse, that determines the trajectory of qubit states within the Hilbert space. Pulse construction algorithms aim to discover an optimal control pulse that transitions a qubit from an initial state to a (possibly unknown) target state. This field of study has gained research attention in recent years [5–7] with a focus on error mitigation, that is, minimizing noise and errors affecting the final state. One of the simplest strategies is to construct time-optimal pulses that reduce the exposure of the system to noise by shortening the interaction time of the system with its environment, as demonstrated in [8,9].

More sophisticated approaches for noise mitigation have also emerged. For example, dynamic decoupling methods apply nearly instantaneous pulses to counteract static system-environment interactions [10,11]. Rather than directly countering the effects of noise, Refs. [12–16] develop pulses that guide the qubit state along paths that are less susceptible to noise. These approaches introduce time-independent noise and use cost functions based on the fidelity with respect to a predefined gate. Of these approaches, Refs. [12,13] optimize gates using gradient estimation techniques, while Ref. [14] takes a supervised learning approach. Other research, such as Refs. [15,16], employ analytical methods like inversion and perturbation techniques. Additionally, recent work in Refs. [17,18] revolves around designing pulses to maximize fidelity while considering the Lindblad equation, using estimated gradients and a known target state. In this context, Ref. [19] investigates the optimal control of the Lindblad equation via Pontryagin's maximum principle [20] as an alternative.

Instead of just looking at the mean behavior of a quantum system described by the Lindblad equation, we found that examining individual realizations improves error analysis for state preparation. This can be achieved using a stochastic unraveling (or stochastic dilation) of the Lindblad equation with the stochastic Schrödinger equation (SSE) [21–23]. Our focus here is on noise sources driven by classical noise processes originating from the control pulses used in qubit state evolution. This includes noise in laser intensity and frequency, which often limit fidelity [24,25]. Considering only classical noise also simplifies the analysis as it fixes one physically accurate unraveling of the Lindblad equation from many possible options [21], allowing for a clearer understanding of how these common factors influence quantum systems. Additionally, this method enables us to calculate the entire distribution of prepared states, moving beyond just the mean value [26]. This results in novel insights into individual state preparations, allowing for more advanced optimal control techniques designed for stochastic processes [27]. We further consider noise intensities that scale with the amplitude of the control pulse, as would be the case in intensity noise in neutral atom systems [28] or flux noise in superconducting systems [29].

Since the SSE serves as a possible unraveling of the evolution of density matrices, (non-Markovian) Lindblad-type master equations can be derived from it, as shown in Refs. [26,30,31]. In Ref. [32], optimal control techniques

*Contact author: r.j.p.t.d.keijzer@tue.nl

have been applied to create specific spatial wave functions through the SSE. Meanwhile, Ref. [33] rigorously shows the existence of solutions from a mathematical standpoint. Two other works have utilized the SSE for enhancing fidelity in quantum operations. Reference [34] does this with a focus on quantum error correction rather than addressing noise mitigation, whereas Ref. [35] uses path-integral approaches. Regarding robust optimal control methods based on the Lindblad equation, Ref. [19] unravels the Lindblad equation into a potentially nonphysical form and uses arbitrary Lindblad decay operators. In contrast, our stochastic process is based on classical control noise, resulting in physically significant realizations. The work most closely aligned with ours is presented in Ref. [17], where closed systems are analyzed and Hamilton-Jacobi-Bellman equations [36] are used for optimality conditions. Our work differs in the fact that we use analytic gradients derived from stochastic optimal control rather than estimated gradients, demonstrating practical applications through numerical examples.

In this work, we present a pulse construction method called fidelity-enhanced variational quantum optimal control (FVQOC) based on the optimal control of the SSE, which we believe offers four key advantages.

(i) The method is based on analytical gradients, allowing for the creation of high-fidelity pulses for systems with arbitrary numbers of qubits.

(ii) It accommodates classical colored noise [37], enabling more realistic noise profiles on practical control systems based on power spectral densities [38] compared to approaches assuming white noise.

(iii) Fidelity optimization can be combined with unknown state preparations, such as ground-state optimization in variational quantum optimal control (VQOC) [39] and other pulse-based variational quantum algorithms [6,7,40], resulting in improved error rates.

(iv) The algorithm generates pulses tailored to specific experimental setups by incorporating the characteristic noise operators of a system, allowing qubits to avoid noise-prone regions of the Hilbert space.

The layout of this paper is as follows. Section II recalls (deterministic) VQOC, which is indifferent to noise. In Sec. III we detail the SSE and establish a framework that enables optimal control techniques. Section IV provides the pulse optimization scheme for FVQOC. In Sec. V we show initial results for our model on single- and multiple-qubit state and gate preparation, using classical white and colored noise.

II. VARIATIONAL QUANTUM OPTIMAL CONTROL

In VQOC [39], a noiseless state[1] ϕ evolves according to the Schrödinger equation as

$$\frac{\mathrm{d}\phi_t}{\mathrm{d}t} = iH(t)\phi_t, \quad \phi(0) = \phi_0,$$
$$H(t) := \sum_{j\in J} z_t^j H_j, \quad t \geqslant 0, \tag{1}$$

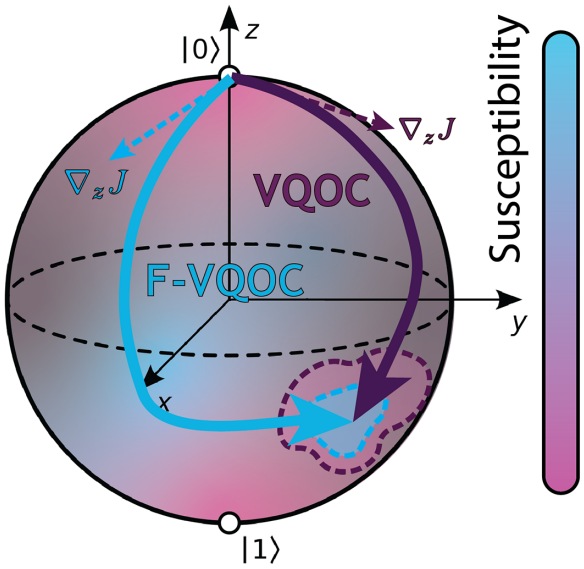

FIG. 1. Illustration of a higher-fidelity path (blue) on the Bloch sphere found by FVQOC, compared to the low-fidelity path (purple) found by standard VQOC. The incorporation of fidelity in the cost function leads to differing gradients $\nabla_z J$ and subsequently lower error outcomes. Shades on the Bloch sphere indicate areas of high or low susceptibility to noise.

where $\mathbf{z} := (z^j)_{j\in J}$ is the vector of control pulses indexed by a finite set J and $H(t)$ is the control Hamiltonian made up of a combination of Hamiltonians H_j, $j \in J$. The goal is to find optimal pulses by minimizing a cost function $J = J_1 + J_2$ with

$$J_1(\mathbf{z}) := \left(\phi_T^{\mathbf{z}}\right)^\dagger H_{\mathrm{targ}} \phi_T^{\mathbf{z}}, \quad J_2(\mathbf{z}) := \frac{\lambda}{2}\|\mathbf{z}\|_2^2, \tag{2}$$

where $\phi_t^{\mathbf{z}}$ is the unique solution corresponding to the Schrödinger equation (1) for a specified control $\mathbf{z}$,[2] H_{targ} is a given target Hamiltonian of which the ground state ϕ_{targ} is to be constructed, with ground-state energy $E_{\mathrm{targ}} = \phi_{\mathrm{targ}}^\dagger H_{\mathrm{targ}} \phi_{\mathrm{targ}}$. The functional J_1 represents the energy of the state at time $t = T$, while the J_2 term penalizes for high-amplitude pulses.

In VQOC [39], the Gâteaux derivative $\nabla_z J$ is calculated via adjoint methods and is determined as

$$\begin{aligned}\nabla J(\mathbf{z})\left(\delta z^j\right) = &-\lambda \int_0^T z_t^j \delta z_t^j \,\mathrm{d}t \\ &- \int_0^T 2i\left(\phi_t^{\mathbf{z}}\right)^\dagger [H_j^\dagger, \Gamma^\dagger(T,t) H_{\mathrm{targ}} \Gamma(T,t)] \\ &\times \phi_t^{\mathbf{z}} \delta z_t^j \,\mathrm{d}t,\end{aligned} \tag{3}$$

where $\Gamma(t,s) := U(t)U^\dagger(s)$, with U the unitary solution operator satisfying $\phi_t^{\mathbf{z}} = U(t)\phi_0$. Following the steepest descent along this gradient yields pulses that will map ϕ_0 to a (local) minimizer of the cost function J via (1).

For every ground-state preparation problem, there are infinitely many possible paths ϕ_t connecting the initial state ϕ_0

[1]For readability purposes, standard bra-ket notation is replaced by daggers to indicate conjugates, as in Refs. [26,30].

[2]For notational purposes, we will write ϕ_t for $\phi_t^{\mathbf{z}}$, where the dependence on $\mathbf{z}$ is implicit.

and the final state ϕ_T. Figure 1 illustrates that some of these paths will be more susceptible to noise than others and thus will lead to worse final fidelity. If $\lambda > 0$, the optimization algorithm will try to minimize the control function amplitude supplied to the system and thus reduce the influence of noise implicitly, as the signal-to-noise ratio of a control function is often constant over the signal strength [41]. This parameter may be increased to ensure that the maximum amplitudes of the generated pulses remain within experimentally feasible limits. We note that other pulse constraints, such as maximum rise and fall times, are inherently dependent on the specific system and therefore fall outside the scope of this investigation. However, the pulses constructed in this study demonstrate no excessive behavior, as detailed in Sec. V. During pulse construction, no direct attention is paid to the specific noise sources in the system. When these optimized pulses are implemented in experimental contexts, noise will continuously introduce errors, which are particularly pertinent in the current NISQ era. This underscores the necessity for a more sophisticated methodology that explicitly considers noise, with the objective of achieving higher-fidelity state preparations.

