Control of noisy quantum systems: Field-theory approach to error mitigation

Rafael Hipolito[*] and Paul M. Goldbart

School of Physics, Georgia Institute of Technology, 837 State Street, Atlanta, Georgia 30332, USA

We consider the basic quantum-control task of obtaining a target unitary operation (i.e., a quantum gate) via control fields that couple to the quantum system and are chosen to best mitigate errors resulting from time-dependent noise, which frustrate this task. We allow for two sources of noise: fluctuations in the control fields and fluctuations arising from the environment. We address the issue of control-error mitigation by means of a formulation rooted in the Martin-Siggia-Rose (MSR) approach to noisy, classical statistical-mechanical systems. To do this, we express the noisy control problem in terms of a path integral, and integrate out the noise to arrive at an effective, noise-free description. We characterize the degree of success in error mitigation via a fidelity metric, which characterizes the proximity of the sought-after evolution to ones that are achievable in the presence of noise. Error mitigation is then best accomplished by applying the optimal control fields, i.e., those that maximize the fidelity subject to any constraints obeyed by the control fields. To make connection with MSR, we reformulate the fidelity in terms of a Schwinger-Keldysh (SK) path integral, with the added twist that the "forward" and "backward" branches of the time contour are inequivalent with respect to the noise. The present approach naturally and readily allows the incorporation of constraints on the control fields—a useful feature in practice, given that constraints feature in all real experiments. In addition to addressing the noise average of the fidelity, we consider its full probability distribution. The information content present in this distribution allows one to address more complex questions regarding error mitigation, including, in principle, questions of extreme value statistics, i.e., the likelihood and impact of rare instances of the fidelity and how to harness or cope with their influence. We illustrate this MSR-SK reformulation by considering a model system consisting of a single spin-s freedom (with s arbitrary), focusing on the case of $1/f$ noise in the weak-noise limit. We discover that optimal error mitigation is accomplished via a universal control field protocol that is valid for all s, from the qubit (i.e., $s = 1/2$) case to the classical (i.e., $s \to \infty$) limit. In principle, this MSR-SK approach provides a transparent framework for addressing quantum control in the presence of noise for systems of arbitrary complexity.

I. INTRODUCTION

The ability to control the fate of quantum systems whose dynamics are subject to noisy influences is a critical factor in numerous settings, including, notably, quantum information processing [1–3]. A typical characteristic of such systems is that in order to *control* them it is necessary to couple them to external time-varying fields that cannot themselves be perfectly controlled and which, therefore, inevitably introduce classical noise into the system. This noise, present in the external fields, alters the time evolution of the quantum system of interest (i.e., the system we wish to control), typically pushing the end result away from the intended target. The external control fields enter the Hamiltonian governing the dynamics of the quantum system by way of coupling directly to the system's degrees of freedom, and thus coupling these sources of noise directly to it. The dynamics, including the contribution from the noisy external fields, can be described via a Hamiltonian containing external *stochastic* parameters in addition to terms describing internal system dynamics. Although the probabilistic properties of the stochastic parameters can be determined, it is generally impossible to predict the values realized in any instance of the control attempt. Even in the absence of these external control fields, quantum systems are still generally subject to external sources of noise, because no system is truly isolated from the environment. Environmental noise is detrimental to the control

mission because it inevitably leads to the unremediable loss of information from the quantum system, due to entanglement between system and environment degrees of freedom, leading to quantum decoherence [4].

In the light of these remarks it is of value to identify and understand how to design schemes capable of mitigating the effects of noise to the greatest extent possible. Several approaches have been developed to address this issue, including dynamical decoupling [5], dynamical control by modulation [6,7] and, more recently, the filter function approach used in Refs. [8,9]. The aim of the present paper is to develop an alternative approach to the task of quantum control in the presence of noise. The approach is rooted in the Schwinger-Keldysh (henceforth SK) path-integral framework [10,11], which has frequently been invoked in treatments of quantum dynamical systems that are not in thermodynamic equilibrium; see, e.g., Refs. [12,13]. The SK framework is especially well suited for providing a transparent account of the effects of quantum fluctuations (i.e., *quantum noise*) via interference between quantum fields that propagate "forward" and "backward" in time. Although the framework was originally developed with *closed* quantum systems in mind, it can readily be modified to account for open quantum systems that are coupled to external fields and sources of noise.

