

Geometric quantum control and the random Schrödinger equation

Rufus Lawrence,^{*} Aleš Wodecki[†], and Johannes Aspman[‡]

Department of Computer Science, *Czech Technical University in Prague, Prague, Czech Republic*

Llorenç Balada Gaggioli[§]

Department of Computer Science, *Czech Technical University in Prague, Prague, Czech Republic*
and *LAAS-CNRS, Université de Toulouse, Toulouse, France*

Jakub Mareček^{||}

Department of Computer Science, *Czech Technical University in Prague, Prague, Czech Republic*



(Received 2 June 2025; accepted 19 December 2025; published 10 February 2026)

Understanding and mitigating noise in quantum systems is a fundamental challenge in achieving scalable and fault-tolerant quantum computation. Error modeling for quantum systems can be formulated in many ways, some of which are very fundamental, but hard to analyze (evolution by general dynamical map) and others perhaps too simplistic to represent physical reality. In this paper, we present an intermediate approach, introducing the random Schrödinger equation, with a noise term given by a time-varying random Hermitian matrix as a means to model noisy quantum systems. We derive bounds on the error of the synthesized unitary in terms of bounds on the norm of the noise, and show that for certain noise processes these bounds are tight. We then show that in certain situations, minimizing the error is equivalent to finding a geodesic on $SU(n)$ with respect to a Riemannian metric encoding the coupling between the control pulse and the noise process, thus connecting our work to the complexity geometry pioneered by Michael Nielsen.

DOI: [10.1103/y8nw-n1wb](https://doi.org/10.1103/y8nw-n1wb)

I. INTRODUCTION

In order to achieve the low gate-error rates necessary for scalable fault-tolerant quantum computation, one needs to mitigate the effect of noise. In the context of quantum optimal control, mitigating the effect of noise on the system requires a model of the noise which faithfully represents the behavior of the quantum system while remaining tractable mathematically and numerically. Some models of open (noisy) quantum systems are hard to analyze, while others are too simplistic to represent physical reality.

The standard approach to modeling open quantum systems is to use the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation:

$$\frac{d\rho}{dt} = [H, \rho] + \sum_j \gamma_j \left(L_j \rho L_j^\dagger - \frac{1}{2} \{L_j L_j^\dagger, \rho\} \right). \quad (1)$$

Here, H denotes the system Hamiltonian, the (non-Hermitian) operators L_j are the *jump operators*, and the real

numbers γ_j denote the *damping rates*. However, it is difficult to apply this approach to a specific quantum system, since in general, it is very difficult to write down the jump operators and damping rates for a given physical system from first principles. Although it is sometimes possible to identify them experimentally (see, for instance, Ref. [1]), this approach also has its limitations.

An alternative approach of representing the state of an open quantum system as a time-dependent random vector in a Hilbert space goes back at least to Belavkin [2]. Subsequently, Gisin and Percival [3] derive the *stochastic Schrödinger equation* as an *unraveling* of the GKSL equation. That is to say, given a GKSL equation for the density operator ρ , the authors construct an Itô-type stochastic differential equation for the state $|\psi\rangle$ such that the expectation $\mathbb{E}[|\psi\rangle\langle\psi|]$ satisfies the original GKSL equation. A Schrödinger equation associated in such a way to the GKSL equation is called a *stochastic Schrödinger equation*. These types of equations have also been studied in the case of non-Markovian open quantum systems (see, e.g., Refs. [4–6]). It should be noted that while the dynamics of the stochastic Schrödinger equation give rise to mixed states after taking an ensemble average over the probability space, the trajectories $|\psi(t, \omega)\rangle$ evolve unitarily in this approach.

The purpose of this paper is twofold. We begin by showing that a large class of open quantum systems, sometimes called *noisy quantum systems* [7–9], can be modeled by a random ordinary differential equation (RODE) which we refer to as the random Schrödinger equation. Just like the stochastic Schrödinger equation, the dynamics of random Schrödinger

^{*}Contact author: lawrejam@fel.cvut.cz

[†]Contact author: wodecki.ales@fel.cvut.cz

[‡]Contact author: johannes.aspman@cvut.cz

[§]Contact author: llorenç.balada.gaggioli@fel.cvut.cz

^{||}Contact author: jakub.marecek@fel.cvut.cz

equation also give rise to mixed states upon taking ensemble averages. We then use this formalism to study robust optimal quantum control.

The random Schrödinger for the time-evolution operator $U_S(t, \cdot)$ is the following RODE:

$$\begin{aligned} \frac{dU_S(t, \cdot)}{dt} &= -i(H_{0,S}(t) + H_{1,S}(t, \cdot))U_S(t, \cdot), \\ U_S(0, \cdot) &= \mathbb{1}. \end{aligned} \quad (2)$$

Here, $H_{0,S}$ is the deterministic (control) Hamiltonian, $H_{1,S}$ is an essentially bounded, Hermitian matrix-valued random process representing the noise, and S denotes that the operators are considered in the Schrödinger picture. Just as for the stochastic Schrödinger equation, the trajectories of solutions evolve unitarily, and mixed states are only obtained after taking ensemble averages. The RODE approach to modeling noisy quantum systems is complementary to the Itô-type SDE modeling introduced in Ref. [3]. In particular, while Itô calculus and the associated SDE yield the appropriate framework for modeling quantum systems subject to “white” Hermitian Gaussian noise, the RODE approach lends itself more readily to modeling systems subject to time-correlated, bounded Hermitian noise processes.

The control theory of open quantum systems presents many challenges, and is not well studied. While Dirr *et al.* [10] have studied the controllability/reachability of density matrices subject to GKSL dynamics in abstract terms using techniques from differential geometry, not much is known about the precision- or fidelity-optimal control problem. The control theory of the stochastic Schrödinger equation has been studied in Refs. [11,12], where the authors derive bounds that are similar to the bounds derived in this paper. Recently, some progress has been made on the robust control problem in Refs. [7,8,13–17], as well as in Ref. [18] in the context of robust quantum annealing. In these works, the noise model is somewhat more ad-hoc, consisting simply of a “generic” error term in the Hamiltonian, and probabilistic questions are not fully addressed. The RODE approach introduced in this paper should be considered as a probabilistic improvement on these models, rather than as a substitute for the GKSL/stochastic Schrödinger approach.

By introducing a coupling between the control $H_{0,S}$ and the noise $H_{1,S}$, we are able to use the random Schrödinger model to place the ideas of Refs. [7,8,18] on a more rigorous mathematical and probabilistic footing. In this paper, we provide a first step in this direction by showing that for certain classes of Hermitian noise processes, it makes sense to model a quantum system as an RODE. Following this, we provide “worst case” bounds that apply to any such noise. Subsequently, we formulate a condition on the noise that makes this bound tight. These bounds allow us to discuss a natural geometric interaction between the control field and the noise, which ultimately allows us to link our results to the seminal work of Nielsen *et al.* [19–22] on the geometry of quantum computation. In particular, we show that for a natural class of couplings between control and noise, the fidelity-optimal control problem reduces to the problem of finding geodesics with respect to an appropriately defined *noise metric*. These noise metrics are analogous to the *complexity metrics* introduced by Nielsen and his coauthors.

Summary of results

In this section, we summarize the main results of the paper, giving physical intuition and sketching proofs, where possible. The full proofs are given in the following sections of the paper.

Our first contribution is a bound on the error induced by the noise term $H_{1,S}$. Let $U_{0,S}$ be the operator that satisfies the noiseless equation

$$\begin{aligned} \frac{dU_{0,S}(t)}{dt} &= -iH_{0,S}(t)U_{0,S}(t), \\ U_{0,S}(0) &= \mathbb{1}. \end{aligned} \quad (3)$$

The following theorem quantifies the difference between the noiseless evolution given by $U_{0,S}$ and the noisy evolution U_S [see Eq. (2)].

Theorem 1. If $\text{ess sup}_{t,\omega} \|H_{1,S}(t, \omega)\| \leq K$, the worst-case error induced by the noise grows linearly in time. That is, $\text{ess sup}_{\omega} \|U_S(t, \omega) - U_{0,S}(t)\| \leq Kt$ with probability 1.

Proof. (Sketch) Let $U_{0,S}(t)$ be the propagator associated with $H_{0,S}$, transform Eq. (2) into the interaction picture via $U_{0,S}$ and integrate. The result then follows by applying standard matrix norm inequalities. ■

The subsequent theorem provides a geometric refinement of the previous theorem.

Theorem 2. If $\Lambda(t) := \text{ess sup}_{\omega} \|H_{1,S}(t, \omega)\|_F$, we have the following inequality:

$$\begin{aligned} d_F(U_I(t, \omega), U_{0,S}(t)) &\leq \int_0^t \|H_{1,S}(s, \omega)\|_F ds \\ &\leq \int_0^t \Lambda(s) ds. \end{aligned} \quad (4)$$

Proof. (Sketch) Similar to the previous proof. ■

Here, $\|\cdot\|_F$ denotes the Frobenius norm and $d_F(\cdot, \cdot)$ denotes the metric (distance function) on $\text{SU}(n)$ induced by the Frobenius norm. We introduce a condition on the noise that makes the bound tight, which we call the *worst-case noise condition* (WCNC; see Definition 5). As the name suggests, the WCNC is a condition that guarantees that the noise is aligned in the worst possible direction with maximal intensity with positive probability. For noise processes satisfying the WCNC the bound in Theorem 2 is tight in a probabilistic sense:

Theorem 3. For noise processes $H_{1,S}$ satisfying the *worst-case noise condition* (see Definition 5), we have that

$$\mathbb{P}\left[\int_0^t \Lambda(s) ds - d_F(U_I(t, \omega), \mathbb{1}) < \epsilon\right] > 0. \quad (5)$$

Proof. (Sketch) In general, we expect the noise to exhibit a degree of self-cancellation (see Remark 6). However, this can fail: If in the interaction picture the noise process is always oriented in the same axis and close to its bound, we show that the solution approaches its bound. The proof is technical and relies on a secondary interaction picture transformation. ■

We then provide a general result that gives conditions on the noise $H_{1,S}(t, \omega)$ under which the *worst-case noise condition* of Definition 5 may be satisfied.

Theorem 4. Let $\mathcal{H}_E$ denote the space of Hermitian matrices with Frobenius norm less than E , and let $H_{1,S}(t, \cdot)$ be a

stochastic process with almost surely continuous sample paths and let $H_{0,S}(t) \in \mathcal{H}_E$ for some $E \in \mathbb{R}$. Suppose further that $H_{1,S}(t, \cdot)$ satisfies the strong Markov property, and that the associated measure on $C^0([0, T], \mathcal{H}_E)$ is positive on cylindrical sets. Then, $H_{1,S}(t, \cdot)$ satisfies the worst-case noise condition with respect to $H_{0,S}$.