III. STOCHASTIC SCHRÖDINGER EQUATION

To take into account the system's error sources during pulse construction, a descriptive model for the noise is necessary. These error sources might include spontaneous emission, dephasing [42], or, as analyzed in this paper, errors in the controls, e.g., caused by intensity or frequency noise in the lasers [43]. If one assumes a white-noise profile for the control noise, the mean behavior of the system can be analyzed by the Lindblad equation, describing open quantum system evolution, and pulses that are more robust to these noises can be estimated numerically as in Refs. [17–19]. However, we instead opt to employ the SSE since it is more versatile in the noise profiles and provides information over the entire distribution of states, which can in turn be used for an analytical method to construct higher-fidelity pulses. The SSE for the noisy state ψ is given by

$$\mathrm{d}\psi_t = iH\psi_t \mathrm{d}t - \frac{1}{2}\sum_l S_l^\dagger S_l \psi_t \mathrm{d}[X^l]_t + i\sum_l S_l \psi_t \mathrm{d}X_t^l, \quad (4)$$

where $S := \{S_l\}_{l\in L}$ is a family of Hermitian noise operators indexed by a finite set L and $X_l = (X_t^l)_{t\geqslant 0}$ are noise processes with finite quadratic variation $[X^l]_t = \gamma_l^2 t$, $\gamma_l > 0$ [30]. This could be, for instance, white noise with $X_t^l = W_t^l$, but also Ornstein-Uhlenbeck (OU) noise [44] with

$$\mathrm{d}X_t^l = -k_l X_t^l \mathrm{d}t + \gamma_l \mathrm{d}W_t^l, \quad k_l > 0,$$

or any noise sampled from an arbitrary power spectral density. As an illustrative case, we employ OU noise for this work, as it has a power spectral density that dominates in the lower frequencies, which is most common in real-world sources of control noise [38]. In an ideal scenario, a pulse minimizing the ground-state energy error

$$J_{\mathrm{err}}(\psi) := \mathbb{E}[\psi_T^\dagger H_{\mathrm{targ}} \psi_T] - E_{\mathrm{targ}} \quad (5)$$

can be found. However, the quadratic dependence of this cost function on the stochastic wave function ψ_T makes identifying an analytic expression for the gradient difficult. Instead, we adapt the VQOC problem (2) by adding a so-called fidelity regularizer to the cost function, i.e.,

$$J = J_1 + J_2 + J_3,$$
$$J_3(\mathbf{z}) = -\mu\mathbb{E}\left[\mathsf{F}_T^{\mathbf{z}} + \nu\int_0^T \mathsf{F}_s^{\mathbf{z}}\mathrm{d}s\right], \quad (6)$$

where $\mu, \nu > 0$ and $\mathsf{F}_t^{\mathbf{z}} := |(\phi_t^{\mathbf{z}})^\dagger \psi_t^{\mathbf{z}}|^2$ is the fidelity of the stochastic state ψ with respect to the deterministic state ϕ. The choice of regularization parameter values (λ, μ, ν) is important in the convergence behavior of the algorithm. Properly tuning these parameters directly influences the algorithm's performance, affecting its ability to reach an optimal solution efficiently. If $\nu > 0$, the algorithm maximizes the fidelity over the entire path instead of just the end time. If $\mu = 0$, the algorithm is inattentive to the fidelity and is identical to VQOC. By choosing $\mu > 0$, we ensure that the algorithm prioritizes fidelity conservation, although it will not fully reach the ground state. It is therefore recommended to gradually decrease μ as the algorithm progresses. This approach allows the algorithm to first establish a high-fidelity path and in the subsequent iterations refine its focus towards the ground state. In essence, FVQOC is a fidelity-regularized version of VQOC. As a generale rule, we calculate $\nabla J_1(\mathbf{z}^{(0)})$ and $\nabla J_3(\mathbf{z}^{(0)})$, with $\mathbf{z}^{(0)}$ the initial guess for the pulses, and choose $\mu = \|\nabla_{\mathbf{z}} J_1(\mathbf{z}^{(0)})\|/\|\nabla_{\mathbf{z}} J_3(\mathbf{z}^{(0)})\|$ to ensure both terms contribute equal orders of magnitude to the cost function. Similarly, we set $\lambda = \|\nabla_{\mathbf{z}} J_1(\mathbf{z}^{(0)})\|/\|\nabla_{\mathbf{z}} J_2(\mathbf{z}^{(0)})\|$.

In Sec. IV, we provide the Gâteaux derivative of J_3, with the state evolving according to the stochastic differential equations of (4), for which we employ tools from stochastic optimal control [27]. To this end, a system of evolution equations is constructed for the vector $\boldsymbol{\eta}_t = (\phi_t^\dagger P_0 \psi_t, \phi_t^\dagger P_1 \psi_t, \ldots, \phi_t^\dagger P_{4^N-1}\psi_t)^\top$ as performed in Ref. [26] and described in detail in Appendix A. Here N is the number of qubits and P_i are the 4^N distinct Pauli matrices. Fixing $P_0 = I$ ensures that the fidelity F_t can be equivalently expressed as

$$\mathsf{F}_t = \boldsymbol{\eta}_t^\dagger \Lambda_0 \boldsymbol{\eta}_t, \quad \Lambda_0 = [1,0,0,\ldots,0][1,0,0,\ldots,0]^\dagger.$$

The system of equations for $\boldsymbol{\eta}$ is given by

$$\mathrm{d}\boldsymbol{\eta}_t = g(\boldsymbol{\eta}_t, \mathbf{z}(t))\mathrm{d}t + \sum_l f_l(\boldsymbol{\eta}_t, \mathbf{z}(t))\mathrm{d}X_{l,t}, \quad (7)$$

with

$$g(\boldsymbol{\eta}, \mathbf{z}) := \sum_{j\in J} z^j A_j \boldsymbol{\eta} - \frac{1}{2}\sum_{l\in L} \gamma_l^2 B_l^\dagger B_l \boldsymbol{\eta},$$
$$f_l(\boldsymbol{\eta}, \mathbf{z}) := B_l \boldsymbol{\eta}, \quad l \in L,$$

where the anti-Hermitian matrices A_j and B_j have elements

$$(A_j)_{m,n} = i\mathrm{Tr}(P_m[H_j, P_n]) = -\overline{(A_j)_{n,m}},$$
$$(B_l)_{m,n} = i\mathrm{Tr}(P_m P_n S_l) = -\overline{(B_l)_{n,m}}.$$

From (7), one can show that $\boldsymbol{\eta}_t^\dagger \boldsymbol{\eta}_t = \boldsymbol{\eta}_0^\dagger \boldsymbol{\eta}_0 = 2$ is a conserved quantity.

IV. FIDELITY-ENHANCED VARIATIONAL QUANTUM OPTIMAL CONTROL

Fidelity-enhanced variational quantum optimal control aims to minimize the cost function presented in (6) by employing an analytic gradient descent approach. To achieve this, we need to determine the Gâteaux derivative of the cost function. For the J_1 and J_2 terms outlined in (6), the Gâteaux derivative has been computed in Ref. [39], as given in (3). For the fidelity term J_3, techniques derived from stochastic optimal control are required to compute the Gâteaux derivative. The calculations follow the general approach established in Ref. [27] and are detailed in Appendix B. The Gâteaux derivative of J_3 takes the form

$$\nabla J_3(\mathbf{z})(\delta z_j) = \mathbb{E}\left[\int_0^T K_{z_j}(\boldsymbol{\eta}_t, \mathbf{z}_t, p_t, r_t)\delta z_j(t)\mathrm{d}t\right],$$

where

$$K(\boldsymbol{\eta}, z, p, r) = \Lambda_0\boldsymbol{\eta}_t + p^\dagger g(\boldsymbol{\eta}, z) + \sum_l r_l^\dagger f_l(\boldsymbol{\eta}, z).$$

Here p and r are variables of backward stochastic differential equations, which are known to be difficult to solve [45], both analytically and numerically [46]. We will examine two cases that are of physical relevance, where the Gâteaux derivative can be explicitly expressed using only forward stochastic differential equations, as elaborated in Appendix B. These are fixed noise, described in Sec. IV A, and scaled noise, as in Sec. IV B. Finally, in Sec. IV C we will address the challenge of optimizing the fidelity over entire unitaries $I \to U_{\text{targ}}$ instead of a single state preparation $\phi_0 \to \phi_{\text{targ}}$.

A. Fixed noise

When the noise strength $\gamma > 0$ is independent of the control functions, as is typical in scenarios involving external noise sources such as auxiliary electric fields [12,47], spontaneous decays, or stimulated emission, the Gâteaux derivative can be expressed as

$$\nabla J_3(\mathbf{z})(\delta z_j) = \mathbb{E}\left[\int_0^T \zeta_t\Psi_t A_j\boldsymbol{\eta}_t\delta z_j(t)\mathrm{d}t\right],$$

$$\zeta_t := \boldsymbol{\eta}_T^\dagger\Lambda_0\Phi_T + \nu\int_t^T \boldsymbol{\eta}_s^\dagger\Lambda_0\Phi_s\mathrm{d}s,$$

where Φ is the solution operator of the stochastic system (7), $\boldsymbol{\eta}_t = \Phi_t x_0$, and $\Psi = \Phi^{-1}$. Both Φ and Ψ follow forward stochastic differential equations and can thus be solved numerically using Monte Carlo simulation (see Appendix C), allowing for accurate and efficient approximation of the analytic expression for the gradient.