The core idea behind the present approach is as follows. We consider noisy quantum systems for which we possess: (i) a complete characterization of the noise-free system via the specification of its freedoms and a Hamiltonian that governs them; and (ii) a complete *statistical* characterization of the noise that perturbs the dynamics from its noise-free form, in

---

[*]rafael.hipolito@physics.gatech.edu

the form of a probability distribution for the history of the noise parameters. We take the goal of the control process to be to guide the system as accurately as possible, i.e., to impart upon it (as accurately as possible) some predetermined unitary transformation, without being informed about the instance of the noise-parameter history. We characterize the performance of the control process—i.e., its ability to impart a predetermined target operation upon the quantum system of interest—via an overlap metric or *fidelity*, which is designed to assess the accuracy of the process, averaged over the noise-parameter history weighted by its known distribution. The guiding of the system is accomplished via a time-dependent Hamiltonian that we select from some menu; typically the menu is incomplete, in the sense that only certain operators are regarded as being available and the time dependence of the classical variables that characterize these operators is restricted by the kinds of constraints present in real experiments. An inevitable consequence of the present framework is that the guiding that will be ascertained will be more accurate for some noise histories and target unitaries than for others. A strength, however, is that it need only be determined once, for any given noisy system and noise distribution.

We formulate the task of controlling a generic noisy quantum system via an *optimization* problem, in which one seeks the control-field history that maximizes the fidelity, i.e., the measure of success alluded to earlier; see also Refs. [8,9] (although the present definition of the fidelity differs slightly from those used in these references). The path-integral formulation of the fidelity (which is the object of primary interest to us) takes a form that is close to the conventional SK path integral but has the following twist: The "'forward-in-time" and "backward-in-time" branches are asymmetrical with respect to the *external* noise (both environmental and due to the controls), which is present in the former branch but absent from the latter. This is a direct consequence of the definition we have chosen for the fidelity. By contrast, the internal *quantum noise*, which originates within the system of interest still appears symmetrically in the two branches, as it does in the usual SK path integral. The SK formulation of control is similar in spirit to the approach introduced by Martin-Siggia-Rose (MSR) [14] to study classical statistical dynamics. Once we have constructed the path integral for our SK-type formulation of the fidelity, we integrate out the environmental noise and thus arrive at an effective description that is completely deterministic (i.e., noise free), to which we can apply the tools available from field theory, such as diagrammatic expansion and even nonperturbative methods.

Via our approach, we show that the optimization procedure used to maximize the fidelity naturally gives rise to an action principle, which leads to equations of motion for the control-field history. The solution of these equations is a continuous deformation of the control scheme relevant to the noise-free case. As one continuously increases the strength of the noise, the optimal control scheme continuously deforms, parametrically, away from the noise-free scheme. The noise-free scheme typically presents an arbitrary number of schemes to choose from. By adopting any one of these schemes and continuously increasing the strength of the noise, we sweep through a continuous *family* of control schemes, with members of the family being parametrized by the noise strength.

Different choices of noise-free schemes typically correspond to different sets of winding numbers, as we show below. Each family can then be labeled by its set of winding numbers, which are invariant under these continuous deformations. In common with more familiar action principles, the present approach has the advantage of allowing the straightforward implementation of many classes of constraints, including those naturally appearing in experiments, via the Lagrange multiplier technique. The path-integral formalism also provides a natural starting point for developing a semiclassical approach to the task of controlling noisy quantum systems, in which quantum effects are introduced as a refinement of an underlying classical process. The literature is extensive on controlling noisy *classical* systems, as it is on the control of two-level quantum systems (i.e., qubits), which constitute the extreme quantal case. *Inter alia*, the formalism we develop here serves to bridge the gap between these two extremes—a regime that has been left relatively unexplored, to date.

Finally, we address the *full probability distribution* of the fidelity, as induced by the noise. This is a question that to the best of our knowledge has not been focused on, to date, in the literature on error mitigation in quantum control scenarios. To obtain this distribution, we make use of a standard route (see, e.g., Ref. [15]), which involves first computing the *characteristic function* corresponding to the fidelity distribution, from which all moments and cumulants of the fidelity follow. With full knowledge of the characteristic function, it is conceptually a simple matter of inversion to find the probability distribution; for the case of weak noise, we obtain an explicit analytical expression. Having the full probability distribution allows one, in principle, to address more complicated questions, such as how to minimize the likelihood of rare (but perhaps catastrophic) events in the control mission, as well as related questions from the realm of extreme value statistics (see, e.g., Ref. [16]); we leave such questions to future work.

We illustrate our approach by analyzing some concrete examples in detail. We specifically develop the formalism for a single quantum spin $\hat{\mathbf{S}}$, keeping the spin quantum number s arbitrary. We couple the spin to external control fields in order to achieve a target operation, and allow the spin to be under the influence of noise (from the environment or the control fields or both) with the statistics of the noise presumed known. Although we develop the formalism with this specific system in mind, we note that it can be readily generalized to more elaborate systems, including interacting systems such as spin chains and ultracold atomic gases. We also note that the SK path integral, if formulated in terms of coherent states, provides a natural starting point for a semiclassical expansion. This is a useful feature, if one is interested in studying control problems for systems where the semiclassical expansion gives an excellent approximation to the full quantum dynamics (e.g., ultracold bosonic gases [17]).