Proof. (Sketch) The Markov property combined with the positivity condition imply that for any sufficiently regular path γ in the space of Hermitian matrices, $\mathbb{P}(\|H_{1,S}(t, \omega) - \gamma(t)\| < \epsilon \mid \gamma) > 0$ is positive. The result then follows. ■

This theorem allows us to conclude that many well-known stochastic processes satisfy the WCNC. In particular, if $X(t, \cdot)$ is an unbounded Markov process (e.g., a Gaussian Markov process), and f is a continuous bounded function, then $H_{0,S}(t, \cdot) = f(X(t, \cdot))$ satisfies the WCNC for any $H_0(t)$ whose image lies in $\mathcal{H}_E$. See Remark 9 for more details.

We then introduce the notion of a control-noise coupling, which quantifies the relationship between the control field and the noise. In particular, we consider control-noise couplings induced by a Riemannian noise metric denoted by g_Λ . Noise metrics are analogous to the complexity metrics introduced by Nielsen *et al.* [19–22]. For these couplings, solving the robust control problem (see Definition 4) is equivalent to finding a g_Λ geodesic between $\mathbb{1}$ and V , in a situation analogous to Refs. [19–22].

Theorem 5. Suppose that $H_{1,S}(t, \cdot)$ satisfies the control-dependent WCNC (see Definition 7) with respect to all $H_{0,S}(t) \in \mathcal{H}_E$, and suppose $\Lambda(t) = \sqrt{g_\Lambda(H_{0,S}(t), H_{0,S}(t))}$. Then, solving the robust optimal control problem of synthesizing a unitary target V with respect to the cost function $c(U_S(T), H_{0,S}(\cdot), T, V) = \text{ess sup}_\omega \|U_S(T, \omega) - V\|$ is equivalent to finding a length-minimizing g_Λ geodesic on $\text{SU}(n)$.

Proof. (Sketch) Plugging $\Lambda(t) = \sqrt{g_\Lambda(H_{0,S}(t), H_{0,S}(t))}$ into Theorem 2 we see that the cost function of a given control is bounded by the length of the path traced out by the control in $\text{SU}(n)$, measured with respect to g_Λ . By Theorem 3, minimizing the length of this path (i.e., finding a geodesic) minimizes the cost function. ■

II. RANDOM ORDINARY DIFFERENTIAL EQUATIONS

In this section, we review the basic theory of RODEs, following Refs. [23,24]. See Remarks 5 and 6 for a comparison of RODEs and SDEs.

Definition 1. Let $\eta : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ be an $\mathbb{R}^m$ -valued random process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ be continuous. A random differential equation is an equation of the form

$$\frac{dx(t, \cdot)}{dt} = f(t, x(t, \cdot), \eta(t, \cdot)), \quad x(0, \cdot) = x_0. \quad (6)$$

We say that Eq. (6) is linear if f is linear in the state space (second argument). It follows that any linear RODE can be written as

$$\frac{dx(t, \cdot)}{dt} = A(t, \cdot)x(t, \cdot), \quad x(0, \cdot) = x_0, \quad (7)$$

where $A : [0, T] \times \Omega \rightarrow \text{Mat}_n(\mathbb{R})$ is an $n \times n$ matrix-valued random process.

Remark 1. Our definition of an RODE is more restrictive than the definitions used by other authors, notably Ref. [24]. Other authors define RODEs using the concept of a random function, which allows f to depend directly on $\omega \in \Omega$. But this requires a significantly more technical treatment in the nonlinear case.

There are several different ways to interpret the derivative on the left-hand side of an RODE, each giving rise to a different notion of what it means for a random process $x(t, \omega)$ to be a solution.

Definition 2. We say that a random process $x(t, \omega)$ is a sample path (SP) solution on $[0, T]$ if for almost all $\omega \in \Omega$, $x(\cdot, \omega)$ is absolutely continuous (in t),

$$\frac{\partial x(t, \omega)}{\partial t} = f(t, x(t, \omega), \eta(t, \omega)), \quad (8)$$

for almost all $t \in [0, T]$ and $x(0, \omega) = x_0(\omega)$.

Remark 2. There is a slight subtlety to this definition. We have assumed that any SP solution is a priori a random process, i.e., that the mapping $x(t, \cdot) : \Omega \rightarrow \mathbb{R}^n$ is measurable for all t . It turns out that this is not necessary. If we can solve the SP equation (8) for almost all $\omega \in \Omega$, then the mapping

$$\begin{aligned} x(t, \cdot) : \Omega &\rightarrow \mathbb{R}^n \\ \omega &\mapsto x(t, \omega) \end{aligned}$$

is defined almost everywhere, and moreover is measurable. For a proof, see Sec. 2.1.2 in Ref. [23].

Definition 3. We say that x is a mean-square (L^2) solution if x is an L^2 random process [i.e., if $x(t, \cdot) \in L^2(\Omega)$ for all t], x is strongly absolutely continuous as a function $x : [0, T] \rightarrow L^2(\Omega)$ and

$$\left[\frac{x(t+h, \cdot) - x(t, \cdot)}{h} \right] \xrightarrow{h \rightarrow 0} f(t, x(t, \cdot), \eta(t, \cdot)) \quad \text{in } L^2(\Omega) \quad (9)$$

for almost all $t \in [0, T]$, and $x(0, \cdot) = x_0(\cdot)$ in $L^2(\Omega)$.

If $x(t, \cdot)$ is an L^2 process, we denote by $\hat{x}(t)$ its equivalence class in $L^2(\Omega)$. Clearly, $\hat{x} : [0, T] \rightarrow L^2(\Omega)$.

Due to the absolute continuity assumptions, both the property of being an SP solution and the property of being an L^2 solution have equivalent integral formulations.

In the first case, $x(t, \cdot)$ is an SP solution if and only if

$$x(t, \omega) = x_0(\omega) + \int_0^t f(s, x(s, \omega), \eta(s, \omega)) ds \quad (10)$$

for almost all $\omega \in \Omega$ and for almost all $[0, T]$. The integral above is just the usual Lebesgue integral of the trajectory indexed by ω .

Similarly, an L^2 process $x(t, \cdot)$ is an L^2 solution if and only if

$$x(t, \cdot) = x_0 + \int_0^t f(s, x(s, \cdot), \eta(s, \cdot)) ds \quad (11)$$

for all $t \in T$. Here, the integral denotes the Bochner integral.

The relationship between sample path and L^2 solutions is less obvious. We have the following theorem:

Theorem 6. Suppose $\hat{x}(t)$ is an L^2 solution. Then, there exists a jointly measurable map $x : [0, T] \times \Omega \rightarrow \mathbb{R}^n$, equivalent to $\hat{x}$ in $L^2(\Omega)$, which is an SP solution.

Conversely, suppose $x : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is an SP solution. Then its equivalence class in $L^2(\Omega)$, $\hat{x} : [0, T] \rightarrow L^2(\Omega)$ is an L^2 solution if furthermore

$$\int_a^b \|f(s, \cdot, x(s, \cdot))\|_{L^2} ds < \infty. \quad (12)$$

In light of this theorem, as well as Remark 2, we will from now on move freely between the SP and L^2 notions of solutions wherever it is convenient and allowed.

The main existence theorem for RODEs was proven in Ref. [24] (Theorem 5). We reproduce it here for completeness:

Theorem 7. Let $\tilde{f} : [0, T] \times L^p(\Omega) \rightarrow L^p(\Omega)$ be the mapping induced by f .

$$\|\tilde{f}(t, x) - \tilde{f}(t, y)\|_{L^p} \leq k(t)\|x - y\|_{L^p} \quad (13)$$

with $\int_a^b |k(t)| dt < \infty$, then the RODE

$$\frac{dx(t, \cdot)}{dt} = f(t, x(t, \eta(t, \cdot))), \quad x(0, \cdot) = x_0 \quad (14)$$

has a unique L^p solution.

Corollary 1. The linear RODE

$$\frac{dx(t, \cdot)}{dt} = A(t, \cdot)x(t, \cdot), \quad x(0, \cdot) = x_0(\cdot) \quad (15)$$

admits a solution if $A(t, \cdot)$ is essentially bounded. By this we mean that there exists an M such that for all $t \in [0, 1]$, $\mathbb{P}(\|A(t, \cdot)\| \leq M) = 1$, where $\|\cdot\|$ denotes any matrix norm.

We conclude this section with some general remarks.

Remark 3. By identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ we can treat complex-valued RODEs on the same footing as real-valued RODEs.

Remark 4. Corollary 1 suggests that, if we want to obtain solutions to a linear RODE with finite moments up to a given order, we should restrict ourselves to studying RODEs with bounded right-hand side. However, there is one case where it is possible and indeed natural to consider unbounded linear RODEs. If $A(t, \omega) \in \mathfrak{g} \subset \text{Mat}_n(\mathbb{C})$, where $\mathfrak{g}$ is a Lie subalgebra of $\text{Mat}_n(\mathbb{C})$ such that the associated Lie group $G \subset GL_n(\mathbb{C})$ is compact, and $A(t, \omega)$ is continuous in t for almost all ω , then the above RODE admits solutions with finite moments. The proof follows from using the escape lemma to solve the SP problem for each ω , measurability then follows from Remark 2 and finiteness of moments is evident since $x(t, \omega) \in G$ for all t for almost all ω , and G is compact.

Remark 5. Certain types of RODEs can be reinterpreted as an Itô-type SDE: If $W(t, \cdot)$ is the classical Wiener process, and μ and σ are integrable functions, the RODE

$$\frac{dx(t, \cdot)}{dt} = \mu(x(t, \cdot), t) + \sigma(x(t, \cdot), t) \frac{dW(t, \cdot)}{dt} \quad (16)$$

is equivalent to the Itô stochastic differential equation

$$dx(t, \cdot) = \mu(x(t, \cdot), t)dt + \sigma(x(t, \cdot), t)dW(t, \cdot). \quad (17)$$

Of course, $\frac{dW(t, \cdot)}{dt}$ only exists in the sense of distributions.

Remark 6. Uncorrelated centered noise does not affect the dynamics. Consider the RODE

$$\frac{dx(t, \cdot)}{dt} = A(t, \cdot)x(t, \cdot), \quad x(0, \cdot) = x_0(\cdot) \quad (18)$$

and suppose that $\mathbb{E}[A(t, \cdot)] = 0$, and $\mathbb{E}[A_{ij}(t)A_{kl}(s)] = 0$ for all $s \neq t$. Then, the only L^2 solution is given by $x(t) = x_0$ for all t .