B. Scaled noise

In realistic quantum systems, the noise associated with control pulses typically scales, most often linearly, with the amplitude of those pulses, e.g., amplitude and frequency noise. This scaling can be factored into the noise model (7) by considering noise processes $\mathrm{d}Y_{l,t} = \gamma_l\sqrt{|z_{c(l)}|}\mathrm{d}X_{l,t}$, which keeps the signal-to-noise ratio $\mathrm{d}[Y]_{l,t}/|z_{c(l)}|$ fixed. Here c maps the pulse $z_j(t)$ to which the noise profile $X_{l,t}$ scales, and

ALGORITHM 1. FVQOC for fixed noise.

input: $z^{(0)}$, H_{targ}, ϕ_0, #iter, #trials
output: $z^{(\#\text{iter})}$, $\mathbb{E}_\omega[\psi_{\omega,T}^\dagger H_{\text{targ}}\psi_{\omega,T}]$
// Pulse Optimization Procedure
for $k = 0$ **to** #iterations **do**
  calculate ∇J_1, ∇J_2;
  for $\omega = 0$ **to** #trials **do**
    generate $X_{l,\omega}$;
    calculate $\boldsymbol{\eta}_\omega$, Φ_ω, Ψ_ω;
    $\nabla J_3^\omega = \zeta_\omega\Psi_\omega A_j\boldsymbol{\eta}_\omega$;
  end
  $z_j^{(k+1)} = z_j^{(k)} - \alpha_k\left(\nabla J_1 + \nabla J_2 + \mathbb{E}\left[\nabla J_3^\omega\right]\right)$;
end
// Final Energy Determination
for $\omega = 0$ **to** #trials **do**
  generate $X_{\omega,l,t}$;
  calculate $U_\omega(z^{(\#\text{iter})})$;
end
return $z^{(\#\text{iter})}$, $\mathbb{E}_\omega[\phi_0^\dagger U_\omega^\dagger(T)H_{\text{targ}}U_\omega(T)\phi_0]$;

$\gamma_l > 0$ serves as the base noise level. The scaled noise SSE takes the form

$$\mathrm{d}\boldsymbol{\eta}_t = \sum_{j\in J} z_j(t)A_j\boldsymbol{\eta}_t\mathrm{d}t - \frac{1}{2}\sum_l \gamma_l^2|z_{c(l)}(t)|B_l^\dagger B_l\boldsymbol{\eta}_t\mathrm{d}[X]_l + \sum_l \gamma_l\sqrt{|z_{c(l)}(t)|}B_l\boldsymbol{\eta}_t\mathrm{d}X_{l,t}.$$

Assuming that $X_{l,t}$ is generated[3] by a white-noise process $W_{l,t}$, the derivative now takes the form

$$\nabla J_3(\mathbf{z})(\delta z_j) = \mathbb{E}\left[\int_0^T \zeta_t\Psi_t\left(A_j\boldsymbol{\eta}_t\mathrm{d}t + \sum_{l|c(l)=j}\frac{1}{2}\frac{\gamma_l}{|z_j(t)|^{1/2}}B_l\boldsymbol{\eta}_t dW_{l,t}\right)\delta z_j(t)\mathrm{d}t\right].$$

For both fixed and scaled noise, the Gâteaux derivatives are determined by first generating noise realizations $X_{l,t}(\omega)$ and subsequently calculating $x(\omega)$, $\Psi(\omega)$, and $\Phi(\omega)$ for each of these realizations. We then obtain the gradients by averaging these results. The pseudocode in Algorithm 1 shows an implementation of FVQOC for fixed noise, which adapts easily to the scaled noise version.

C. Gate optimization

Alongside state preparation (transitioning ϕ_0 to ϕ_{targ}), a key challenge is the construction of a complete unitary transformation U_{targ}. At first sight, one may not expect VQOC to benefit significantly from fidelity optimization because of the following naive intuition: Rotating one state out of a noisy region would merely rotate another into that same noisy region. Yet we will study the problem as it reveals surprising and insightful conclusions.

[3]Note that this does not imply $X_{l,t} = W_{l,t}$; instead, $X_{l,t}$'s only random source is $W_{l,t}$, e.g., Ornstein-Uhlenbeck noise [22].

The evolution of both the noisy unitary U and the noiseless unitary V is characterized by

$$\mathrm{d}U_t = iHU_t\mathrm{d}t - \frac{1}{2}\sum_l S_l^\dagger S_l U_t \mathrm{d}[X]_{l,t} + i\sum_l S_l U_t \mathrm{d}X_{l,t},$$

$$\mathrm{d}V_t = iHV_t\mathrm{d}t, \quad U(0) = V(0) = I.$$

For this problem, the ground-state energy reads

$$J_1(V) = -|\mathrm{Tr}(U_{\mathrm{targ}}^\dagger V_T)|^2.$$

The fidelity equivalent for gates F_t is given by an integral over all possible initial states ϕ_0 distributed according to the Haar measure μ_{Haar} [48]:

$$\begin{aligned} J_3(U,V) &= \mathbb{E}\left[\int \phi_0^\dagger V_T^\dagger U_T \phi_0 \mu_{\mathrm{Haar}}(\mathrm{d}\phi_0)\right] \\ &= \mathbb{E}\left[\left\langle 0\middle|\int_{\bar{U}} \bar{U}^\dagger V_T^\dagger U_T \bar{U}\,\mu_{\mathrm{Haar}}(\mathrm{d}\bar{U})\middle|0\right\rangle\right] \\ &= \frac{1}{2^N}\mathbb{E}[\mathrm{Tr}(V_T^\dagger U_T)], \end{aligned}$$

where the last equality holds by the averaging property of the Haar measure, stating that, for any operator O,

$$\int \bar{U}^\dagger O \bar{U}\,\mu_{\mathrm{Haar}}(\mathrm{d}\bar{U}) = \frac{1}{2^N}\mathrm{Tr}(O)I.$$

The operator $Q_t := V_t^\dagger U_t$ evolves according to

$$\mathrm{d}Q_t = i[H, Q_t]\mathrm{d}t - \frac{1}{2}\sum_l \gamma^2 S_l^\dagger S_l Q_t \mathrm{d}t + i\sum_l S_l Q_t \mathrm{d}X_{l,t},$$

with initial data $Q(0) = I$. The continuous-time case ($\nu > 0$) may be considered in a similar fashion.

The minimization for J_3, in this case, can be done using stochastic optimal control with

$$\boldsymbol{\eta}_t = \{\mathrm{Tr}(P_0 Q_t), \mathrm{Tr}(P_1 Q_t), \ldots, \mathrm{Tr}(P_{4^N-1} Q_t)\}.$$

Noticeably, this system evolves exactly as (7), but with $\boldsymbol{\eta}_0 = \{1, 0, 0, \ldots, 0\}$, and thus acts as if the initial states $\boldsymbol{\eta}_0 = \{1, \eta_{01}, \ldots, \eta_{0N}\}$ in (7), where $\{\eta_{01}, \ldots, \eta_{0N}\}$ are points on the $(4^N - 1)$-dimensional sphere, are averaged out in the full gate optimization case.

Furthermore, considering white noise $X_{l,t} = \gamma_l W_{l,t}$ and Pauli-type noise operators $S_l^\dagger S_l = I$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\mathrm{Tr}(Q_t)] = -\frac{1}{2}\gamma^2\mathbb{E}[\mathrm{Tr}(S^\dagger S Q_t)] = -\frac{1}{2}\gamma^2\mathbb{E}[\mathrm{Tr}(Q_t)],$$

indicating that the fidelity term $J_3(U, V)$ remains independent of the choice of pulses, as hypothesized at the beginning of this section. Nevertheless, when $S^\dagger S \neq I$ or in scenarios involving noise sources beyond white noise, one could expect to find a nonzero gradient of J_3.

V. RESULTS

In this section, we present several examples to illustrate the performance of FVQOC and provide a comparison with its deterministic counterpart, VQOC [39]. In all cases, we will compare FVQOC both with $\nu > 0$ (for the continuous-time cost) and with $\nu = 0$ (for the end-time cost) to VQOC ($\mu = 0$, no SSE). Throughout this section, we incorporate relatively high noise strengths, characterized by signal-to-noise ratios on the order of 0.01, which is high but not unreasonable for NISQ devices [28,49,50]. We do this to distinguish errors induced by the SSE from those arising from the pulse construction algorithm, which might struggle to find the exact minimizing pulse for the ground-state problem. Furthermore, it showcases the effect that FVQOC can have on paths taken through Hilbert space. All averages are taken over $N = 200$ runs. For the single-qubit cases, convergence is typically observed after 20 pulse optimization iterations. Therefore, the maximum number of iterations is fixed at this number. As a figure of merit, we will generally consider the relative increase in error $[J_{\mathrm{err}}(\psi)_{\mathrm{FVQOC}} - J_{\mathrm{err}}(\psi)_{\mathrm{VQOC}}]/J_{\mathrm{err}}(\psi)_{\mathrm{VQOC}}$ as in (5).

A. Fixed noise

In the case of fixed noise, we analyze a one-qubit problem with the target Hamiltonian defined as $H_{\mathrm{targ}} = -\sigma_Y$. Hence, the goal is to find control pulses that map the initial state $\phi_0 = |0\rangle$ to the ground state $\phi_g = |+\rangle$. We assume full control over the Bloch sphere, i.e., we can apply $H_i = \sigma_i$ for $i \in \{X, Y, Z\}$. The control Hamiltonians are subjected to OU control noise, characterized by $S_i = H_i$, with values $(\gamma_X, \gamma_Y, \gamma_Z) = (0.07, 0.01, 0.01)$ and $k_i = 0.1$. To optimize the control, we implement the FVQOC with the parameters $(\lambda, \mu, \nu) = (0.1, 250, 1)$ for the continuous cost function, while setting either $\nu = 0$ or $\mu = 0$ for the end cost and the VQOC, respectively. After ten iterations, we set $\mu = 0$ in all cases, aiming to home in on achieving a high-fidelity path toward the approximate ground state.