Continuing with the case of a single spin, we use our formalism to find the fidelity and the corresponding optimized control fields for various choices of noise distributions, with a special focus on $1/f$ noise sources, which arise in many systems of interest [18]. For the specific case of a single spin, we find results for arbitrary s, ranging from the qubit limit ($s = 1/2$) up to the classical limit $s \to \infty$.

The paper is organized as follows. In Sec. II we give a discussion of the general setting that we shall be concerned with, and also review the various questions that we shall be exploring in detail. In Sec. III we formulate the control problem in terms of a modifed SK approach, and apply it to the case of a single spin $\hat{\mathbf{S}}$ of arbitrary spin quantum number s in the presence of noise. We also construct the expression for the fidelity and find the optimal control scheme for the case of weak $1/f$ noise and constraints being imposed on the strength of the control fields. Finally, we find the expression for the full probability distribution associated with the fidelity for this specific case. In Sec. IV we summarize our results and briefly discuss future directions. The technical details for the derivation of our formalism are mostly relegated to appendixes for the sake of the clarity of the presentation, and will be referred to where relevant.

II. ELEMENTS

Our task is to control a given quantum system—specifically, we wish to complete a *predetermined unitary transformation*, which we call the *target unitary operator* U_T, upon the quantum state. The state is not necessarily known ahead of time. In order to accomplish this, we invoke external time-dependent *control fields*, which couple directly to the system degrees of freedom and are used to steer the quantum system in such a way that, at the end of some transit time τ, the quantum system has evolved to a final state that is equivalently described by U_T in the sense that the *net effect* of the two is the same.

The quantum system we wish to control is not generally isolated. It is coupled to an external environment and is thus subject to environmental noise. In addition, there may be fluctuations in the control fields themselves (i.e., originating in the devices used to generate these fields), which provide another source of noise. The net effect originating from all sources of noise will generally interfere with the control scheme. Assuming one chooses the control fields such that one obtains a unitary equivalent to U_T at the end of the transit time τ *in the absence of noise*, these same control fields will generally give rise to a unitary transformation that differs from U_T.

The question we address can be stated as follows: Given many histories of *sets* of control fields to choose from, each of which would give rise to U_T in the absence of noise, which particular *set* gives rise to a unitary transformation that is *closest* to U_T? In the presence of noise, it is not possible to predict how the quantum system that we aim to control will evolve in each given instance of the noise field history. The best we can do is determine with a *statistical* measure that tells us how close we are to reaching our stated goal, which is to have unitary evolution that is as close as possible to U_T at the end of the transit time τ.

The simplest such measure, viz., the *fidelity* (which we define below), reports how close we get *on average*, i.e., after averaging over all sources of noise. In other words, we choose a *single* set of control fields; then, for each instance of noise, we find the corresponding unitary transformation at the end of the transit time τ; we *average* these unitary transformations over all instances of noise histories; and, finally, we compare

the resulting averaged unitary transformation with U_T. We define the best single set of control field histories—what we are searching for—to be the set from which one obtains the (averaged) transformation that lies as close as possible to the one resulting from U_T. We rigorously define all of the quantities of interest below, as we develop our methodology.

Before developing our methodology, let us spend some time to obtain a better quantitative understanding of the problem at hand in terms of the various Hamiltonian terms involved.

A. Hamiltonian

Consider a quantum system described by a Hamiltonian $H(t)$, which can be written as a sum of several parts:

$$H(t) = H_c(t) + H_s(t) + H_e(t) + \epsilon_{se} H_{se}(t); \qquad (1)$$

we allow all parts of the Hamiltonian to be explicitly time dependent, and $H_c(t)$, $H_s(t)$, $H_e(t)$, and $H_{se}(t)$ are, respectively, taken to be the control Hamiltonian, system Hamiltonian, environmental Hamiltonian, and system-environment coupling Hamiltonian. The parameter ϵ_{se} determines the strength of the system-environment coupling, which we take to be weak.