First, note that if $B_{ij}(t) = \int_0^t A_{ij}(s)ds$, then $B_{ij}(t) = 0$ in $L^2(\Omega)$ for all t , since

$$\begin{aligned} \mathbb{E}[B_{ij}(t)^2] &= \mathbb{E}\left[\int_0^t A_{ij}(s_1)ds_1 \int_0^t A_{ij}(s_2)ds_2\right] \\ &= \int_0^t \int_0^t \mathbb{E}[A_{ij}(s_1)A_{ij}(s_2)]ds_1ds_2 = 0. \end{aligned} \quad (19)$$

Here, we have used Fubini's theorem to interchange integration with respect to s_i and $\mathbb{E}$, and the fact that, by assumption, $\mathbb{E}[A_{ij}(s_1)A_{ij}(s_2)]$ is supported only on the diagonal $\{s_1 = s_2\}$, which has measure 0.

Now, using equivalence of L^2 and SP solutions (Theorem 1), we can solve Eq. (18) using the Magnus expansion:

$$x(t, \omega) = \exp\left(\sum_{k=1}^{\infty} M_k\right), \quad (20)$$

with

$$M_1 = \int_0^t A(t_1)dt_1, \quad (21)$$

$$M_2 = \frac{1}{2} \int_0^t \int_0^{t_1} [A(t_1), A(t_2)]dt_1dt_2. \quad (22)$$

$$\vdots \quad (23)$$

Equation (19) shows that M_1 vanishes for all t , and a similar argument shows that all the higher order terms vanish too.

There is however one type of uncorrelated noise that does affect the dynamics, namely, the Gaussian white noise of Remark 5, which of course is not an L^2 process. Making sense of this rigorously is difficult, requiring techniques from Malliavin calculus. Setting aside technical details, if we denote $G(t, \cdot) = \frac{dW}{dt}(t, \cdot)$, we have that $\mathbb{E}[G(t), G(s)] = \delta(t - s)$. When we try to apply the same argument as in Eq. (19), we instead find that

$$\begin{aligned} \mathbb{E}[W(t, \cdot), W(s, \cdot)] &= \int_0^t \int_0^s \delta(t_1 - s_1)dt_1ds_1 \\ &= \min(t, s) \end{aligned} \quad (24)$$

as expected. This illustrates the main difference between SDEs and RODEs:

(1) RODEs are suitable for modeling systems subject to bounded noise, since otherwise we cannot in general expect that a solution even exists. If we want to have nontrivial stochastic behavior, we need the noise to exhibit some degree of correlation at different times.

(2) SDEs are suitable for modeling systems subject to uncorrelated noise. However, in order for the stochastics to be nontrivial, the noise process must live in a more exotic class of stochastic process than the class of L^2 processes. Another drawback of SDEs is that, unlike RODEs, they are not amenable to pathwise analysis.

Robust control for RODEs

Just as there is a notion of control theory for ODEs, there is a natural notion of *robust control* theory for RODEs. Suppose that we have an RODE

$$\begin{aligned} \frac{dx(t, \cdot)}{dt} &= f(a_1(t), \dots, a_k(t), x(t, \cdot), \eta(a_1(t), \dots, a_k(t), t, \cdot)), \\ x(0, \cdot) &= x_0 \in L^2(\Omega, \mathbb{R}^n), \end{aligned} \quad (25)$$

where the $a_j(t)$ are real-valued control functions. Note that even for complex RODEs, we can still assume that the control functions are real valued, by splitting into real and imaginary parts if necessary. Denote by A the function space of admissible control functions, i.e., $a := (a_1(\cdot), \dots, a_k(\cdot)) \in A$. Note also that the noise process η can depend on the control functions $a_j(t)$. We refer to this dependency as a *control-noise coupling*.

Definition 4. Given some vector y in $\mathbb{R}^n$, the objective of the robust control problem is to find control functions $a_j(t)$ and a stopping time $T \in \mathbb{R}$ minimizing the *cost function*

$$c(x(T, \cdot), a_j(\cdot), T, y). \quad (26)$$

We will mostly restrict our attention to cost functions that depend only on the random variable $x(T, \cdot)$ and the target y . We also restrict our attention to the situation where $\mathbb{E}[x(T, \cdot)] = y$, and we assume that we can find a control $a \in A$ such that this occurs. We refer to this last condition as *operator controllability in expectation*.

III. THE RANDOM SCHRÖDINGER EQUATION

The RODE we will study in this paper is the *random Schrödinger equation* for the time-evolution operator:

$$\begin{aligned} \frac{dU_S(t, \cdot)}{dt} &= -i(H_{0,S}(t) + H_{1,S}(t, \cdot))U_S(t, \cdot), \\ U_S(0, \cdot) &= \mathbb{1}, \end{aligned} \quad (27)$$

where the subscript S denotes that we are working in the Schrödinger picture. Here, $H_{0,S}(t)$ is the deterministic part of the Hamiltonian, while $H_{1,S}(t, \cdot)$ is a Hermitian matrix-valued random process. In what follows, we will assume that $H_{1,S}$ is *essentially bounded*; although in view of Remark 4, this is not necessary to guarantee the existence of a solution. We may further assume that $H_{1,S}$ is centered, since we can always absorb the drift term into $H_{0,S}$.

As expected, there is also random Schrödinger equation for (random) states:

$$\begin{aligned} \frac{d|\varphi(t, \cdot)\rangle}{dt} &= -i(H_{0,S}(t) + H_{1,S}(t, \cdot))|\varphi(t, \cdot)\rangle, \\ |\varphi(0, \cdot)\rangle &= |\varphi_0(\cdot)\rangle, \end{aligned} \quad (28)$$

and the usual relation between states and operators holds pathwise: $|\varphi(t, \omega)\rangle = U_S(t, \omega)|\varphi_0(\omega)\rangle$. Figure 1 shows plots of solutions to the random Schrödinger equation for states, visualized on the Bloch sphere using QuTiP Ref. [25].

Before discussing the theoretical guarantees for systems modelled by the random Schrödinger equation, we discuss the relationship between the random Schrödinger equation and other more traditional noise models.

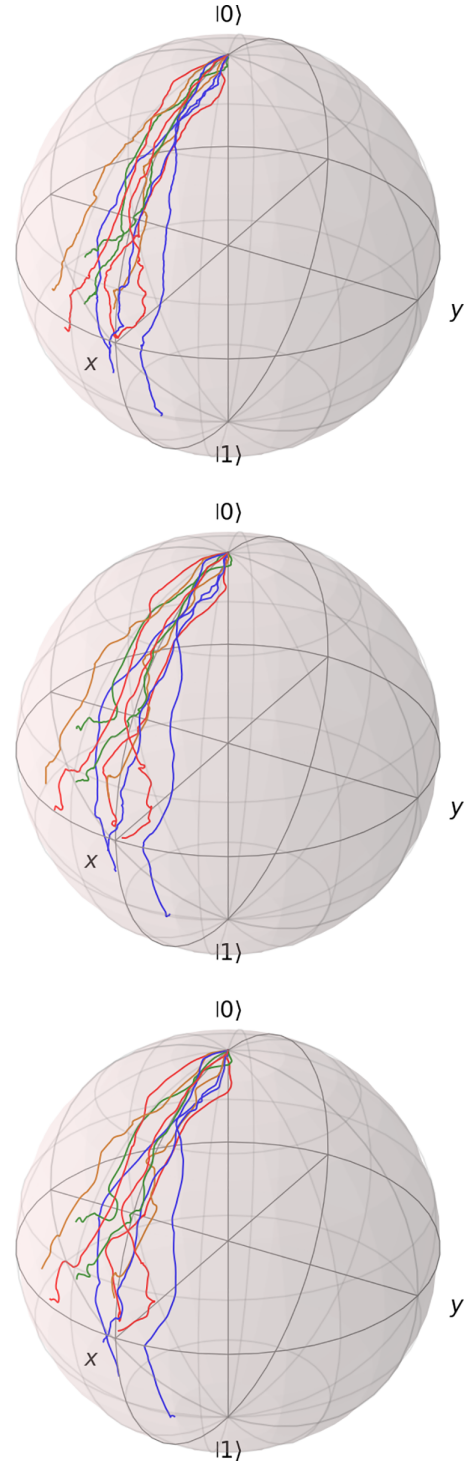


FIG. 1. Plots of 10 pathwise solutions to the random Schrödinger equation for a single qubit with $H_{0,S}(t) = \cos(t)\sigma_y$ and $H_{1,S}(t, \omega) = \frac{1}{2}M_\nu(t, \cdot)$, where $M_\nu(t, \cdot)$ is a Gaussian process with Matérn covariance. The parameter ν controls the smoothness of sample paths. We take to ν to be 0.2, 0.6, and 1.0 from top to bottom. A Gaussian process with Matérn covariance is in general $\lceil \nu \rceil - 1$ times L^2 differentiable. Trajectories were sampled using *sci-kit learn* and plotted using *QuTiP* [25].

A. Mixed-state behaviour

Since $H_{0,S}(t) + H_{1,S}(t, \omega)$ is Hermitian, the sample trajectories $U_S(t, \omega)$ are always unitary. However, by considering

ensemble averages over the probability space Ω , we obtain the mixed-state behavior that we expect open quantum systems to exhibit.

Consider a (normalized) random state $\psi \in L^2(\Omega, \mathcal{H})$. From this, we can form a random density operator

$$\rho : \Omega \rightarrow \mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{B}(\mathcal{H}), \rho = \rho^\dagger, \rho \succeq 0, \text{Tr}(\rho) = 1\},$$

$$\omega \mapsto |\psi(\omega)\rangle\langle\psi(\omega)|. \quad (29)$$

For each $\omega \in \Omega$, $\rho(\omega)$ is a pure state, however, taking an ensemble average, we can obtain a matrix $\hat{\rho} = \mathbb{E}_\omega(\rho(\omega))$.

Now, $\hat{\rho}$ is in fact a density matrix, since $\text{Tr}(\hat{\rho}) = 1$, and it is in general a mixed state. To see this, note that $\rho(\omega)$ is supported on the space of pure states,

$$\mathcal{D}(\mathcal{H}) \cap \{\text{Tr}\rho^2 = 1\}. \quad (30)$$

We know that the expectation $\mathbb{E}(\rho)$ is contained in the convex hull of the support. Since the space of pure states contains no nontrivial convex subsets, it is clear that $\mathbb{E}(\rho)$ is a mixed state as soon as $\rho(\omega)$ is nonconstant.

B. Generating standard error processes

Despite its apparent simplicity, the random Schrödinger equation captures most “standard” quantum errors considered in quantum information theory and quantum error correction.

Consider first the case of a bit flip channel on a single qubit. This can be described in terms of the Kraus maps $K_1 = \sqrt{p}X$, $K_0 = \sqrt{1-p}\mathbb{1}$, for some parameter $p \in [0, 1]$. It is not difficult to write down a random Hamiltonian that generates this channel upon taking ensemble averages. If we set

$$H_{0,S}(t, \omega) = \begin{pmatrix} 0 & \frac{\pi}{2} \\ -\frac{\pi}{2} & 0 \end{pmatrix} \quad \text{with probability } p,$$

$$H_{0,S}(t, \omega) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{with probability } 1 - p, \quad (31)$$

then evolving for one unit of time generates the bit flip channel. Of course this choice is nonunique.