The energies and control paths identified in this problem are illustrated in Fig. 2. One observes that the end-time cost demonstrates both higher fidelity and lower-energy error compared to VQOC, with the continuous-time cost performing even better on both metrics. As anticipated, both methods converge towards the ground state after ten iterations when $\mu = 0$. The promising performance can be attributed to the paths found on the Bloch sphere, as shown in Fig. 2(b). Both FVQOC methods effectively guide the qubit state towards the σ_X eigenstates, where it is least susceptible to noise due to the condition $\gamma_X > \gamma_Y, \gamma_Z$. In the final segment of the trajectory, the state transitions to the target Hamiltonian $H_{\mathrm{targ}} = -\sigma_Y$ ground state. Figure 3 shows the calculated pulses for FVQOC and VQOC. For VQOC, there is no preference for the σ_X and σ_Y directions, resulting in uniform pulses. In contrast, FVQOC produces more intricate pulses due to fidelity regularization. Nevertheless, these pulses are continuous and, with adequate experimental control, should be realizable on contemporary quantum systems.

In this specific example, we see that the choice of the target Hamiltonian and noise strengths leads to FVQOC obtaining an intuitively higher-fidelity path than VQOC. However, it is important to note that FVQOC consistently outperforms VQOC, even in cases where the optimal control pulse is less apparent. To illustrate this point, we perform FVQOC simulations using randomized one-qubit target Hamiltonians paired with uniform random noise strengths $\gamma_i \in [0, 0.1]$. Instead of a random target state and setting $H_{\mathrm{targ}} = I - |\phi_{\mathrm{targ}}\rangle\langle\phi_{\mathrm{targ}}|$, we choose to randomize the entire Hamiltonian H_{targ}. This Hamiltonian has a broader spectrum, making the ground state harder

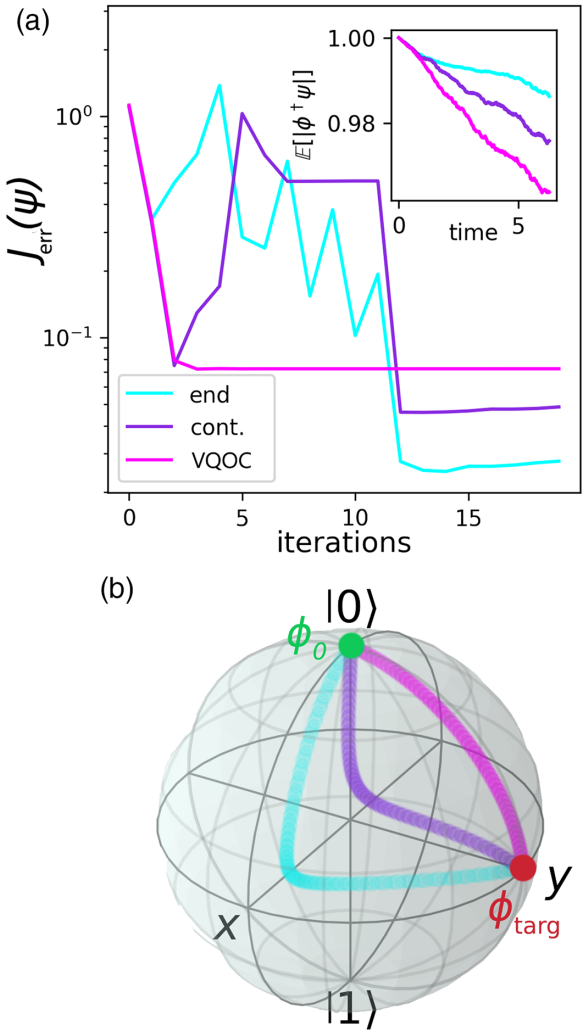

FIG. 2. (a) Empirical average energy found per iteration for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 250, 1)$, $H_i = \sigma_i$, $S_i = H_i$ with OU noise with $(\gamma_X, \gamma_Y, \gamma_Z) = (0.07, 0.01, 0.01)$, and $k_i = 0.1$, for $i \in \{X, Y, Z\}$. (b) Paths found on the Bloch sphere from $\phi_0 = |0\rangle$ to $\phi_g = |+\rangle$ by the various methods.

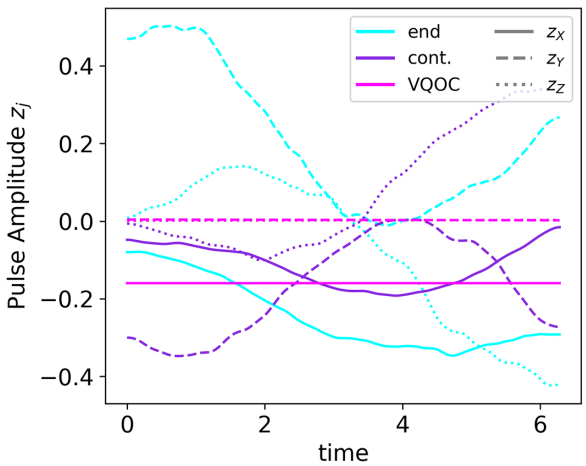

FIG. 3. Pulses generated for the example of Fig. 2. Here $z_j = \sigma_j$ is shown for FVQOC with continuous and end times, as well as for VQOC.

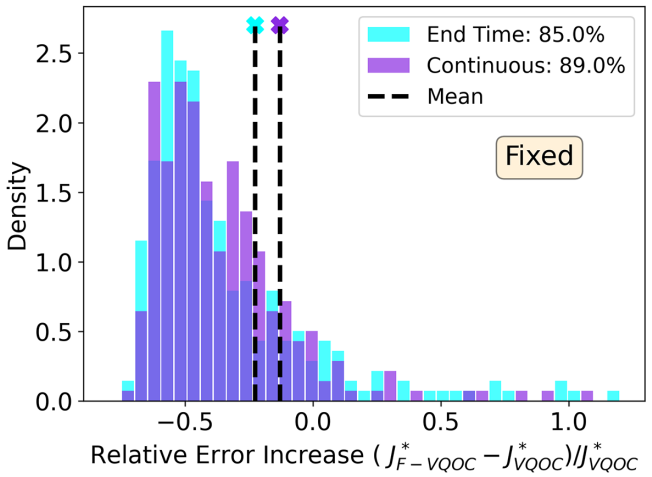

FIG. 4. Distribution of relative ground-state energy error increase for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 250, 1)$ for random one-qubit Hamiltonians H_{targ}, $H_i = S_i = \sigma_i$ with OU noise, $(\gamma_X, \gamma_Y, \gamma_Z) \sim \text{Unif}[0, 0.1]^3$, and $k_i = 0.1$, $i \in \{X, Y, Z\}$. The legend indicates the percentage of FVQOC trials outperforming VQOC counterparts. Mean lines indicate mean relative error increase.

to find [51], showcasing the strength of our method. Figure 4 displays the distribution of relative error between FVQOC and VQOC for the end- and continuous-time costs for 300 of these random problem initializations. We observe similar performance improvements both for end- and continuous-time costs, achieving an average reduction in error by 20% and a decrease in error in 87% of the cases examined. Notably, the relative error increases for FVQOC seem to mostly happen when the absolute fidelity of VQOC is already high.

It is worth mentioning that these results could likely be further improved by fine-tuning the regularization parameters within a fixed experimental setup, where the values of γ_i are kept constant. Overall, these findings demonstrate that the FVQOC method effectively balances the minimization of ground-state energy while conserving fidelity, thereby enabling the identification of lower error paths through Hilbert space.

B. Scaled noise

For scaled noise, the experiment proceeds similarly, with modifications only to the regularization parameters set to $(\lambda, \mu, \nu) = (0.1, 60, 1)$, and $\mu = 0$ is set only after 15 iterations, as this leads to better convergence performance. Furthermore, the target Hamiltonian is specified as $H_{\text{targ}} = -|1\rangle\langle 1|$.

Figure 5 shows the results of the algorithm addressing this problem. Here VQOC drives σ_X and σ_Y simultaneously, as this leads to a lower pulse norm $J_2 = \|z\|_2$. Consistent with the findings of the fixed noise case, FVQOC outperforms VQOC on both occasions by navigating a path closer to the noise immune eigenstates of σ_X.

Interestingly, as seen in Fig. 6, while the end-time cost performs similarly to VQOC, the continuous-time cost significantly outperforms it in most of the randomized trials. We hypothesize that this behavior emerges from the scaling of noise with pulse strength, making the direction of the path

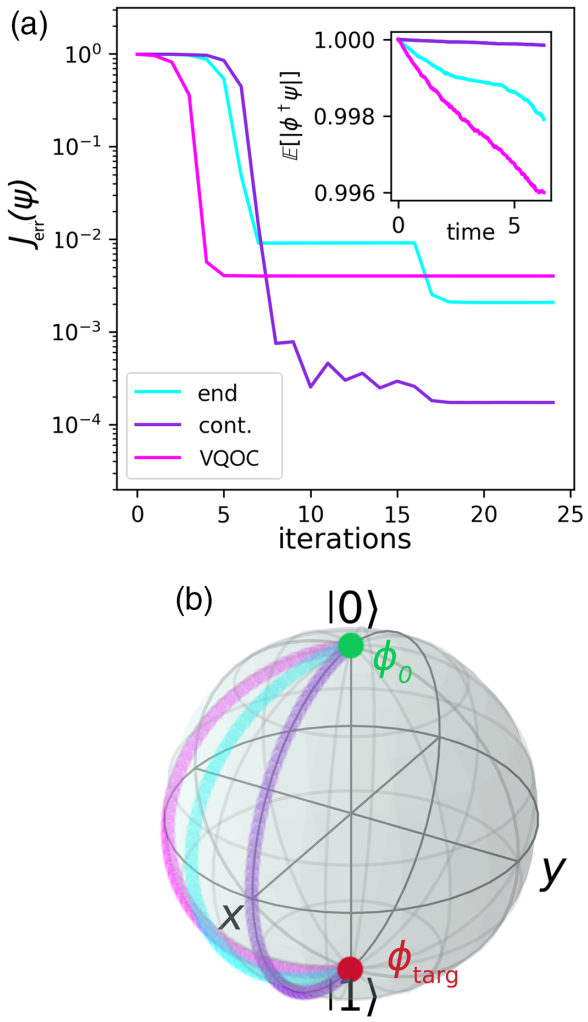

FIG. 5. (a) Average energy found per iteration for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 60, 1)$, $H_i = \sigma_i$, $S_i = H_i$ with OU noise with $(\gamma_X, \gamma_Y, \gamma_Z) = (0.07, 0.01, 0.01)$, and $k_i = 0.1$, for $i \in \{X, Y, Z\}$. (b) Paths found on the Bloch sphere from $\phi_0 = |0\rangle$ to $\phi_g = |1\rangle$ by the various methods.