Let us denote by $\{\hat{q}_i\}$ the quantum degrees of freedom of the system (i.e., the part that we wish to *control*), and by $\hat{Q}_i$ the degrees of freedom of the environment. The various Hamiltonian terms have the following characteristics: $H_c(t) = \sum_i c_i(t) \hat{q}_i$ couples the *control fields* $c_i(t)$ directly to the quantum degrees of freedom $\hat{q}_i$, which we are interested in controlling; $H_s(t) = H_s(t, \hat{q}_i)$ determines the internal system dynamics; $H_e(t) = H_e(t, \hat{Q}_i)$ determines the dynamics of the environment; and $H_{se}(t) = H_{se}(t, \hat{q}_i, \hat{Q}_i)$ couples the system and environment to one another. From the perspective of RG, the terms that linearly couple the system and environment degrees of freedom are the most relevant (see, e.g., Ref. [19]). Taking the system-environment coupling to be weak (i.e., ϵ_{se} small), and temporarily ignoring all other terms in $H_{se}(t)$ we get

$$H_{se}(t) \cong \sum_{ij} \hat{q}_i \, \eta_{ij}(t) \, \hat{Q}_j. \qquad (2)$$

Next, by using the path integral language we can proceed to *integrate out* the environment degrees of freedom, and thus obtain an effective Hamiltonian, in which only the system degrees of freedom remain. Carrying out this procedure, and assuming that one can treat the environmental degrees of freedom using a semiclassical approximation, we obtain an effective system-environment Hamiltonian $\widetilde{H}_{se}(t)$ which in general has the following form:

$$\widetilde{H}_{se}(t) \simeq \sum_i \tilde{\eta}_i(t) \, \hat{q}_i, \qquad (3)$$

where the histories $\{\tilde{\eta}_i(\cdot)\}$ are *stochastic*, with a histories functional probability density $P_e[\{\tilde{\eta}_i(\cdot)\}]$ (where $\{\tilde{\eta}_i(\cdot)\}$ designates the set of $\tilde{\eta}_i(\cdot)$ for all i), which in principle can be determined from the state of the environment and the environment Hamiltonian H_e.

In actual experiments, there is also some degree of uncertainty associated with the control fields $c_i(t)$ themselves, such as fluctuations in the fields due to noise generated in the experimental apparatus responsible for the generation of these

fields. In general, we can write

$$c_i(t) = \tilde{c}_i(t) + \epsilon_c \, \delta c_i(t), \qquad (4)$$

where $\tilde{c}_i(t)$ is the control field as given in the absence of any fluctuations, the fluctuations $\delta c_i(t)$ are described by some probability distribution $P_c(\{\delta c_i(\cdot)\})$, and ϵ_c is a parameter describing the strength of fluctuations. As experimental measurements are sensitive to the *net* contribution of noise, we combine the environmental noise and noise due to fluctuations in the control fields into a single effective noise term, $H_n(t)$, given by

$$\epsilon H_n(t) = \epsilon \sum_i n_i(t)\, \hat{q}_i, \qquad (5)$$

where $\epsilon \equiv \epsilon_{se} + \epsilon_c$ and

$$n_i(t) \equiv \frac{\epsilon_{se}}{\epsilon_{se} + \epsilon_c} \tilde{\eta}_i(t) + \frac{\epsilon_c}{\epsilon_{se} + \epsilon_c} \delta c_i(t), \qquad (6)$$

so that $\{n_i(\cdot)\}$ are effective stochastic fields described by a (classical) effective probability distribution $P[\{n_i(\cdot)\}]$, which can in principle be determined via experiment. In what follows, we assume that we know $P[\{n_i(\cdot)\}]$.

Let us replace $\tilde{c}_i$ by c_i, where c_i is now understood to be the control field. We can now specify the effective Hamiltonian, which we will refer to as $H_\epsilon(t)$, so as to remind us of its dependence on ϵ [note this Hamiltonian is different from Eq. (1)]:

$$H_\epsilon(t) = H_s(t,\hat{q}_i) + \sum_i [c_i(t) + \epsilon n_i(t)]\hat{q}_i$$
$$\equiv H_s(t) + H_c(t) + \epsilon H_n(t). \qquad (7)$$

This effective Hamiltonian H_ϵ contains the full quantum description of the system degrees of freedom $\hat{q}_i$, which are coupled to both the *stochastic* fields $n_i(t)$ and the *control* fields $c_i(t)$, as well as to one another via the internal system dynamics as described by H_s.

B. Fidelity

We now construct the *fidelity*, i.e., a measure of success in effecting the specified unitary transformations, averaged over the noise history. At the end of the transit time τ, our aim is to have the system evolution to be as close as possible to U_T. The evolution operator $U_\epsilon(t,\{c_i(\cdot)\},\{n_i(\cdot)\})$ corresponding to H_ϵ, cf. Eq. (7), is given by

$$U_\epsilon(\tau,\{c_i(\cdot)\},\{n_i(\cdot)\}) = \mathcal{T} \exp\left[-i \int_0^\tau dt\, H_\epsilon(t)\right], \qquad (8)$$