Indeed, the same trick works for any mixed unitary error admitting Kraus maps $\{p_j U_j\}$ with $\sum_j p_j = 1$. We can choose Hamiltonians $H_j(t)$ generating U_j , and let $H_{1,S}(t, \omega)$ be a random whose trajectory is given by $H_j(t)$ with probability p_j . Again, this choice is (very) nonunique. Furthermore, one could generate arbitrary quantum channels by evolving via non-Hermitian noise processes, as for instance in Ref. [26]. However, this is beyond the scope of this paper.

Another common error encountered in the literature is the so-called $\frac{1}{f}$ noise. This fits naturally into our framework. For instance, in Ref. [27], the authors consider the random

Hamiltonian

$$\hat{H}_q = \frac{1}{2} \mathbf{B} \cdot \vec{\sigma}. \quad (32)$$

Here, $\mathbf{B}$ denotes the magnetic field, and $\vec{\sigma}$ is the Pauli-vector. The authors model the magnetic field as

$$\mathbf{B} = \mathbf{B}_0(t) + \mathbf{b}(t, \omega), \quad (33)$$

where $\mathbf{b}(t, \omega)$ is a stochastic process, which coincides with the random Schrödinger formalism so long as $\mathbf{b}(t, \omega)$ is essentially bounded.

C. Basic error estimate

In this section, $\|\cdot\|$ may be taken to denote the operator norm induced by the Hilbert norm on $\mathcal{H}$, although exactly the same results holds for the Frobenius norm $\|\cdot\|_F$. The first natural question to ask is how the noise term affects the solution. In other words, if we denote by $U_{0,S}(t)$ the propagator associated with $H_{0,S}(t)$, how does the random variable $\|U_S(t, \cdot) - U_{0,S}(t)\|$ evolve over time?

The random-variable $\|U_S(t, \cdot) - U_{0,S}(t)\|$ can be thought of as the error induced by perturbing the ideal control Hamiltonian $H_{0,S}$ by the additive noise $H_{1,S}(t, \cdot)$. We have the following theorem:

Theorem 1. If $\text{ess sup}_{t,\omega} \|H_{1,S}(t, \omega)\| \leq K$, the worst case error induced by the noise grows linearly in time. That is, $\text{ess sup}_\omega \|U_S(t, \omega) - U_{0,S}(t)\| \leq Kt$ with probability 1.

Remark 7. This is much better than we can expect using standard results such as Gronwall’s inequality. This is because the dynamics are constrained to $\text{SU}(n)$, which is compact.

Proof. The proof proceeds via a transformation to the interaction picture. For a Schrödinger state $|\varphi_S(t)\rangle$, define the corresponding interaction state

$$|\varphi_I(t)\rangle = U_{0,S}(t)^\dagger |\varphi_S(t)\rangle. \quad (34)$$

For operators, we recall that for a Hermitian operator A_S in the Schrödinger picture, the corresponding operator in the interaction picture is given by

$$A_I = U_{0,S}(t)^\dagger A_S U_{0,S}(t). \quad (35)$$

It follows from the standard theory that time-evolution operator $U_I(t)$ satisfies the interaction picture Schrödinger equation

$$\frac{d}{dt} U_I(t) = -i H_{1,I}(t) U_I(t), \quad (36)$$

where $H_{1,I} = U_{0,S}(t)^\dagger H_{1,S}(t) U_{0,S}(t)$.

Of course, if $H_{1,S}$ is 0; i.e., if there is no noise, then there is no state evolution (and only operators evolve with time). Moreover, since the transformation to the interaction picture is unitary, we have that

$$\|U_S(t, \cdot) - U_{0,S}(t)\| = \|U_I(t, \cdot) - \mathbb{1}\|. \quad (37)$$

The quantity on the right-hand side is easy to estimate. Fixing $\omega \in \Omega$, we rewrite the interaction picture Schrödinger

equation as an integral equation:

$$\begin{aligned} U_I(t, \omega) - \mathbb{1} &= -i \int_0^t H_{1,I}(s, \omega) U_I(s, \omega) ds \\ &= -i \int_0^t U_{0,S}(s)^\dagger H_{1,S}(s, \omega) U_{0,S}(s) U_I(s, \omega) ds. \end{aligned} \quad (38)$$

Hence,

$$\begin{aligned} \|U_I(t, \omega) - \mathbb{1}\| &= \left\| \int_0^t U_{0,S}(s)^\dagger H_{1,S}(s, \omega) U_{0,S}(s) U_I(s, \omega) ds \right\| \\ &\leq \int_0^t \|U_{0,S}(s)^\dagger H_{1,S}(s, \omega) U_{0,S}(s) U_I(s, \omega)\| ds \\ &= \int_0^t \|U_{0,S}(s)^\dagger\| \|H_{1,S}(s, \omega)\| \\ &\quad \times \|U_{0,S}(s)\| \|U_I(s, \omega)\| ds \\ &= \int_0^t \|H_{1,S}(s, \omega)\| ds \leq Kt, \end{aligned} \quad (39)$$

where the last inequality holds for almost all ω . ■

Remark 8. We have a similar result for any submultiplicative matrix norm, not just the operator norm or Frobenius norm, potentially at the cost of replacing K by a larger constant, since multiplying by a unitary matrix will not preserve arbitrary matrix norms. However, since the only norms we consider in this paper are the operator norm and the Frobenius norm, both of which are preserved by unitary matrices.

D. A geometric error estimate

Unfortunately, the inequality proven in the previous section is not tight, since the left-hand side measures the distance between $U_I(t, \omega)$ and the identity in the space of matrices, whereas the integral on the right-hand side measures the length of a path on $SU(n)$. In order to get a tighter bound, we need to work intrinsically on $SU(n)$.

When working intrinsically on $SU(n)$, it is most natural to use the Frobenius norm $\|\cdot\|_F$ on matrices, since the associated inner product pulls back via the inclusion $SU(n) \rightarrow M_n(\mathbb{C})$ to the unique bi-invariant metric on $SU(n)$ that makes the inclusion an isometric embedding.

Let us assume that $\|H_{1,I}(t)\|_F \leq K$ almost surely for all t . For $U, V \in SU(n)$, let $d_F(U, V)$ denote the distance between U and V on $SU(n)$ measured with respect to the metric induced by the Frobenius norm (also known as the Killing metric).

We have the following result:

$$\begin{aligned} \|U_I(t, \omega) - \mathbb{1}\|_{op} &\leq \|U_I(t, \omega) - \mathbb{1}\|_F \\ &\leq d_F(U_I(t, \omega), \mathbb{1}) \\ &\leq \int_0^t \|H_{1,I}(s, \omega) U_I(s, \omega)\|_F ds \\ &= \int_0^t \|H_{1,I}(s, \omega)\|_F ds \leq Kt \quad a.s. \end{aligned} \quad (40)$$

Of course, we may wish to study noise whose amplitude varies with time. Defining

$\Lambda(t) := \text{ess sup}_\omega \|H_{1,I}(t, \omega)\|_F$, the previous is easily adapted to show that

$$d_F(U_I(t, \omega), \mathbb{1}) \leq \int_0^t \|H_{1,I}(s, \omega)\|_F ds \leq \int_0^t \Lambda(s) ds. \quad (41)$$

We have proven the following theorem:

Theorem 2. If $\Lambda(t) := \text{ess sup}_\omega \|H_{1,I}(t, \omega)\|_F$, we have the following inequality:

$$\begin{aligned} d_F(U_I(t, \omega), U_{0,S}(t)) &\leq \int_0^t \|H_{1,S}(s, \omega)\|_F ds \\ &\leq \int_0^t \Lambda(s) ds. \end{aligned} \quad (4)$$

E. Tightness of error bound

In fact, for certain noise processes, the error bound [Eq. (41)] is optimal; i.e., the noise can push the error arbitrarily close to the bound given by Eq. (41).

Theorem 3. For noise processes $H_{1,S}$ satisfying the *worst-case noise condition* (see Definition 5), we have

$$\mathbb{P}\left[\int_0^t \Lambda(s) ds - d_F(U_I(t, \omega), \mathbb{1}) < \epsilon\right] > 0. \quad (5)$$

Definition 5. We say that $H_{1,I}(t, \cdot)$ satisfies the worst-case noise condition with respect to $H_{0,S}$ if the following two conditions hold. The first condition is

$$\mathbb{P}(H_{1,I}(t, \cdot) = 0) = 0 \quad (42)$$

for all t . In this case, we write

$$H_{1,I}(t, \omega) = \lambda(t, \omega) \hat{H}_I(t, \omega), \quad (43)$$

where $\|\hat{H}_{1,I}(t, \omega)\|_F = 1$.

The second condition is that for all $\epsilon > 0$ and for all $t \geq 0$,

$$\begin{aligned} \mathbb{P}[|\lambda(t, \omega) - \Lambda(t)| < \epsilon \\ \text{and } \|\hat{H}_{1,I}(t, \omega) - \hat{H}_{1,I}(0, \omega)\|_F < \epsilon] > 0. \end{aligned} \quad (44)$$

For convenience, we denote

$$\begin{aligned} E_\epsilon^I &= \{\omega \in \Omega \text{ s.t. } |\lambda(t, \omega) - \Lambda(t)| < \epsilon \\ \text{and } \|\hat{H}_{1,I}(t, \omega) - \hat{H}_{1,I}(0, \omega)\|_F < \epsilon\}. \end{aligned} \quad (45)$$

If $H_{1,I}(t, \cdot)$ satisfied the WCNC with respect to $H_{0,I}(t, \cdot)$, then we have that $\mathbb{P}(E_\epsilon^I) > 0$ for all $\epsilon > 0$.

Suppose that E_ϵ^I occurs, that is, $\omega \in \Omega$. Then $H_{1,I}(t, \omega) = \Lambda(t) \hat{H}_{1,I}(0, \omega) + G(t, \omega)$, where $\|G(t)\|_F$ is $\mathcal{O}(\epsilon)$.

First, define $\tilde{U}_I(t) = \exp(-i \int_0^t \Lambda(s) \hat{H}_{1,I}(0) ds)$. Note that $\tilde{U}_I(t)$ defines a geodesic on $SU(n)$ and moreover, that $d_F(I, \tilde{U}_I(t)) = \int_0^t \Lambda(s) ds$. Hence, $\tilde{U}_I(t)$ achieves the bound [Eq. (41)].