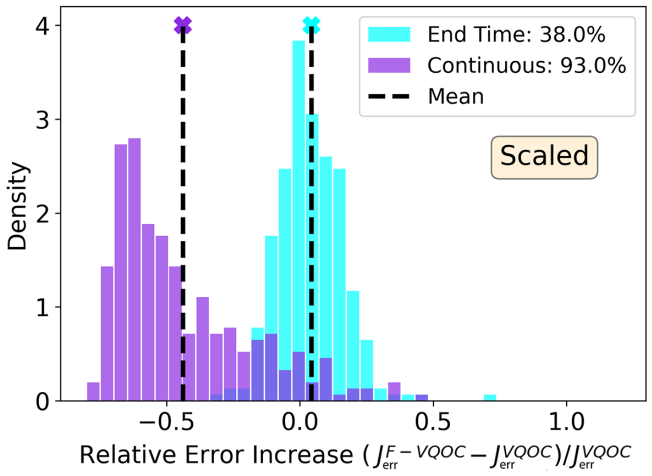

FIG. 6. Relative ground-state energy error increase density plots for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 60, 1)$, $H_i = \sigma_i$, $S_i = H_i$ with OU noise with $(\gamma_X, \gamma_Y, \gamma_Z) \sim \text{Unif}[0, 0.1]^3$, and $k_i = 0.1$, for $i \in \{X, Y, Z\}$ and H_{targ} a random single-qubit Hermitian. The legend indicates what percentage of FVQOC trials outperform their VQOC counterpart. Mean lines indicate mean relative error increase.

critically important, rather than its strength. A continuous cost would then find a path that goes to a noise-insensitive area early in the evolution, contrasting with the delayed responses of an end-time cost. Another reason could be due to bad choices of the regularization parameters (λ, μ, ν). However, variations of these parameters were tested, and initial findings suggest that this is not likely.

C. Gate optimization

We implement the gate construction method outlined in Sec. IV C with the same control Hamiltonians as in the previous experiments. We consider white-noise processes $S_0 = |0\rangle\langle 0|$ and $S_1 = |1\rangle\langle 1|$, with $\gamma_0 = 0.14$ and $\gamma_1 = 0.07$ respectively. To quantify errors, we analyze the mean and variance of the quantity $J_{\text{err}} := 1 - \text{Tr}(U_T^\dagger V_T)$. Note that $S_l^\dagger S_l \neq I$, indicating that the fidelity J_3 may depend on the control Hamiltonian, as shown in Sec. IV C.

Figure 7 illustrates one specific random target unitary demonstrating a significant decrease in both infidelity J and variance VJ over all possible initial states of the gate. Interestingly, we observe that the paths tend to loop around a certain axis of the Bloch sphere, resembling trajectories that stay close to the noiseless eigenstates, as shown in Fig. 2.

However, Fig. 8 shows that when sampling many possible target unitaries, the mean fidelities for VQOC and FVQOC are similar. This is not surprising given the small gradients observed in the fidelity term J_3. Still, we do see different paths being taken through the Hilbert space (see Fig. 7), highlighting the effect of the method, which is notable in the variances of the fidelities. Fidelity-enhanced VQOC shows a preference towards constructing pulses that result in similar fidelities for

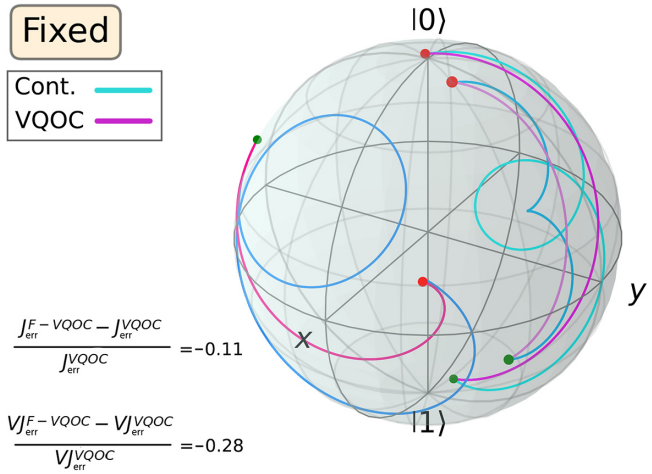

FIG. 7. Example of gate optimization for a single random target unitary U_{targ} done using FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 60, 1)$, $H_i = \sigma_i$, $S_0 = |0\rangle\langle 0|$, and $S_1 = |1\rangle\langle 1|$ with white noise with $(\gamma_0, \gamma_1) = (0.14, 0.07)$. Shown are three random initial states (green) and their paths taken by FVQOC (blue lines) and VQOC (pink lines) to their target states (red).

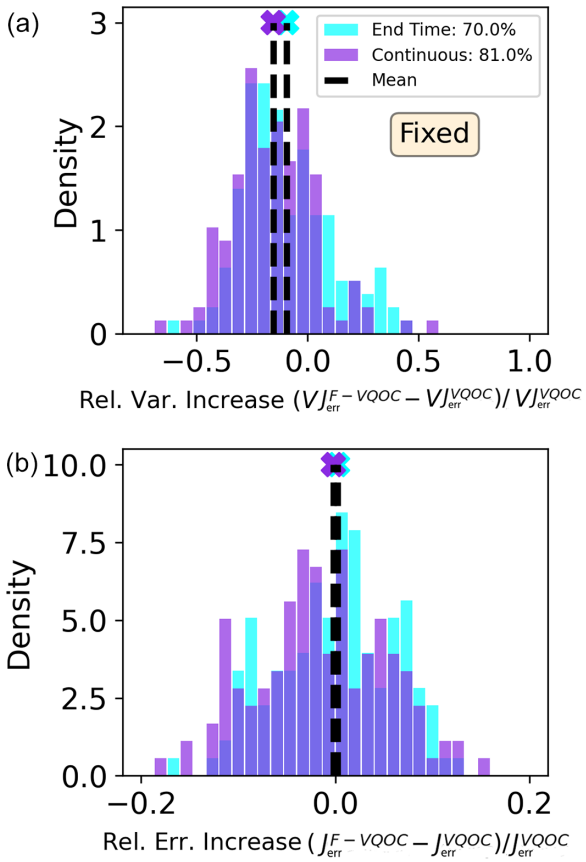

FIG. 8. Relative unitary overlap error increase density plots for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 60, 1)$, $H_i = \sigma_i$, $S_0 = |0\rangle\langle 0|$, and $S_1 = |1\rangle\langle 1|$ with white noise with $(\gamma_0, \gamma_1) = (0.14, 0.07)$ and 300 randomly sampled single-qubit unitaries U_{targ} (one of which is shown in detail in Fig. 7): (a) relative variance increase and (b) relative mean increase. The legend indicates what percentage of FVQOC trials outperform VQOC counterparts. Mean lines indicate mean relative error increase.

each input state, in contrast to having widely varying fidelity levels.

D. Multiple qubits

To demonstrate the effectiveness of our methods for multiqubit state preparation, we focus on generating Greenberger-Horne-Zeilinger (GHZ) states on two qubits, represented as $H_{\text{targ}} = |\text{GHZ}\rangle\langle\text{GHZ}|$. For each run, we use the parameters $(\lambda, \mu, \nu) = (0.1, 600, 1)$. The control Hamiltonian is given by $H = \prod_i \sigma_{Z,i} + \sum_i z_{X,i}\sigma_{X,i} + z_{Z,i}\sigma_{Z,i}$, where index i corresponds to qubit i. For the noise operators, we consider $S_i = \sigma_{X,i}$ with $\gamma = 0.07$, and $S_{ij} = \sigma_{Z,i}\sigma_{Z,j}$ with $\gamma = 0.01$, accounting for noise at both single-qubit and multiqubit levels. We typically find that for two-qubit problems, the convergence of the FVQOC algorithm is reached after 40 iterations. This is more than the single-qubit case, which is not surprising given the additional control functions and the higher dimensionality of the space. This scaling behavior is also reported for VQOC [39].

Figure 9 shows that our method is capable of outperforming VQOC on a noisy multiqubit system. As in the single-qubit case, for general target states, the end-time condition works well only with fixed noise, while the continuous case performs better in both fixed and scaled noise situations. This indicates that our method can effectively create control pulses for preparing important multiqubit states, such as the GHZ state.