where the notation reminds us that U_ϵ is a function of τ as well as a *functional* of the control and effective noise fields. We shall often use the shorthand $U_\epsilon(t)$, provided there is no risk of confusion. The quantity $U_0(t) \equiv U_\epsilon(t)|_{\epsilon=0}$ corresponds to the evolution operator in the absence of noise. Recall that at the transit time τ, we have $U_0(\tau) = U_T$ (i.e., the target unitary). We elect to define the fidelity $\mathcal{F}$ as follows:

$$\mathcal{F}(\{c_i(\cdot)\}) = \langle\, \mathrm{Tr}\, U_T^\dagger\, U_\epsilon(\tau)\, \rangle_n \qquad (9a)$$
$$= \langle\, \mathrm{Tr}\, U_0^\dagger(\tau) U_\epsilon(\tau)\rangle_n, \qquad (9b)$$

where the brackets $\langle \cdots \rangle_n$ denote averaging over $\{n_i(\cdot)\}$, the trace operation is taken over the entire Hilbert space of the system, normalized by the dimension of the Hilbert space so that $\mathrm{Tr}\, \mathbf{1} = 1$. Observe that the fidelity obeys $0 \leqslant |\mathcal{F}| \leqslant 1$ and reaches its upper bound of unity if and only if $U_\epsilon = U_T$ (i.e., perfect control is achieved regardless of initial state). The fidelity is a *functional* of the control fields $\{c_i(\cdot)\}$. In order to best accomplish the sought for unitary transformation, we solve the *variational* problem to find the set $\{c_i(\cdot)\}$ that *maximizes* $\mathcal{F}$. We shall make use of the formulation given in the second line of Eq. (9b), which turns out to be efficacious when expressed in terms of a path integral. In terms of the Hamiltonian, the fidelity is given by

$$\mathcal{F}[\{c_i(\cdot)\}] = \left\langle \mathrm{Tr}\, \mathcal{T}_K\, e^{i \int_0^\tau dt\, H_0(t)}\, e^{-i \int_0^\tau dt\, H_\epsilon(t)} \right\rangle_n. \qquad (10)$$

Here the operation $\mathcal{T}_K$ corresponds to time ordering on the Keldysh contour (see, e.g., Ref. [12]): The first exponential factor is *anti*-time ordered; the second one is time ordered in the usual way. Equation (10) can be expressed in terms of a Schwinger-Keldysh path integral over a closed-time contour, but there is a twist: Unlike the mere usual quantum-dynamical problems (e.g., for isolated quantum systems), for which this technique is commonly applied, in the present setting there is an *asymmetry* between the forward and backward branches of the time contour. Specifically, the stochastic degrees of freedom $\{n_i(\cdot)\}$ are only present in the forward branch.

As we shall see, for some purposes, it is useful to refer to the *fidelity amplitude* $\mathcal{A}$ instead of the fidelity. This is defined via

$$\mathcal{A}[\{c_i(\cdot)\},\{n_i(\cdot)\}] = \mathrm{Tr}\, \mathcal{T}_K\, e^{i \int_0^\tau dt\, H_0(t)}\, e^{-i \int_0^\tau dt\, H_\epsilon(t)}, \qquad (11)$$

which is just the expression for the fidelity *before* taking the average over the stochastic fields, and thus is a functional of *both* $\{c_i(\cdot)\}$ and $\{n_i(\cdot)\}$. In terms of $\mathcal{A}$ we have

$$\mathcal{F}[\{c_i(\cdot)\}] = \langle \mathcal{A}(\{c_i(\cdot)\},\{n_i(\cdot)\}) \rangle_n. \qquad (12)$$

C. Constraints

Determining the control sequence that best mitigates noise amounts to finding the set of control fields $\{c_i(\cdot)\}$ that maximize $\mathcal{F}$ and yet satisfy all *constraints* imposed on the set $\{c_i(\cdot)\}$ *together with* the boundary condition at the end of the transit time τ, viz. $U_0(\tau) = U_T$. In any realistic situation, there will be physical limitations on $\{c_i(\cdot)\}$ that we may consider. For instance, each $c_i(t)$ must be of *finite magnitude*, and its functional dependence on time would be constrained by the experimental devices in use.

There are certain classes of constraints that can be entirely accounted for by using Lagrange multipliers. These include, but are not limited to, holonomic constraints. Explicit examples will be worked out in the following sections.

D. Probability distribution for the fidelity amplitude

In the previous subsections, we have focused on developing methods to determine $\mathcal{F}$, the fidelity, which is simply the *noise average* of the fidelity amplitude $\langle \mathcal{A} \rangle_n$, taken with respect to the stochastic noise fields $\{n_i(\cdot)\}$. In applications, it may be of interest to compute higher moments of $\mathcal{A}$, and even its full probability distribution function. For instance, consider a

typical problem in extreme value statistics (see, e.g., Ref. [16]), where one may be interested in finding a control protocol that reduces the likelihood of highly detrimental events, however unlikely they may be, since even a single occurrence may have unacceptable consequences for the control task. In order to find the likelihood of extreme events, it is sufficient to have the full probability distribution.