Next, we will show that $d_F(U_I(t, \omega), \tilde{U}_I(t))$ is $\mathcal{O}(\epsilon)$. One relatively straightforward way to do this is via a second interaction picture transformation. Define $\tilde{H}_{1,I}(t) = \Lambda(t) \hat{H}_{1,I}(0, \omega)$. Notice that $\tilde{H}_{1,I}(t)$ commutes with itself at all times, and hence by definition $\tilde{U}_{1,I}(t)$ is the unitary propagator associated with $\tilde{H}_{1,I}(t, \omega)$. Now, perform another transformation to the interaction picture, and denote operators in the double interaction picture by a subscript I' . By definition, $H_{1,I'}(t, \omega) = \tilde{U}_I(t, \omega)^\dagger G(t, \omega) \tilde{U}_I(t, \omega)$, and letting

$U_I(t, \omega)$ denote the time-evolution operator in the (double) interaction picture, we have the equation

$$\frac{dU_I(t, \omega)}{dt} = -iH_{1,I'}(t, \omega)U_I(t, \omega). \quad (46)$$

Now, since $\tilde{U}_I(t, \omega)$ is unitary, we have that $\|H_{1,I'}(t, \omega)\|_F = \mathcal{O}(\epsilon)$, and we also know that $d_F(U_I(t, \omega), \tilde{U}_I(t, \omega)) = d_F(U_I(t, \omega), \mathbb{1})$. But since $H_{1,I'}(t, \omega)$ is $\mathcal{O}(\epsilon)$, we see that $d_F(U_I(t, \omega), \mathbb{1})$ is $\mathcal{O}(\epsilon)$, for example by considering the Magnus expansion.

Putting this all together and using the triangle inequality for $d_F(\cdot, \cdot)$ on $SU(n)$, we see that

$$\int_0^t \Lambda(s)ds - d_F(I, U_I(t)) < d_F(U_I(t), \tilde{U}_I(t)) = \mathcal{O}(\epsilon). \quad (47)$$

But since by assumption, for all ϵ , $\mathbb{P}(E_\epsilon^I) > 0$, we conclude that

$$\mathbb{P}\left[\int_0^t \Lambda(s)ds - d_F(U_I(t, \omega), \mathbb{1}) < \epsilon\right] > 0 \quad (48)$$

for all ϵ .

F. Deriving a condition on the Schrödinger noise Hamiltonian

One slightly unsatisfactory part of this reasoning is that the worst-case noise condition is defined in terms of the interaction-picture noise Hamiltonian $H_{1,I}(t, \omega)$, which of course depends on the $H_{0,S}$. In the context of control theory, this is problematic. In this section, we derive a sufficient condition on the Schrödinger noise Hamiltonian.

First note that if $H_{1,I}(t, \omega) = 0$ if and only if $H_{1,S}(t, \omega) = 0$, and if $H_{1,I}(t, \omega) = \lambda(t, \omega)\hat{H}_{1,I}(t, \omega)$, then similarly $H_{1,S}(t, \omega) = \lambda(t, \omega)\hat{H}_{1,S}(t, \omega)$. Now note that, by the definition of the interaction picture,

$$\begin{aligned} \hat{H}_{1,I}(t, \omega) &\in B_{F,\epsilon}(\hat{H}_{1,I}(0, \omega)) \\ &\Updownarrow \\ \hat{H}_{1,S}(t, \omega) &\in B_{F,\epsilon}(U_{0,S}(t)\hat{H}_{1,S}(0, \omega)U_{0,S}(t)^\dagger). \end{aligned}$$

The next step is to remove the dependence on the control $U_{0,S}(t)$. Denote by CT the space of all allowable control trajectories $[0, T] \rightarrow SU(n)$, i.e., the space of all trajectories generated by admissible controls $H_{0,S}(t)$. Of course, precisely which controls are admissible depends on the context.

This motivates the following definition:

Definition 6. We say that $H_{1,S}(t, \omega)$ satisfies the worst-case noise condition with respect to CT if

$$\begin{aligned} \mathbb{P}[\hat{H}_{1,S}(t, \omega) \in B_{F,\epsilon}(U_{0,S}(t)\hat{H}_{1,S}(0, \omega)U_{0,S}(t)^\dagger); \\ \Lambda(t) - \lambda(t, \omega) < \epsilon; \forall t] \end{aligned} \quad (49)$$

is strictly positive for all ϵ . A simpler way to view this condition is that $H_{1,S}$ satisfies the WCNC with respect to CT if and only if $H_{1,S}$ satisfies the WCNC with respect to all $H_{0,S}(t)$ generating trajectories in CT .

Of course, given a class of admissible control trajectories CT , we might ask which noise processes satisfy this condition. In general, this is a difficult question, but we will show that in most cases this class contains several physically reasonable noise processes.

For bounded controls $H_{0,S}(t)$, there is a sufficient condition to check whether $H_{1,S}(t, \cdot)$ satisfies the WCNC with respect to $H_{0,S}(t)$.

Theorem 4. Let $\mathcal{H}_E$ denote the space of Hermitian matrices with Frobenius norm less than E , and let $H_{1,S}(t, \cdot)$ be a stochastic process with almost surely continuous sample paths and let $H_{0,S}(t) \in \mathcal{H}_E$ for some $E \in \mathbb{R}$. Suppose further that $H_{1,S}(t, \cdot)$ satisfies the strong Markov property, and that the associated measure on $C^0([0, T], \mathcal{H}_E)$ is positive on cylindrical sets. Then $H_{1,S}(t, \cdot)$ satisfies the worst-case noise condition with respect to $H_{0,S}$.

Proof. Suppose $H_{1,S}$ satisfies the assumptions stated above.

We begin by observing that the WCNC is equivalent to the condition that

$$\mathbb{P}[H_{1,S}(t, \omega) \in B_{F,\Lambda(t)\epsilon}(U_{0,S}(t)H_{1,S}(0, \omega)U_{0,S}(t)^\dagger) \forall t] > 0. \quad (50)$$

The next step is to find an inner approximation of

$$\begin{aligned} \bigcup_{t \in [0, T]} B_{F,\Lambda(t)\epsilon}(U_{0,S}(t)H_{1,S}(0, \omega)U_{0,S}(t)^\dagger) \times \{t\} \\ \subset \mathcal{H} \times [0, T] \end{aligned} \quad (51)$$

by a finite concatenation of cylindrical sets. This is always possible, since $\frac{d}{dt}U_{0,S}(t)H_{1,S}(0, \omega)U_{0,S}(t)^\dagger$ is bounded by $E\lambda(0, \omega)$. Precisely, there exists a partition $0 = t_0 < \dots < t_K = T$ of $[0, T]$ and $\epsilon_i < \epsilon$ such that

$$\begin{aligned} \bigcup_{i=0}^{K-1} \bigcup_{t \in [t_i, t_{i+1}]} B_{F,\Lambda(t_i)\epsilon_i}(U_{0,S}(t_i)H_{1,S}(0, \omega)U_{0,S}(t_i)^\dagger) \times \{t\} \\ \subset \bigcup_{t \in [0, T]} B_{F,\Lambda(t)\epsilon}(U_{0,S}(t)H_{1,S}(0, \omega)U_{0,S}(t)^\dagger) \times \{t\}, \end{aligned} \quad (52)$$

and such that

$$\begin{aligned} B_{F,\Lambda(t_i)\epsilon_i}(U_{0,S}(t_i)H_{1,S}(0, \omega)U_{0,S}(t_i)^\dagger) \\ \cap B_{F,\Lambda(t_{i+1})\epsilon_{i+1}}(U_{0,S}(t_{i+1})H_{1,S}(0, \omega)U_{0,S}(t_{i+1})^\dagger). \end{aligned} \quad (53)$$

is nonempty.

To ease notation, let

$$B_i = B_{F,\Lambda(t_i)\epsilon_i}(U_{0,S}(t_i)H_{1,S}(0, \omega)U_{0,S}(t_i)^\dagger) \quad (54)$$

for all $i = 0, \dots, K-1$. Now, note that

$$\begin{aligned} \mathbb{P}\left[H_{1,S}(t, \omega)([0, T]) \subset \bigcup_{i=0}^{K-1} \bigcup_{t \in [t_i, t_{i+1}]} B_i \times \{t\}\right] \\ = \mathbb{P}[H_{1,S}(t, \omega) \in B_i \text{ for } t \in [t_i, t_{i+1}] \forall i]. \end{aligned} \quad (55)$$

Denote by F_i the event

$$H_{1,S}(t, \omega) \in B_i \text{ for } t \in [t_i, t_{i+1}]. \quad (56)$$

By the assumption that the path measure associated with $H_{1,S}$ is positive on cylindrical sets, $\mathbb{P}(E_i) > 0$ for all i . By the strong Markov assumption, we have that

$$\mathbb{P}[F_{i+1}|F_i] = \mathbb{P}[F_{i+1}|H_{1,S}(t_i, \omega) \in B_{i+1} \cap B_i]. \quad (57)$$

Now, for all i , $\mathbb{P}[H_{1,S}(t_i, \omega) \in B_{i+1} \cap B_i]$ is nonzero. By the definition of conditional probability, we conclude that

$$\mathbb{P}[H_{1,S}(t, \omega) \in B_i \text{ for } t \in [t_i, t_{i+1}] \forall i] > 0. \quad (58)$$

But since

$$\begin{aligned} & \bigcup_{i=0}^K \bigcup_{t \in [t_i, t_{i+1}]} B_i \times \{t\} \\ & \subset \bigcup_{t \in [0, T]} B_{F, \Lambda(t)}(U_{0,S}(t)H_{1,S}(0, \omega)U_{0,S}(t)^\dagger) \times \{t\} \\ & \subset \mathcal{H} \times [0, T], \end{aligned} \quad (59)$$

the result follows. $\blacksquare$

Remark 9. Identify the space of Hermitian matrices with $\mathbb{R}^k$ for the appropriate k . Suppose that $X(t, \cdot)$ is an k -dimensional unbounded Markov process, for example a Wiener process, an Ornstein-Uhlenbeck process, or a geometric Brownian motion. Let $f: \mathbb{R}^k \rightarrow B_E(0)$ be a continuous bijection. Then, it is a standard fact from stochastic analysis that $f(X_t)$ is still a Markov process. Moreover, it is straightforward to check that if the path measure associated X_t is positive on cylindrical sets in $[0, T] \times \mathbb{R}^k$, then the path measure associated with $f(X_t)$ is positive on cylindrical sets in $[0, T] \times \mathcal{H}_E \subset [0, T] \times \mathbb{R}^k$.

Hence, for example, in one dimension, letting $W(t, \cdot)$ denote the Wiener process, the random process

$$\frac{2E}{\pi} \arctan(W(t, \cdot)) \quad (60)$$

satisfies the assumptions of Theorem 4 and hence the WCNC.