VI. DISCUSSION

This work introduces a versatile pulse optimization strategy based on the stochastic Schrödinger equation. It offers insights into the fidelity landscape by utilizing the full distribution of states. The information this gives on individual realizations proves vital in developing noise-informed control methods. Our gradient-based method, derived from stochastic optimal control, constructs high-fidelity pulses for a wide

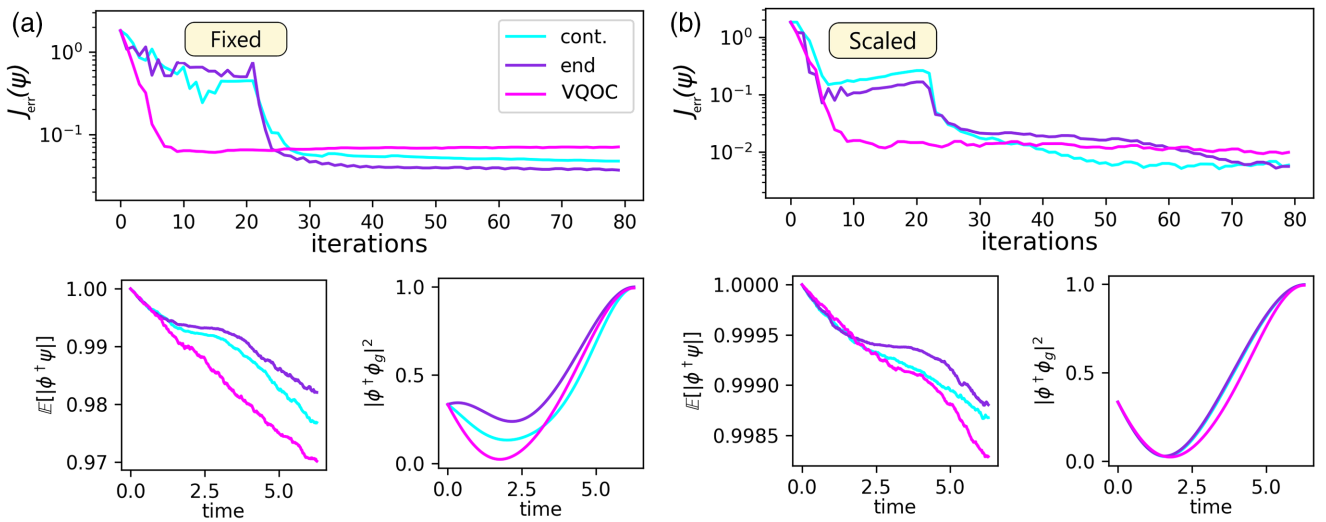

FIG. 9. Average energy found per iteration for FVQOC vs VQOC with $(\lambda, \mu, \nu) = (0.1, 600, 1)$, $H = \sum_i z_{X,i}\sigma_{X,i} + z_{Z,i}\sigma_{Z,i}$, $S_i = \sigma_{X,i}$, and $S_{ij} = \sigma_{Z,i}\sigma_{Z,j}$ with white noise with $\gamma_X = 0.07$ and $\gamma_{ZZ} = 0.01$: (a) fixed noise strength and (b) scaled noise strength. The lower plots show fidelities plus overlap between the target-state GHZ state and the noiseless state, indicating the path taken through state space.

range of problems affected by various colored-noise processes. This approach significantly reduces errors compared to noise-ignorant methods such as VQOC, particularly in ground-state preparation tasks. A key advantage of this algorithm is its ability to tailor pulses to the specific noise characteristics of a system. By inputting the noise operators and their strengths, the algorithm generates pulses that guide qubits through the least noise-sensitive areas of the Hilbert space.

In future research, our objective is to establish clearer relations between the optimal regularization constants (λ, μ, ν) and the system parameters. We also plan to expand our investigation to include arbitrary noise scaling. Additionally, we seek to gather more evidence supporting our hypothesis that continuous-time costs outperform end-time costs in cases of linear noise scaling. We intend to experimentally validate our methods by comparing the fidelities of VQOC with FVQOC pulses on a physical quantum computer. This could be achieved in an early-stage NISQ device by introducing artificial noise profiles to enhance the observed effects.

ACKNOWLEDGMENTS

We thank Jasper Postema, Raul F. Santos, Jasper van de Kraats, Madhav Mohan, and Emre Akaturk for fruitful discussions. This research was financially supported by the Dutch Ministry of Economic Affairs and Climate Policy (EZK), as part of the Quantum Delta NL program, the Horizon Europe program HORIZON-CL4-2021-DIGITAL-EMERGING-01-30 via Project No. 101070144 (EuRyQa), and by the Netherlands Organisation for Scientific Research (NWO) under Grants No. 680.92.18.05 and No. NGF.1582.22.009.

The authors declare no competing interests.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request. All code used is available publicly from [52].

APPENDIX A: MATRIX TRANSFORMATIONS

Our quantity of interest is the fidelity $\mathsf{F} := |\phi^\dagger \psi|^2$, as the overlap between ϕ (the desired state without noise) and the noisy state ψ evolving according to (4). To derive an explicit formula for F, a system of real-valued stochastic differential equations for a vector $\boldsymbol{\eta} \in \mathbb{C}^m$, $m \geqslant 1$, is derived where $\boldsymbol{\eta}_t = [\phi_t^\dagger P_0 \psi_t, \phi_t^\dagger P_1 \psi_t, \ldots, \phi_t^\dagger P_{4^N-1} \psi_t]$. Here N is the number of qubits and P_i is one of the 4^N individual Pauli matrices. Fixing $P_0 = I^{\otimes N}$ ensures $\boldsymbol{\eta}_t^\dagger \Lambda_0 \boldsymbol{\eta}_t = F_t$, with $\Lambda_0 = [1, 0, 0, \ldots, 0][1, 0, 0, \ldots, 0]^\dagger$. The system of equations for $\boldsymbol{\eta}$ is given by

$$\begin{aligned} d\boldsymbol{\eta}_t &= \sum_j z_j(t) A_j \boldsymbol{\eta}_t dt - \frac{1}{2} \sum_l \gamma_l^2 B_l^\dagger B_l \boldsymbol{\eta}_t dt + \sum_l \gamma_l B_l \boldsymbol{\eta}_t dX_{l,t} \\ &=: b(t, \boldsymbol{\eta}, z) dt + \sum_l \sigma_l(t, \boldsymbol{\eta}, z) dX_{l,t}, \end{aligned}$$

where the anti-Hermitian matrices A_j and B_j have elements

$$\begin{aligned} (A_j)_{m,n} &= i\mathrm{Tr}(P_m[H_j, P_n]) = -\overline{(A_j)_{n,m}}, \\ (B_l)_{m,n} &= i\mathrm{Tr}(P_m P_n S_l) = -\overline{(B_l)_{n,m}}. \end{aligned}$$

From (7) one can show that $\boldsymbol{\eta}_t^\dagger \boldsymbol{\eta}_t = \boldsymbol{\eta}_0^\dagger \boldsymbol{\eta}_0 = 2$ is a conserved quantity. As $x_{0,0} = 1$, the rest of the initial state can be seen as a point on the $(4^N - 1)$-dimensional sphere. In certain cases, it is easier to consider the variable $Q := \boldsymbol{\eta}\boldsymbol{\eta}^\dagger$, which evolves according to

$$dQ = \sum_j z_j(t)(A_j Q - Q A_j^\dagger) dt + \frac{1}{2} \sum_l \gamma_l^2 (-2B_l Q B_l^\dagger + B_l^\dagger B_l Q + Q B_l^\dagger B_l) dt + \sum_l \gamma_l (B_l Q - Q B_l^\dagger) dX_{l,t},$$

which through the Fock-Liouville isomorphism [53] is equivalent to

$$\mathrm{d}\mathsf{V} = \mathbf{A}\mathsf{V}\mathrm{d}t + \sum_l \mathbf{B}_l \mathsf{V} \mathrm{d}X_{l,t} + \mathbf{a}\mathrm{d}t + \sum_l \mathbf{b}_l \mathrm{d}X_{l,t},$$

where constant entries of V have been incorporated in the inhomogeneous parts $\mathbf{a}$ and $\mathbf{b}$. Note that the first component of V now directly corresponds to the fidelity $\mathsf{F} = |\phi^\dagger \psi|^2$. The size of the system varies depending on the properties of the noise operator S.

APPENDIX B: DERIVATION OF THE GÂTEAUX DERIVATIVE

Here we establish the derivation of the Gâteaux derivative $\nabla_z J_3$ term in (6). We want to find the gradient of

$$J_3(\mathbf{z}) = -\mu \mathbb{E}\left[\boldsymbol{\eta}_t^\dagger \Lambda_0 \boldsymbol{\eta}_t + \nu \int_0^T \boldsymbol{\eta}_s^\dagger \Lambda_0 \boldsymbol{\eta}_s ds \right],$$

where x evolves under the system of controlled stochastic differential equations given by

$$d\boldsymbol{\eta}_t = \sum_j z_j(t) A_j \boldsymbol{\eta}_t \mathrm{d}t - \frac{1}{2}\sum_l \gamma_l^2 |z_{c(l)}(t)|^q B_l^\dagger B_l \boldsymbol{\eta}_t d[X]_{l,t} + \sum_l \gamma_l |z_{c(l)}(t)|^{q/2} B_l \boldsymbol{\eta}_t dX_{l,t}.$$

Setting $q = 0$ or 1, respectively, results in the fixed noise and scaled noise versions. We denote the filtration of all combined noise processes $X_{l,t}$ by $\mathcal{F}_t$ [54]. For this derivation, we combine the approaches in Ref. [27] (Chap. 11) and Ref. [55], which generalizes the results in Ref. [27] from white-noise processes to general colored-noise processes.