In this subsection, we will briefly lay out the general procedure that we shall use to compute both the characteristic function and the full probability distribution. The formalism developed here makes use of standard methods, i.e., as seen in Ref. [15].

In the case of weak noise ($\epsilon \ll 1$), as far as the full probability distribution is concerned, the physics is clearer when working with the *infidelity amplitude* $\mathcal{I}$, defined as

$$\mathcal{I} = 1 - \mathcal{A}, \tag{13}$$

in terms of which the *infidelity* I is given by

$$I = \langle \mathcal{I} \rangle_n. \tag{14}$$

Clearly, with these definitions, we can find the fidelity (which is just an average) via the simple relation,

$$\mathcal{F} = 1 - I. \tag{15}$$

The optimal control-field histories, as we have defined them, will then *minimize* the infidelity I (subject to constraints—see previous subsection); this amounts to getting I as close as possible to fully vanishing.

The first step we take in obtaining the probability distribution $P(\mathcal{I})$ for the infidelity amplitude is to compute the *characteristic function* $\chi(k)$, which is defined via

$$\chi(k) = \langle e^{-ik\mathcal{I}[\{c_i(\cdot)\},\{n_i(\cdot)\}]} \rangle_n, \tag{16}$$

where the notation reminds us that $\mathcal{I}$ is a functional of both the controls $\{c_i(\cdot)\}$ and the noise $\{n_i(\cdot)\}$.

The characteristic function χ is an interesting object in its own right, as it allows us to easily compute all the *moments* of $\mathcal{I}$, via

$$\langle \mathcal{I}^m \rangle_n = (i\,\partial_k)^m \chi(k)|_{k=0}, \tag{17}$$

and it is also straightforward to calculate all the cumulants using the expression,

$$\langle\langle \mathcal{I}^m \rangle\rangle_n = (i\,\partial_k)^m \ln \chi(k)|_{k=0}. \tag{18}$$

Once we have $\chi(k)$, the full probability distribution $P(\mathcal{I})$ can be obtained in terms of the Fourier integral,

$$P(\mathcal{I}) = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\,\chi(k)e^{ik\mathcal{I}}. \tag{19}$$

Specific applications of this formalism and the results obtained with it will be given in detail in the following section, where we will be working with a specific system (a single spin in the presence of time-dependent noise).

We note here a tendency in the behavior of $P(\mathcal{I})$, which while not generic, does show up in the examples we consider. Since the values of $|\mathcal{I}|$ are constrained to lie in the interval $0 \leqslant |\mathcal{I}| \leqslant 1$, we find that when minimizing the *average* I, we simultaneously minimize the *width* of $P(\mathcal{I})$.

III. SINGLE SPIN $\hat{\mathbf{S}}$

We now spell out explicitly how the abstract formalism presented in Sec. II applies in the case of a single spin $\hat{\mathbf{S}}$, for now leaving the spin quantum number s arbitrary. In this setting, the total Hamiltonian is given by

$$H_\epsilon(t) = (\boldsymbol{\omega}(t) + \epsilon \mathbf{n}(t)) \cdot \hat{\mathbf{S}}, \tag{20}$$

where the field $\boldsymbol{\omega}(t)$ corresponds to the external control field and $\mathbf{n}(t)$ is the stochastic field, discussed in Sec. II, that represents the effect of environmental noise and also accounts for any fluctuations in the control field inherently present in the devices used to generate it. In addition, $\boldsymbol{\omega}(t)$ is the control in the absence of such fluctuations. In the present example, the system Hamiltonian, $H_s(t)$, vanishes.

Our goal, then, is to determine the $\boldsymbol{\omega}(t)$ that maximizes the fidelity $\mathcal{F}$ at the end of the transit time τ. In other words, we seek $\boldsymbol{\omega}(t)$ such that, after averaging over different realizations of $\mathbf{n}(t)$, the evolution operator is as close as possible to some prescribed target U_T.