We leave it to the reader to come up with an n -dimensional example by applying their favorite continuous bijection $\mathbb{R}^n \leftrightarrow B_E(0)$ to their favorite Markov process that satisfies the positivity condition on cylindrical sets. As a heuristic, one expects that most Markov processes with almost surely continuous sample paths will satisfy this last condition.

IV. ROBUST QUANTUM CONTROL

A. Control noise-couplings

In this final section, we consider the problem of synthesizing a target unitary V while minimizing the worst-case error $\text{ess sup}_\omega \|U_S(t, \omega) - V\|$, in the presence of a control noise coupling $H_{1,S}(t, \omega) = H_{1,S}(H_{0,S}(t), t, \omega)$. We will restrict our attention to optimizing over control Hamiltonians $H_{0,S}(t)$ generating V exactly in the absence of noise.

Remark 10. It would be tempting to talk about control Hamiltonians $H_{0,S}$ generating V in expectation. But this would be incorrect, unless the expectation is taken to be a *Fréchet mean*.

There are many natural couplings one could consider. The most comprehensive treatment of control noise couplings is given in Refs. [11,12], in the context of the stochastic Schrödinger equations, though the ideas easily translate to the RODE setting.

In each of these papers, the amplitude of the the noise is assumed to be proportional to the intensity of the control function, and in Ref. [12], the authors allow for both multiple

noise terms and multiple control terms. It is the latter approach that we consider here.

We are interested in situations where moving in certain unfavorable Hamiltonian directions induces more noise than moving in more favorable directions. For example, in the multiqubit setting, we generally expect that applying a Hamiltonian corresponding to multiple qubit operations will induce more noise than applying a Hamiltonian corresponding to one and two-qubit interactions, as in Refs. [19–22]. These couplings are natural even in the single qubit case. For example, when controlling a single spin qubit, [28] suggests a situation where a magnetic field can be easily applied in the x - y plane, but only applied with more difficulty in the z axis.

B. Couplings via noise metrics

We would like a coupling that captures both the proportionality of the amplitude of the noise to the intensity of the control functions, while also allowing for a “directional dependency.” Fix an orthonormal global frame $\{H_j\}$. Then any control field can be written as a sum $H_{0,S} = \sum_j h_j(t)H_j$. We will consider couplings $H_{1,S}(t, \omega) = H_{1,S}(H_{0,S}, t, \omega)$ such that

$$\begin{aligned} \text{ess sup}_\omega \|H_{1,S}(t, \omega)\|_F &= \Lambda(t) = \Lambda(H_{0,S}(t)) \\ &= \sqrt{\sum_j l_j h_j(t)^2}. \end{aligned} \quad (61)$$

If, for example, moving in the H_j direction induces a times as much noise as moving in the H_k direction, then we should set $\Lambda_k = a^2 \Lambda_j$, and so on. Of course, this means that $\Lambda(H_{0,S}(t))$ is actually induced by a Riemannian metric g_Λ on $\text{SU}(n)$. With respect to the orthonormal global frame $\{H_j\}$, g_Λ is represented by the diagonal matrix

$$\begin{pmatrix} l_1 & & \\ & \ddots & \\ & & l_m \end{pmatrix}. \quad (62)$$

We call such a metric a *noise metric*. Couplings arising from such a metric satisfy both the property of proportionality of the amplitude of the noise to the intensity of the signal, as well as allowing for directional dependence. Moreover, as alluded to in Ref. [12], the components of g_Λ can be either calculated or inferred experimentally. For a visualization of the geometry of g_Λ , see Ref. [28] and Fig. 2.

Next, recall that by Theorem 2, we know that

$$\|U_S(t, \omega) - V\|_F \leq d_F(U_S(t, \omega), V) \leq \int_0^t \|H_1(s, \omega)\|_F ds \quad (63)$$

for almsot all ω .

But since, by definition, $\|H_{1,S}(t, \omega)\| \leq \Lambda(t) = \sqrt{g_\Lambda(H_{0,S}(t), H_{0,S}(t))}$, we see that

$$\|U_S(t, \omega) - V\|_F \leq \int_0^t \sqrt{g_\Lambda(H_{0,S}(s), H_{0,S}(s))} ds = l_{g_\Lambda}(\gamma), \quad (64)$$

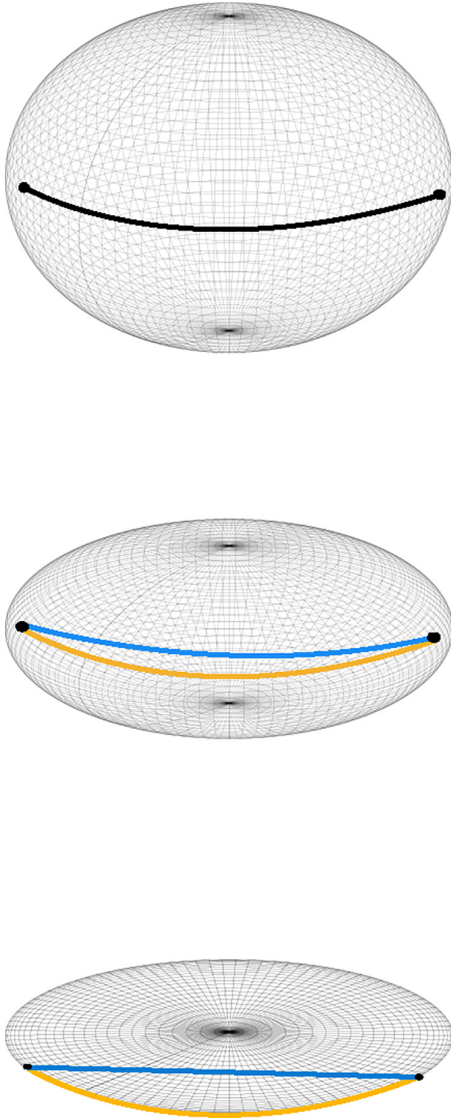


FIG. 2. Schematic of a deformation of $SU(n)$ corresponding to penalizing one direction. The picture on the top represents (a hyperplane section of) $SU(n)$ with the usual round metric, showing two “equatorial” points joined by a geodesic. The middle picture shows a deformed $SU(2)$, with the blue line representing the geodesic with respect to the warped metric, and the orange line shows the image of the geodesic with respect to the round metric. The bottom picture shows the limiting case, where one direction has been “infinitely penalized” effectively reducing the number of tangent directions (and hence the dimension) by one. Note that this diagram is merely a schematic: Each g_Λ is a homogeneous metric, meaning that the curvature should be identical at each point. Unfortunately, it is not possible to satisfactorily depict this in a sketch. For a more detailed discussion (see Ref. [28]).

where $l_{g_\Lambda}(\gamma)$ is the length, with respect to g_Λ , of the path on $SU(n)$ generated by the control field $H_{0,S}$.

This would suggest that minimizing the error

$$\text{ess sup}_\omega \|U_S(t, \omega) - V\|_F \quad (65)$$

is equivalent to finding a length-minimizing g_Λ geodesic from the identity to V . It turns out that this is true, as we might

expect from Theorem 3. However, we cannot apply Theorem 3 directly, due to the presence of the control-noise coupling. This motivates the following definition:

Definition 7. We say that $H_{1,S}(H_{0,S}(t), t, \omega)$ satisfies the *control-dependent worst-case noise condition* if, for each $U_0 \in CT$ generated by a control Hamiltonian $H_{0,S}(t)$, $H_{1,S}(H_{0,S}(t), t, \omega)$ satisfies the WCNC with respect to $H_0(t)$, as in Definition 5.

Remark 11. We observe that if the noise is independent of the control, we recover the WCNC with respect to CT .

With these definitions in place, we can state the main theorem of this section.

Theorem 5. Suppose that $H_{1,S}(t, \cdot)$ satisfies the control-dependent WCNC (see Definition 7) with respect to all $H_{0,S}(t) \in \mathcal{H}_E$, and suppose $\Lambda(t) = \sqrt{g_\Lambda(H_{0,S}(t), H_{0,S}(t))}$. Then solving the robust optimal control problem of synthesizing a unitary target V with respect to the cost function $c(U_S(T), H_{0,S}(\cdot), T, V) = \text{ess sup}_\omega \|U_S(T, \omega) - V\|$ is equivalent to finding a length-minimizing g_Λ geodesic on $SU(n)$.

Proof. First note that minimizing $d_F(U_S, V)$ minimizes $\|U_S(t, \omega) - V\|_F$. The result then follows by applying Theorem 3 to the Schrödinger RODE for each $H_{0,S}(t)$. ■

C. Calculating g_Λ geodesics

In theory, calculating geodesics with respect to g_Λ connecting two points U and V on $SU(n)$ is straightforward. We simply need to calculate the Christoffel symbols Γ_{bc}^a of the Levi-Civita connection associated with g_Λ , and then pass the geodesic equation

$$\ddot{x}^a + \Gamma_{bc}^a \dot{x}^b \dot{x}^c = 0, \quad (66)$$

$$x(0) = U, \quad x(1) = V, \quad (67)$$

where $x(t)$ denotes the geodesic path, to an ODE solver capable of handling Dirichlet boundary conditions (for instance, via shooting methods). Now, although the geodesic equation is formally a second-order equation, it is actually a first-order equation in $\dot{x}$. This means that it is tempting to calculate the Christoffel symbols with respect to some orthogonal (or indeed orthonormal) frame of vector fields $\{e_a\}$ in $SU(n)$. With respect to such a basis, and writing $g_\Lambda = g$ to simplify notation, the Christoffel symbols are given by

$$\begin{aligned} \Gamma_{bc}^a = \frac{1}{2} g^{da} [e_b(g_{cd}) + e_c(g_{db}) - e_d(g_{bc}) \\ + C_{dbc} + C_{cdb} + C_{bdc}]. \end{aligned} \quad (68)$$

Now, by the homogeneity of g , the first three terms vanish, since g is constant with respect to this frame. Moreover, taking $\{e_a\}$ to be (normalized) generalized Pauli vector fields, computing the structure constants is straightforward. However, it turns out that this approach does not work for our purposes: while we can easily obtain the geodesic equation for $\dot{x} = \dot{x}^a e_a$, we cannot impose Dirichlet boundary conditions $x(0) = U$, $x(1) = V$ in these coordinates.

If we want to impose Dirichlet boundary conditions, we need to work with respect to a coordinate basis. Working in a coordinate basis still leads to some simplifications, since the structure constants C_{abc} vanish due to symmetry of mixed partials. However, expressing g_{ab} with respect to coordinates on $SU(n)$ is delicate, since we in general expect to have a

description of $g = g_\Lambda$ with respect to a basis $\mathfrak{su}(n) = \text{TSU}(n)$ of generalized Pauli matrices.

Recall that the exponential map $\mathfrak{su}(n) \rightarrow \text{SU}(n)$ is a local diffeomorphism, and since $\mathfrak{su}(n) \cong \mathbb{R}^{n^2-1}$, the local inverse of the exponential map, corresponding to a choice of matrix logarithm (in the natural base) $\xi = \ln U$ gives local coordinates on $\text{SU}(n)$.