According to Ref. [55], the Gâteaux derivative can be expressed as

$$\begin{aligned}
\nabla J(\mathbf{z})(\delta z_j) &= \mathbb{E}\left[\int_0^T K_{z_j}(\boldsymbol{\eta}_t, z(t), p(t), r(t))\delta z_j(t) dt\right] \\
&= \mathbb{E}\left[\int_0^T \left[p^\dagger(t) A_j \boldsymbol{\eta}_t - \sum_{l|c(l)=j}\left(\frac{q}{2}\gamma_l^2 p^\dagger(t) B_l^\dagger B_l \boldsymbol{\eta}_t + \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} x^\dagger(t) B_l^\dagger r_l(t)\right)\right]\delta z_j(t)\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^T p^\dagger(t) A_j \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] + \sum_{l|c(l)=j}\mathbb{E}\left[\int_0^T -\frac{q}{2}\gamma_l^2 p^\dagger(t) B_l^\dagger B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} r_l^\dagger(t) B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] \\
&=: \nabla J_{3,a}(\mathbf{z})(\delta z_j) + \nabla J_{3,b}(\mathbf{z})(\delta z_j) + \nabla J_{3,c}(\mathbf{z})(\delta z_j),
\end{aligned} \tag{B1}$$

where $r_l(t) := \gamma_l |z_{c(l)}(t)|^{q/2} \tilde{r}(t) B_l \boldsymbol{\eta}_t$, and p and $\tilde{r}$ are the unique adjoint processes satisfying

$$\begin{aligned}
-dp(t) &= \left(\sum_j z_j(t) A_j^\dagger p + \sum_l \gamma_l^2 |z_{c(l)}(t)|^q B_l^\dagger \tilde{r}(t) B_l \boldsymbol{\eta}_t + \Lambda_0 \boldsymbol{\eta}_t\right) dt - \tilde{r}(t)\sum_l \gamma_l |z_{c(l)}(t)|^{q/2} B_l \boldsymbol{\eta}_t dX_{l,t} + dN_t \\
&= \left(\sum_j z_j(t) A_j^\dagger p + \sum_l \gamma_l |z_{c(l)}(t)|^{q/2} B_l^\dagger r_l(t) + \Lambda_0 \boldsymbol{\eta}_t\right) dt - \sum_l r_l(t) dX_{l,t} + dN_t, \quad p(T) = \Lambda_0 \boldsymbol{\eta}_t.
\end{aligned}$$

According to Ref. [55], the solution triplet $(p, \tilde{r}, N)$ is uniquely determined. Here N is a mean-zero local martingale orthogonal to all white-noise processes generating the noises $X_{l,t}$. As we assume our noise processes are fully generated by white-noise processes, this can only be true if $N = 0$.

From Ref. [27] we find

$$\begin{aligned}
p(t) &= -\nu \Psi^\dagger(t) \int_0^t \Phi^\dagger(s) \Lambda_0 \boldsymbol{\eta}_s ds + \Psi^\dagger(t) \mathbb{E}[\zeta(0)|\mathcal{F}^t], \\
r_l(t) &= \Psi^\dagger(t) G_l(t) - \gamma_l |z_l(t)|^{q/2} B_l^\dagger p(t), \\
\zeta(t) &= \Phi^\dagger(T) \Lambda_0 \boldsymbol{\eta}_t + \nu \int_t^T \Phi^\dagger(s) \Lambda_0 \boldsymbol{\eta}_s ds,
\end{aligned}$$

where Φ and Ψ are solutions to the forward stochastic differential equations given by

$$\begin{aligned}
d\Phi(t) &= \sum_j z_j(t) A_j \Phi(t)\mathrm{d}t - \frac{1}{2}\sum_l \gamma_l^2 |z_{c(l)}(t)|^q B_l^\dagger B_l \Phi(t)\mathrm{d}t + \sum_l \gamma_l |z_{c(l)}(t)|^{q/2} B_l \Phi(t) dX_{l,t}, \quad \Phi(0) = I, \\
-d\Psi(t) &= \sum_j z_j(t) \Psi(t) A_j \mathrm{d}t + \frac{1}{2}\sum_l \gamma_l^2 |z_{c(l)}(t)|^q \Psi(t) B_l^\dagger B_l \mathrm{d}t + \Psi(t)\sum_l \gamma_l |z_{c(l)}(t)|^{q/2} B_l dX_{l,t}, \quad \Psi(0) = I,
\end{aligned}$$

and G_l are martingales defined using the representation theorem of martingales [56] by

$$\mathbb{E}[\zeta^\dagger(0)|\mathcal{F}^t] - \mathbb{E}[\zeta^\dagger(0)|\mathcal{F}^0] = \sum_l \int_0^t G_l(s) dW_{l,s}.$$

Filling this into (B1) yields

$$\begin{aligned}\nabla_z J_{3,a}(\delta z_j) &= \mathbb{E}\left[\int_0^T p^\dagger(t) A_j \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^T \left(-\nu\Psi^\dagger(t)\int_0^t \Phi^\dagger(s)\Lambda_0 x(s)ds + \Psi^\dagger(t)\mathbb{E}[\zeta^\dagger(0)|\mathcal{F}^t]\right)^\dagger A_j \boldsymbol{\eta}_t \delta z_j(t)\right] \\ &= \mathbb{E}\left[\int_0^T \zeta(t)\Psi(t) A_j \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right]\end{aligned}$$

and

$$\begin{aligned}\nabla_z J_{3,b}(\delta z_j) + \nabla_z J_{3,c}(\delta z_j) &= \sum_{l|c(l)=j} \mathbb{E}\left[\int_0^T -\frac{q}{2}\gamma_l^2 p^\dagger(t) B_l^\dagger B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} r_l^\dagger(t) B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] \\ &= \sum_{l|c(l)=j} \mathbb{E}\left[\int_0^T -\frac{q}{2}\gamma_l^2 p^\dagger(t) B_l^\dagger B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T \frac{q}{2}\gamma_l^2 p^\dagger(t) B_l^\dagger B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] \\ &\quad + \mathbb{E}\left[\int_0^T \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} G_l{}^\dagger(t)\Psi(t) B_l \boldsymbol{\eta}_t \delta z_j(t)\mathrm{d}t\right] \\ &= \mathbb{E}\left[\sum_{l'}\int_0^T G_{l'}^\dagger(t) dW_{l',t} \sum_{l|c(l)=j}\int_0^T \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}}\Psi(t) B_l \boldsymbol{\eta}_t \delta z_j(t) dW_{l,t}\right] \\ &= \mathbb{E}\left[\{\mathbb{E}[\zeta(0)|\mathcal{F}_T] - \mathbb{E}[\zeta(0)|\mathcal{F}_0]\}\sum_{l|c(l)=j}\int_0^T \frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}}\Psi(t) B_l \boldsymbol{\eta}_t \delta z_j(t) dW_{l,t}\right] \\ &= \mathbb{E}\left[\int_0^T \sum_{l|c(l)=j}\zeta(0)\Psi(t)\frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} B_l \boldsymbol{\eta}_t \delta z_j(t) dW_{l,t}\right] \\ &= \mathbb{E}\left[\int_0^T \sum_{l|c(l)=j}\zeta(t)\Psi(t)\frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} B_l \boldsymbol{\eta}_t \delta z_j(t) dW_{l,t}\right].\end{aligned}$$

Combining these terms gives

$$\begin{aligned}\nabla_z J_3(\delta z_j) &= \nabla_z J_{3,a}(\delta z_j) + \nabla_z J_{3,b}(\delta z_j) + \nabla_z J_{3,c}(\delta z_j) \\ &= \mathbb{E}\left[\int_0^T \zeta(t)\Psi(t)\left(A_j\mathrm{d}t + \sum_{l|c(l)=j}\frac{q}{2}\frac{\gamma_l}{|z_j(t)|^{q/2}} B_l dW_{l,t}\right)\boldsymbol{\eta}_t \delta z_j(t)\right].\end{aligned}$$

APPENDIX C: STOCHASTIC INTEGRATION

Numerical calculation of the gradients is performed using stochastic integration. Throughout this work, stochastic differential equations of the form

$$\mathrm{d}\Upsilon = a(\Upsilon)\mathrm{d}t + b(\Upsilon)\mathrm{d}W_t$$

are solved using the explicit (weak) second-order scheme due to Platen [57,58]. This scheme is given by

$$\Upsilon_{n+1} = \Upsilon_n + \tfrac{1}{2}[a(\bar{\Upsilon}) + a(\Upsilon_n)]\Delta t + \tfrac{1}{4}[b(\bar{\Upsilon}^+) + b(\bar{\Upsilon}^-) + 2b(\Upsilon_n)]\mathcal{N}\sqrt{\Delta t} + \tfrac{1}{4}[b(\tilde{\Upsilon}^+) - b(\tilde{\Upsilon}^-)](\mathcal{N}^2 - 1)\sqrt{\Delta t},$$

with supporting values $\bar{\Upsilon} = \Upsilon_n + a(\Upsilon_n)\Delta t + b(\Upsilon_n)\mathcal{N}\sqrt{\Delta t}$ and $\bar{\Upsilon}^\pm = \Upsilon_n + a(\Upsilon_n)\Delta t \pm b(\Upsilon_n)\sqrt{\Delta t}$. Here $\mathcal{N}$ is a sample from a standard normal distribution and $\Delta t > 0$ is a time step. Heuristically, for our types of problems, this scheme leads to better convergence than standard Euler-Maruyama methods [57], most likely due to non-Euclidean nature of the Hilbert space and non-Lipschitz behavior of colored noise [26].

[1] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[2] M. Morgado and S. Whitlock, Quantum simulation and computing with Rydberg-interacting qubits, AVS Quantum Sci. **3**, 023501 (2021).

[3] U. Aseguinolaza, N. Sobrino, G. Sobrino, J. Jornet-Somoza, and J. Borge, Error estimation in current noisy quantum computers, Quantum Inf. Process. **23**, 181 (2024).