As discussed in Sec. II, the stochastic fields $\{n_i(\cdot)\}$ are governed by a distribution functional $P[\{n_i(\cdot)\}]$ which is presumed to be known. In the present case, we take P to be Gaussian with zero mean and covariance given by

$$\langle n_i(t)n_j(t) \rangle_n = \mathcal{N}_{ij}(t,t'). \tag{21}$$

We restrict our attention to stochastic fields that are stationary in time and time-reversal invariant, in which case,

$$\mathcal{N}_{ij}(t,t') = \mathcal{N}_{ij}(|t - t'|) = \mathcal{N}_{ji}(|t - t'|). \tag{22}$$

We make use of the Schwinger-Keldysh (SK) path-integral formulation (see, e.g., Refs. [10,11] for original work by Schwinger and Keldysh and Ref. [12] for a modern treatment), which for the present system can be done either in terms of spin coherent states or bosonic coherent states (if one makes use of the Schwinger representation for spins; see, e.g., Ref. [20]. Choosing the latter, we make use of the mapping between the spin operator $\hat{\mathbf{S}}$ and the two-component bosons $\hat{\mathbf{a}}^\dagger \equiv (a_1^\dagger, a_2^\dagger)$, given by

$$\hat{\mathbf{S}} = \tfrac{1}{2}\hat{\mathbf{a}}^\dagger \cdot \boldsymbol{\sigma} \cdot \hat{\mathbf{a}}, \tag{23}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. The total spin quantum number s is conserved by the dynamics, and is related to the total number of bosons in the Schwinger representation: $s = \tfrac{1}{2}\mathbf{a}^\dagger \cdot \mathbf{a}$. The total Hamiltonian then takes the form,

$$H_\epsilon(t) = \mathbf{a}^\dagger \cdot \mathcal{H}_\epsilon(t) \cdot \mathbf{a}, \tag{24}$$

where

$$\mathcal{H}_\epsilon = \tfrac{1}{2}\boldsymbol{\omega}(t) \cdot \boldsymbol{\sigma} + \frac{\epsilon}{2}\mathbf{n}(t) \cdot \boldsymbol{\sigma}$$
$$\equiv \mathcal{H}_c(t) + \epsilon \mathcal{H}_n(t). \tag{25}$$

A. Schwinger-Keldysh (SK) path integral

We now construct the path-integral expression for the fidelity amplitude $\mathcal{A}_s$, Eq. (11), where s indicates the (arbitrary) spin quantum number. We make use of the coherent state basis $|\boldsymbol{\alpha}\rangle \equiv |\alpha_1, \alpha_2\rangle$, for complex α_1 and α_2, defined by $\mathbf{a}|\boldsymbol{\alpha}\rangle = \boldsymbol{\alpha}|\boldsymbol{\alpha}\rangle$ (see Ref. [20]), in order to evaluate the SK path integral along the SK contour. We use the labels $\boldsymbol{\alpha}_{f,b}$ for

the forward-in-time and backward-in-time branches along the contour, respectively. We thus obtain the expression,

$$
\mathcal{A}_s = \text{Tr} \int D^2 \boldsymbol{\alpha}_f(t) D^2 \boldsymbol{\alpha}_b(t) \times |\boldsymbol{\alpha}_b(0)\rangle
$$
$$
\times e^{-\frac{1}{2}[|\boldsymbol{\alpha}_b(\tau)|^2 + |\boldsymbol{\alpha}_f(0)|^2] + \boldsymbol{\alpha}_b^\star(\tau) \cdot \boldsymbol{\alpha}_f(\tau)}
$$
$$
\times e^{i\int_0^\tau dt\, [\boldsymbol{\alpha}_f^\star \cdot (i\partial_t - \mathcal{H}_\epsilon) \cdot \boldsymbol{\alpha}_f - \boldsymbol{\alpha}_b^\star \cdot (i\partial_t - \mathcal{H}_c) \cdot \boldsymbol{\alpha}_b]} \langle \boldsymbol{\alpha}_f(0)|, \quad (26)
$$

where the normalized trace Tr is taken over the complete set of two-mode bosonic number states $|n, 2s - n\rangle_{\text{num}}$ for $n = 0, 1, \ldots, 2s$, satisfying the constraint $\langle n, 2s - n|\frac{1}{2}\hat{\mathbf{a}}^\dagger \cdot \hat{\mathbf{a}}|n, 2s - n\rangle_{\text{num}} = s$, so as to fix the quantum spin number s. For further details on the meaning of Eq. (26), as well as explicit expressions for the measures $D^2\boldsymbol{\alpha}_{b,f}$ and other details concerning the path integral, we refer the reader to Appendix C. Note that our approach differs from the conventional SK approach [12] used to study nonequilibrium quantum dynamics, in that the forward and backward branches of the time evolution are asymmetric with respect to noise: It is present in the forward branch and completely absent in the backward branch.