Now, in these coordinates, standard results from Lie theory (see, for example, Sec. 6.2 of Ref. [29]) allow us to derive expressions for left-invariant vector fields on $\text{SU}(n)$ with respect to these coordinates. Precisely, given $\eta \in \mathfrak{su}(n)$, there is an associated left-invariant invariant vector field η^L on $\text{SU}(n)$. On an open set $A \subset \mathfrak{su}(n)$ such that $\exp : A \rightarrow \text{SU}(n)$ is a diffeomorphism onto its image, the coordinate representation is given by pulling back η^L by the exponential map. We then have a formula (see, e.g., Appendix C.4. of Ref. [30]) in terms of the adjoint homomorphism:

$$(\exp|_A)^*(\eta^L)|_\xi = \frac{ad_\xi}{1 - \exp(-ad_\xi)} \eta. \quad (69)$$

Now, note that with respect to the coordinates, which we denote ξ^a , we have that $g_{ab} = g(\frac{\partial}{\partial \xi^a}, \frac{\partial}{\partial \xi^b})$, and so we can express g with respect to these vector fields, and hence compute (approximately) the Christoffel symbols. In the context of optimizing one-, two-, and three-qubit entangling gates, this approach is tractable, but it becomes impractical for large systems. For a k -qubit system, we need to perform this procedure on $\text{SU}(2^k)$, whose dimension grows exponentially with respect to the number of qubits. For this reason, numerical solution of this equation is impractical for large numbers of qubits without additional assumptions.

When working in this approach, we also need to contend with two sources of numerical error: clearly, solving the geodesic equation via numerical methods will induce some numerical error, but perhaps more fundamentally, in order to write down the geodesic equation itself, we are obliged to truncate a power series expression for the generalized Pauli-vector fields.

Nevertheless, for larger systems, Nielsen and Dowling [19] present an alternative method for calculating geodesics, based on the *lifted Hamilton-Jacobi-Bellman equation* and the notion of a *geodesic derivative*. The basic idea is to compute the geodesics for the round (Killing) metric, which correspond to time-independent Hamiltonians, and can hence be computed by finding a matrix logarithm of the target gate V , and then deforming these geodesics with geodesics of g_Λ .

D. The sub-Riemannian limit

Of course, even if one could solve the geodesic equation efficiently and accurately for a large number of qubits, the resulting fidelity-optimal Hamiltonian would in general contain many high-weight generalized Pauli matrices that are impracticable or indeed impossible to apply to any physical realization of the system. In most physical systems, we can only apply certain Hamiltonians that lie in the span of a subbundle $B \hookrightarrow \text{TSU}(n)$ of the tangent bundle spanned by certain low-weight Pauli matrices. It is a classical theorem in control theory and sub-Riemannian geometry, due to Chow and Rashevskii, that if this subbundle generates the whole

tangent bundle under the Lie bracket, then any target can be reached by a path whose tangent lies in the subbundle. The subbundle B is said to be *bracket-generating*, or *totally nonintegrable* (by contrast with subbundles that are closed under the Lie bracket, which are referred to as integrable).

In this context, it makes sense to look for length-minimizing curves between the identity and the target whose tangent vector lies in B . This gives rise to the so-called Carnot-Carathéodory metric:

$$d(U, V) = \inf_\gamma \int_0^1 \sqrt{g(\gamma(t), \gamma(t))} dt, \quad (70)$$

$$\dot{\gamma}(t) \in B, \quad \gamma(0) = U, \quad \gamma(1) = V. \quad (71)$$

The Carnot-Carathéodory metric on $\text{SU}(n)$ can be thought of as the limit of a noise metric when certain Hamiltonian directions are infinitely penalized. In Ref. [31], the authors make this claim rigorous. The work has important numerical and practical (hardware) implications.

From a numerical perspective, computing the sub-Riemannian geodesics (i.e., minimizers of the Carnot-Carathéodory distance) should in general be easier than computing the geodesics of the noise metrics introduced here (or, equivalently, of the complexity metrics appearing in Ref. [19]). This is because sub-Riemannian geodesics are “sparse” in the following sense: If B denotes the bundle of one- and two-qubit entangling Hamiltonians (corresponding to weight one and two generalized Pauli matrices) on k qubits, then we need to compute only $\mathcal{O}(k^2)$ coefficients, since there are only $\binom{k}{2} + \binom{k}{1}$ generalized Pauli matrices of weight one or two. On the other hand, computing the geodesics of noise metric requires us to compute 2^k coefficients.

From a practical perspective, the sub-Riemannian approach also has a significant advantage, since applying a Hamiltonian to a large number of qubits at once is either impractical or impossible. Computing sub-Riemannian geodesics allows one to restrict the optimal control problem to the space of Hamiltonians that are actually available to the controller, removing the need for the idealized assumption that the controller is able to apply any Hamiltonian to the system. Of course, the sub-Riemannian approach and the penalty factor approach can be applied at the same time in a natural way. Letting B denote the subbundle of available Hamiltonians, one can deform the metric on B to penalize some of the available directions. This allows us to distinguish “good”, “bad,” and “impossible” directions in the control problem.

V. NOISE-INFORMED QUANTUM CONTROL WITHOUT CALCULATING GEODESICS

It turns out that we can use the insight provided by Theorem 3 even when we cannot explicitly compute geodesics, using standard techniques from the theory of quantum optimal control. The basic idea is to minimize the error due to the noise by adding a soft constraint in the cost function of the quantum optimal control problem. We refer to this approach noise-informed quantum control. Moreover, in this setting we can easily incorporate a drift Hamiltonian.

Given $\{H_j\} \in \mathfrak{su}(n)$ and a target $V \in \text{SU}(n)$, we seek control functions $u_j(t) : [0, T] \rightarrow \mathbb{R}$ solving the robust optimal

control problem

$$\begin{aligned} & \min c(U(T, \cdot), V) \\ & \text{subject to } \frac{dU(t, \omega)}{dt} = -i \left(H_d + \sum_j u_j(t) H_j \right) U(t, \omega). \end{aligned} \quad (72)$$

Here, H_d denotes the drift Hamiltonian of the system, and the cost function $c(U(T), V)$ is generally some distance function on $SU(n)$. Common choices include the operator norm of the difference, the Frobenius norm of the distance, or their squares. Another common choice is to choose $C(U(T), V) = -\text{Tr}(U(T)^\dagger V)$, such that solving the optimization problem maximizes the *unitary fidelity*.

In this paper, we are interested in the quantity $\text{ess sup}_\omega \|U_S(T, \omega) - V\|$. In this section, we no longer consider controls generating V in (Fréchet) expectation; i.e., we don't require that $U_{0,S}(t) = V$.

Instead, we apply the triangle inequality:

$$\begin{aligned} & \text{ess sup}_\omega \|U_S(T, \omega) - V\| \\ & \leq \|U_{0,S}(T) - V\| + \text{ess sup}_\omega \|U_S(T, \omega) - U_{0,S}(T)\| \\ & \leq \|U_{0,S}(T) - V\| + \text{ess sup}_\omega \int_0^T \|H_{1,S}(t, \omega)\| dt \\ & = \|U_{0,S}(T) - V\| + \int_0^T \sqrt{g_\Lambda(H_{0,S}(t)), g_\Lambda(H_{0,S}(t))} dt. \end{aligned} \quad (73)$$

In particular, the first term captures the error due to only being able to solve the deterministic control problem approximately, and the second term corresponds to the error induced by the noise.

Now, we no longer have an explicit dependence on ω , so we have successfully recast the stochastic (robust) control problem as a deterministic control problem, sometimes called a *robust counterpart*. The cost function is given by

$$\begin{aligned} c_1(H_0(\cdot)) &= \|U_{0,S}(T) - V\| \\ &+ \int_0^T \sqrt{g_\Lambda(H_{0,S}(t)), g_\Lambda(H_{0,S}(t))} dt. \end{aligned} \quad (74)$$

In fact, minimizing the sum of the squares gives a cost function that is more amenable to standard numerical methods:

$$\begin{aligned} c_2(H_0(\cdot)) &= \|U_{0,S}(T) - V\|^2 \\ &+ \left[\int_0^T \sqrt{g_\Lambda(H_{0,S}(t)), g_\Lambda(H_{0,S}(t))} dt \right]^2. \end{aligned} \quad (75)$$

Minimizing c_2 yields a minimizer that provides an upper bound on the true minimum of c_1 , and provides satisfactory numerical results.

A. A one-qubit example

To illustrate the impact of the approximation [Eq. (75)], compared to not ignoring noise altogether, let us consider the

control of a single qubit, with Hamiltonian

$$H = H_d + \sum_{j=1}^3 u_j \sigma_j + H_n(t, \omega). \quad (76)$$

Here, H_d denotes the drift Hamiltonian, which we take to be $\frac{1}{2}\sigma_y$. The space of admissible control functions is given by fifth-degree polynomials of the form $u_j = \sum_{k=0}^5 c_j t^k$, which we initialize by choosing the c_k randomly.

We assume that the coupling between H_n is given by the g_Λ from the previous example,

$$g_\Lambda = \begin{pmatrix} 1 & & \\ & \frac{1}{100} & \\ & & \frac{1}{100} \end{pmatrix}, \quad (77)$$

so applying σ_x induces ten times as much noise as applying σ_y or σ_z .

We now run two numerical experiments [32], where we solve the quantum control problem in JULIA [33] using the numerical optimization package OPTIM [34]. In the first one, we minimize the cost function given by

$$\|U_{0,S}(T) - V\|^2, \quad (78)$$

which does not take the noise into account. In the second, we perform noise informed optimization and minimize

$$\begin{aligned} c_2(H_0(\cdot)) &= \|U_{0,S}(T) - V\|^2 \\ &+ \left[\int_0^T \sqrt{g_\Lambda(H_{0,S}(t)), g_\Lambda(H_{0,S}(t))} dt \right]^2. \end{aligned} \quad (79)$$

For this, we sample from a noise process in a way that satisfies the coupling. We simulate 100 noise trajectories, setting the entries of $H_{1,S}(t, \cdot)$ to the coordinatewise arctangent of a Gaussian process with Matérn covariance. We then multiply by the envelope function $\frac{1}{2\pi} g_\Lambda(H_0(t), H_0(t))$, which guarantees that $\text{ess sup}_\omega H_n(t) = g_\Lambda(H_0(t), H_0(t))$.

Finally, we evolve the system with each control (noise-aware and noise-blind) and compare the difference in fidelity, as shown in a violin plot in Fig. 3.