[4] A. Pagano, S. Weber, D. Jaschke, T. Pfau, F. Meinert, S. Montangero, and H. P. Büchler, Error budgeting for a controlled-phase gate with strontium-88 Rydberg atoms, Phys. Rev. Res. **4**, 033019 (2022).

[5] A. S. Wiening, J. Bergendahl, V. Leyton-Ortega, and P. Nalbach, Optimizing qubit control pulses for state preparation, arXiv:2409.08204

[6] O. R. Meitei, B. T. Gard, G. S. Barron, D. P. Pappas, S. E. Economou, E. Barnes, and N. J. Mayhall, Gate-free state preparation for fast variational quantum eigensolver simulations, npj Quantum Inf. **7**, 155 (2021).

[7] A. Choquette, A. Di Paolo, P. K. Barkoutsos, D. Sénéchal, I. Tavernelli, and A. Blais, Quantum-optimal-control-inspired ansatz for variational quantum algorithms, Phys. Rev. Res. **3**, 023092 (2021).

[8] S. Jandura and G. Pupillo, Time-optimal two- and three-qubit gates for Rydberg atoms, Quantum **6**, 712 (2022).

[9] E. Dionis and D. Sugny, Time-optimal control of two-level quantum systems by piecewise constant pulses, Phys. Rev. A **107**, 032613 (2023).

[10] A. K. Gupta and J. M. Tembulkar, Dynamic decoupling of secondary systems, Nucl. Eng. Des. **81**, 359 (1984).

[11] N. Ezzell, B. Pokharel, L. Tewala, G. Quiroz, and D. A. Lidar, Dynamical decoupling for superconducting qubits: A performance survey, Phys. Rev. Appl. **20**, 064027 (2023).

[12] M. Mohan, R. de Keijzer, and S. Kokkelmans, Robust control and optimal Rydberg states for neutral atom two-qubit gates, Phys. Rev. Res. **5**, 033052 (2023).

[13] R. L. Kosut, G. Bhole, and H. Rabitz, Robust quantum control: Analysis & synthesis via averaging, arXiv:2208.14193.

[14] Z.-C. Shi, J.-T. Ding, Y.-H. Chen, J. Song, Y. Xia, X. X. Yi, and F. Nori, Supervised learning for robust quantum control in composite-pulse systems, Phys. Rev. Appl. **21**, 044012 (2024).

[15] J.-J. Zhu, X. Laforgue, X. Chen, and S. Guérin, Robust quantum control by smooth quasi-square pulses, J. Phys. B **55**, 194001 (2022).

[16] P. M. Poggi, G. D. Chiara, S. Campbell, and A. Kiely, Universally robust quantum control, Phys. Rev. Lett. **132**, 193801 (2024).

[17] C. A. Weidner, E. A. Reed, J. Monroe, B. Sheller, S. O'Neil, E. Maas, E. A. Jonckheere, F. C. Langbein, and S. Schirmer, Robust quantum control in closed and open systems: Theory and practice, Automatica **172**, 111987 (2025).

[18] A. Aroch, R. Kosloff, and S. Kallush, Mitigating controller noise in quantum gates using optimal control theory, Quantum **8**, 1482 (2024).

[19] C. Lin, D. Sels, Y. Ma, and Y. Wang, Stochastic optimal control formalism for an open quantum system, Phys. Rev. A **102**, 052605 (2020).

[20] R. B. Vinter, in *Encyclopedia of Systems and Control*, edited by J. Baillieul and T. Samad (Springer, London, 2013), pp. 1–9.

[21] M. Moodley and F. Petruccione, Stochastic wave-function unraveling of the generalized Lindblad master equation, Phys. Rev. A **79**, 042103 (2009).

[22] A. Barchielli, C. Pellegrini, and F. Petruccione, Stochastic Schrödinger equations with coloured noise, Europhys. Lett. **91**, 24001 (2010).

[23] C. W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, Springer Series in Synergetics (Springer, Berlin, 2004).

[24] I. S. Madjarov, J. P. Covey, A. L. Shaw, J. Choi, A. Kale, A. Cooper, H. Pichler, V. Schkolnik, J. R. Williams, and M. Endres, High-fidelity entanglement and detection of alkaline-earth Rydberg atoms, Nat. Phys. **16**, 857 (2020).

[25] Q.-L. Hou, H. Wang, and J. Qian, Active robustness against detuning error for Rydberg quantum gates, Phys. Rev. Appl. **22**, 034054 (2024).

[26] R. J. P. T. de Keijzer, L. Y. Visser, O. Tse, and S. J. J. M. F. Kokkelmans, Qubit fidelity distribution under stochastic Schrödinger equations driven by classical noise, Phys. Rev. Res. **7**, 023063 (2025).

[27] A. Bensoussan, *Estimation and Control of Dynamical Systems* (Springer, Berlin, 2018).

[28] M. L. Day, P. J. Low, B. White, R. Islam, and C. Senko, Limits on atomic qubit control from laser noise, npj Quantum Inf. **8**, 72 (2022).

[29] D. A. Rower, L. Ateshian, L. H. Li, M. Hays, D. Bluvstein, L. Ding, B. Kannan, A. Almanakly, J. Braumüller, D. K. Kim *et al.*, Evolution of $1/f$ flux noise in superconducting qubits with weak magnetic fields, Phys. Rev. Lett. **130**, 220602 (2023).

[30] I. Semina, V. Semin, F. Petruccione, and A. Barchielli, Stochastic Schrödinger equations for Markovian and non-Markovian cases, Open Syst. Inform. Dyn. **21**, 1440008 (2014).

[31] H.-P. Breuer and J. Piilo, Stochastic jump processes for non-Markovian quantum dynamics, Europhys. Lett. **85**, 50004 (2009).

[32] J. Cui, S. Liu, and H. Zhou, Optimal control for stochastic nonlinear Schrödinger equation on graph, SIAM J. Control Optim. **61**, 2021 (2023).

[33] B. E. Breckner, H. Lisei, and G. Ionuţ Şimon, Optimal control results for a class of stochastic Schrödinger equations, Appl. Math. Comput. **407**, 126310 (2021).

[34] C. Ahn, A. C. Doherty, and A. J. Landahl, Continuous quantum error correction via quantum feedback control, Phys. Rev. A **65**, 042301 (2002).

[35] A. Villanueva and H. Kappen, Stochastic optimal control of open quantum systems, arXiv:2410.18635.

[36] C. Fromonteil, R. Tricarico, F. Cesa, and H. Pichler, Hamilton-Jacobi-Bellman equations for Rydberg-blockade processes, Phys. Rev. Res. **6**, 033333 (2024).

[37] O. D. Street and D. Crisan, Semi-martingale driven variational principles, Proc. R. Soc. A **477**, 20200957 (2021).

[38] P. Welch, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. **15**, 70 (1967).

[39] R. de Keijzer, O. Tse, and S. Kokkelmans, Pulse based variational quantum optimal control for hybrid quantum computing, Quantum **7**, 908 (2023).

[40] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, J. Magn. Reson. **172**, 296 (2005).

[41] R. M. Howard, *Principles of Random Signal Analysis and Low Noise Design: The Power Spectral Density and Its Applications* (Wiley, New York, 2002), Chap. 4, pp. 92–137.

[42] K. C. Young and K. B. Whaley, Qubits as spectrometers of dephasing noise, Phys. Rev. A **86**, 012314 (2012).

[43] J. Bergli, Y. M. Galperin, and B. Altshuler, Decoherence in qubits due to low-frequency noise, New J. Phys. **11**, 025002 (2009).

[44] R. A. Maller, G. Müller, and A. Szimayer, in *Handbook of Financial Time Series*, edited by T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (Springer, Berlin, 2009), pp. 421–437.

[45] S. Peng, Backward stochastic differential equations and applications to optimal control, Appl. Math. Optim. **27**, 125 (1993).

[46] J. Douglas Jr., J. Ma, and P. Protter, Numerical methods for forward-backward stochastic differential equations, Ann. Appl. Probab. **6**, 940 (1996).

[47] J. P. J. van Dijk, E. Kawakami, R. N. Schouten, M. Veldhorst, L. M. K. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano, Impact of classical control electronics on qubit fidelity, Phys. Rev. Appl. **12**, 044054 (2019).

[48] A. Daele, The Haar measure on finite quantum groups, Proc. Am. Math. Soc. **125**, 3489 (1997).

[49] S. de Léséleuc, D. Barredo, V. Lienhard, A. Browaeys, and T. Lahaye, Analysis of imperfections in the coherent optical excitation of single atoms to Rydberg states, Phys. Rev. A **97**, 053803 (2018).

[50] F. Wudarski, Y. Zhang, A. N. Korotkov, A. G. Petukhov, and M. I. Dykman, Characterizing low-frequency qubit noise, Phys. Rev. Appl. **19**, 064066 (2023).

[51] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Variational quantum algorithms for discovering Hamiltonian spectra, Phys. Rev. A **99**, 062304 (2019).

[52] https://gitlab.tue.nl/s1658271/sse.git.

[53] J. Polonyi and I. Rachid, Elementary open quantum states, Symmetry **13**, 1624 (2021).

[54] C. Dellacherie and P. A. Meyer, *Probabilities and Potential* (Hermann, Paris, 1978), Vol. 1.

[55] J. Song and M. Wang, Stochastic maximum principle for systems driven by local martingales with spatial parameters, Probab. Uncertainty Quant. Risk **6**, 213 (2021).

[56] H. Kunita and S. Watanabe, On square integrable martingales, Nagoya Math. J. **30**, 209 (1967).

[57] M. Bayram, T. Partal, and G. O. Buyukoz, Numerical methods for simulation of stochastic differential equations, Adv. Differential Equ. **2018**, 17 (2018).

[58] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1999).