In this example, the worst-case signal-to-noise ratio, η , corresponding to the largest entry of g_Λ , is one. However, it turns out that the reduction in unitary error is robust across a fairly large range of signal-to-noise ratios. To study this, we consider signal-to-noise ratios varying in $[0, 2.0]$, where 0 implies no noise and 2.0 implies a major impact of noise as the system evolves, as we see in the noise term

$$\int_0^T \eta \sqrt{g_\Lambda(H_{0,S}(t)), g_\Lambda(H_{0,S}(t))} dt. \quad (80)$$

In Fig. 4, we follow a similar setting as for Fig. 3, but we the signal-to-noise ratio η . We plot the average unitary error, both in the noise aware setting, and ignoring the noise.

Both Figs. 3 and 4 illustrate how adding a noise term in the optimization problem helps find a more robust solution. We observe a linear increase in the average error with the signal-to-noise ratio, and a more regular behavior from the noise-aware results.

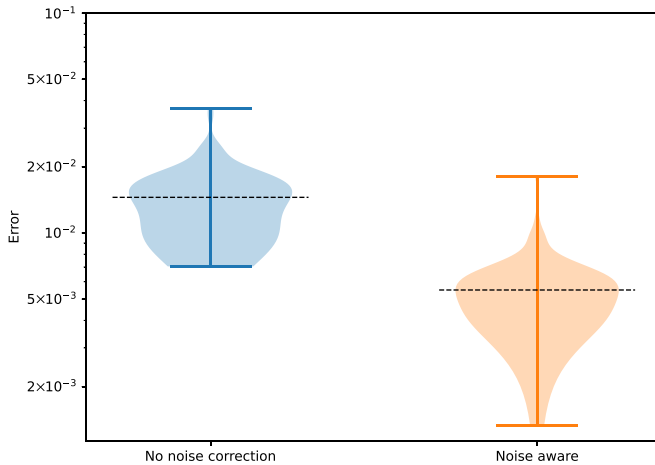


FIG. 3. Violin plots of the error, accounting for the noise [i.e., minimizing Eq. (80)] and not accounting for the noise [i.e., minimizing Eq. (78)]. Each distribution consists of 100 random samples for the control functions, with the dashed lines indicating the mean. We see an approximately 2.6-fold reduction in operator-norm error.

VI. CONCLUSION

Using the theory of random ordinary differential equations, we have proposed a new model of open quantum systems describing a large class of quantum systems subject to external noise. In particular, we prove tight bounds on the error induced by the noise. Inspired by the work of Nielsen, we have introduced a natural class of couplings between the control field and the noise term, which are described by a left-invariant metric on $SU(n)$ called a noise metric.

Additionally, we have demonstrated an approach to synthesizing a target unitary while minimizing the worst-case error, which can be construed as robust quantum control. By minimizing the worst-case error, we are able to transform a stochastic optimal control problem into a more tractable deterministic control problem. This is similar to the approach of constructing a *robust counterpart* in optimization theory. When the noise is coupled to the control via a noise metric, the optimal trajectory is a geodesic on $SU(n)$ with respect to the noise metric.

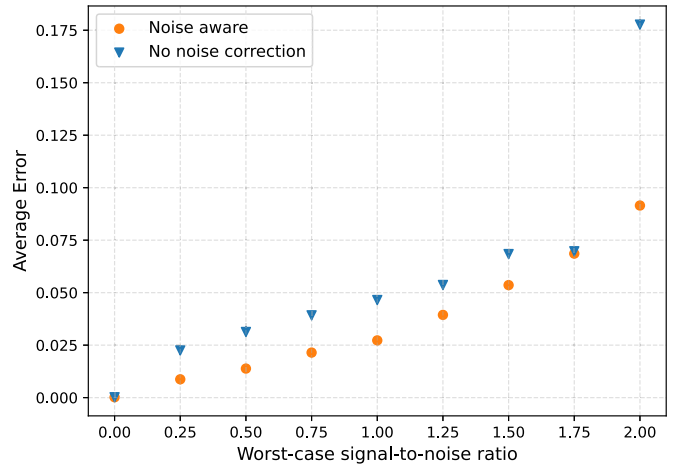


FIG. 4. A plot of the average unitary error accounting for the noise (orange circles) and not accounting for the noise (blue triangles), for 100 random control functions.

More work is needed to explicitly compute these geodesics, particularly in higher dimension, and in the sub-Riemannian limit. We hope to address these questions in future work. We have presented an alternative approach, referred to as noise-informed quantum optimal control, which does not involve computing geodesics. Instead, the method accounts for noise by incorporating a term corresponding to the noise as a soft constraint in classical optimal control problem, which can be solved using standard methods.

ACKNOWLEDGMENTS

The work of R.L., A.W., J.A., and J.M. has been supported by the Czech Science Foundation (23-07947S). L.B.G. has been supported by European Union's HORIZON-MSCA-2023-DN-JD program under the Horizon Europe (HORIZON) Marie Skłodowska-Curie Actions, Grant Agreement No. 101120296 (TENORS).

DATA AVAILABILITY

The data that support the findings of this article are openly available [27].

- [1] W. Parvaiz, J. Aspman, A. Wodecki, G. Korpas, and J. Marecek, Identifiability of autonomous and controlled open quantum systems, [arXiv:2501.05270](#).
- [2] V. P. Belavkin, A stochastic posterior Schrödinger equation for counting nondemolition measurement, *Lett. Math. Phys.* **20**, 85 (1990).
- [3] N. Gisin and I. C. Percival, The quantum-state diffusion model applied to open systems, *J. Phys. A: Math. Gen.* **25**, 5677 (1992).
- [4] W. T. Strunz, The Brownian motion stochastic Schrödinger equation, *Chem. Phys.* **268**, 237 (2001).
- [5] I. de Vega and D. Alonso, Dynamics of non-Markovian open quantum systems, *Rev. Mod. Phys.* **89**, 015001 (2017).
- [6] A. Chenu, M. Beau, J. Cao, and A. del Campo, Quantum simulation of generic many-body open system dynamics using classical noise, *Phys. Rev. Lett.* **118**, 140403 (2017).
- [7] H. Xue and X.-H. Deng, Traversing quantum control robustness landscapes: A new paradigm for quantum gate engineering, [arXiv:2412.19473](#).
- [8] R. L. Kosut, G. Bhole, and H. Rabitz, Robust quantum control: Analysis & synthesis via averaging, [arXiv:2208.14193](#).
- [9] J. Aspman, V. Kungurtsev, and J. Marecek, Robust quantum gate complexity: Foundations, [arXiv:2404.15828](#).
- [10] G. Dirr, U. Helmke, I. Kurniawan, and T. Schulte-Herbrüggen, Lie-semigroup structures for reachability and control of open quantum systems: Kossakowski-Lindblad generators form Lie wedge to Markovian channels, *Rep. Math. Phys.* **64**, 93 (2009).

- [11] S. Kallush, M. Khasin, and R. Kosloff, Quantum control with noisy fields: Computational complexity versus sensitivity to noise, *New J. Phys.* **16**, 015008 (2014).
- [12] M. M. Müller, S. Gherardini, T. Calarco, S. Montangero, and F. Caruso, Information theoretical limits for quantum optimal control solutions: Error scaling of noisy control channels, *Sci. Rep.* **12**, 21405 (2022).
- [13] R. L. Kosut and H. Rabitz, Robust quantum control for set-membership Hamiltonian uncertainty, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Bellevue, WA, USA (IEEE, New York, NY, 2023), Vol. 1, pp. 1304–1307.
- [14] C. Altafini, Coherent control of open quantum dynamical systems, *Phys. Rev. A* **70**, 062321 (2004).
- [15] S. Kallush, R. Dann, and R. Kosloff, Controlling the uncontrollable: Quantum control of open-system dynamics, *Sci. Adv.* **8**, eadd0828 (2022).
- [16] Z. Yang, R. L. Kosut, and K. B. Whaley, Hamiltonian switching control of noisy bipartite qubit systems, *New J. Phys.* **26**, 023051 (2024).
- [17] H. N. Nguyen, F. Motzoi, M. Metcalf, K. B. Whaley, M. Bukov, and M. Schmitt, Reinforcement learning pulses for transmon qubit entangling gates, *Mach. Learn.: Sci. Technol.* **5**, 025066 (2024).
- [18] N. Funcke and J. Berberich, Robustness of optimal quantum annealing protocols, *New J. Phys.* **26**, 093040 (2024).
- [19] M. A. Nielsen, A geometric approach to quantum circuit lower bounds, *Quantum Inf. Comput.* **6**, 213 (2006).
- [20] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, Optimal control, geometry, and quantum computing, *Phys. Rev. A* **73**, 062323 (2006).
- [21] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, Quantum computation as geometry, *Science* **311**, 1133 (2006).
- [22] M. R. Dowling and M. A. Nielsen, The geometry of quantum computation, *Quantum Inf. Comput.* **8**, 861 (2008).
- [23] X. Han and P. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution* (Springer, 2017), Vol. 85, p. 12.
- [24] J. L. Strand, Random ordinary differential equations, *J. Diff. Equ.* **7**, 538 (1970).
- [25] N. Lambert, E. Giguère, P. Menczel, B. Li, P. Hopf, G. Suárez, M. Gali, J. Lishman, R. Gadhvi, R. Agarwal, A. Galicia, N. Shammah, P. Nation, J. R. Johansson, S. Ahmed, S. Cross, A. Pitchford, and F. Nori, QuTiP 5: The quantum toolbox in Python, [arXiv:2412.04705](https://arxiv.org/abs/2412.04705).
- [26] P. Martinez-Azcona, A. Kundu, A. Saxena, A. del Campo, and A. Chenu, Quantum dynamics with stochastic non-Hermitian Hamiltonians, *Phys. Rev. Lett.* **135**, 010402 (2025).
- [27] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, 1/f noise: Implications for solid-state quantum information, *Rev. Mod. Phys.* **86**, 361 (2014).
- [28] A. R. Brown and L. Susskind, Complexity geometry of a single qubit, *Phys. Rev. D* **100**, 046020 (2019).
- [29] E. Meinrenken, Lie groups and Lie algebras, [arXiv:2201.09397](https://arxiv.org/abs/2201.09397).
- [30] E. Meinrenken, *Clifford Algebras and Lie Theory* (Springer, 2013).
- [31] K. Shizume, T. Nakajima, R. Nakayama, and Y. Takahashi, Quantum computational Riemannian and sub-Riemannian geodesics, *Prog. Theor. Phys.* **127**, 997 (2012).
- [32] L. B. Gaggioli and R. Lawrence, Numerical study of the random Schrödinger equation, GitHub (2025), <https://github.com/llorebaga/RandomSE>.
- [33] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, *SIAM Rev.* **59**, 65 (2017).
- [34] P. K. Mogensen and A. N. Riseth, Optim: A mathematical optimization package for Julia, *J. Open Source Software* **3**, 615 (2018).