The expression for A_s in Eq. (26) can be evaluated with the help of a generating functional $G_s[\mathbf{J}]$, which we define as follows:

$$
G_s[\mathbf{J}] = \text{Tr} \int D^2\boldsymbol{\alpha}_f(t) D^2\boldsymbol{\alpha}_b(t) \times |\boldsymbol{\alpha}_b(0)\rangle \langle \boldsymbol{\alpha}_f(0)|
$$
$$
\times e^{-\frac{1}{2}[|\boldsymbol{\alpha}_b(\tau)|^2 + |\boldsymbol{\alpha}_f(0)|^2] + \boldsymbol{\alpha}_b^\star(\tau) \cdot \boldsymbol{\alpha}_f(\tau)}
$$
$$
\times e^{i\int_0^\tau dt\, [\boldsymbol{\alpha}_f^\star \cdot (i\partial_t - \mathcal{H}_c) \cdot \boldsymbol{\alpha}_f - \boldsymbol{\alpha}_b^\star \cdot (i\partial_t - \mathcal{H}_c) \cdot \boldsymbol{\alpha}_b]}
$$
$$
\times e^{\int_0^\tau dt\, [\mathbf{J}^\star \cdot \boldsymbol{\alpha}_f + \boldsymbol{\alpha}_f \cdot \mathbf{J}]}, \quad (27)
$$

where, as anticipated, this expression contains the *noise-free* Hamiltonian on both branches of the Keldysh contour. Moreover, we have coupled the two-component external source $\mathbf{J}(t)$ to the forward branch only. This allows us to calculate quantum averages involving the forward branch fields. Recall that noise couples exclusively to this branch. Observe that we have $G_s[0] = 1$ since in the absence of a source, G_s is just the usual SK partition function and the asymmetry due to the noise no longer arises. Physically, this corresponds to the feature that in the absence of noise the fidelity is exactly unity, as is natural.

In terms of $G_s[\mathbf{J}]$, we may evaluate the fidelity amplitude as follows:

$$
\mathcal{A}_s = e^{-i\int_0^\tau dt\, \frac{\delta}{\delta \mathbf{J}(t)} \cdot \mathcal{H}_n(t) \cdot \frac{\delta}{\delta \mathbf{J}^\star(t)}} G_s[\mathbf{J}(t)]|_{\mathbf{J}(t)=\mathbf{0}}. \quad (28)
$$

$G_s[\mathbf{J}]$ is evaluated in Appendix C, giving

$$
G_s[\mathbf{J}] = \oint \frac{dz}{2\pi i} \frac{e^{\int_0^\tau \int_0^\tau dt\, dt'\, \mathbf{J}^\star(t) \cdot \mathcal{G}_c(z,t,t') \cdot \mathbf{J}(t')}}{(2s+1)(1-z)^2 z^{2s+1}}, \quad (29)
$$

where the integral over the auxiliary complex variable z is taken over any closed contour that encircles the origin once, but does not include the pole at $z = 1$. The variable z plays the role of a conjugate variable to the discrete spin number s—this integral can be interpreted as an integral transform between the z representation and s representation.

In the exponent of Eq. (29) we have the Green function $\mathcal{G}_c(z,t,t')$, which is given by

$$
\mathcal{G}_c(z,t,t') \equiv \left(\frac{z}{1-z} + \Theta(t-t')\right) \mathcal{U}_c(t,t'), \quad (30)
$$

where $\Theta(t)$ is the Heaviside step function (zero and unity for negative and positive arguments, respectively), and $\mathcal{U}_c(t,t') \equiv \mathcal{T} \exp[-i\int_{t'}^t dt''\, \mathcal{H}_c(t'')]$ is the unitary matrix corresponding to the control Hamiltonian. Note that in this expression and formula (28) for $\mathcal{A}_s$, for an arbitrary spin s, we only need consider the dynamics of a two-level system (i.e., spin 1/2).

To make the meaning (and evaluation) of A_s more transparent, we can re-express Eq. (28) in terms of a two-component Gaussian quantum field $\boldsymbol{\phi}(t)^\dagger \equiv (\phi_1^\star(t), \phi_2^\star(t))$ whose description is completely given in terms of the expectation value:

$$
\langle \boldsymbol{\phi}(t) \boldsymbol{\phi}^\star(t')\rangle_q = \mathcal{G}_c(z,t,t'), \quad (31)
$$

where $\langle \cdots \rangle_q$ denotes a quantum average over $\boldsymbol{\phi}(t)$, and higher-order averages are determined by means of Wick's theorem. Thus,

$$
\mathcal{A}_s = \oint \frac{dz}{2\pi i} \frac{\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\star \mathcal{H}_n \boldsymbol{\phi}} \rangle_q}{(2s+1)(1-z)^2 z^{2s+1}}. \quad (32)
$$

Note the the expressions for $\mathcal{A}_s$ given in Eqs. (32) and (28) are entirely equivalent.

The quantum expectation value for the fields $\boldsymbol{\phi}(t)$ in Eq. (32) can be evaluated exactly with the use of diagrammatic techniques, which are developed in Sec. 2 of Appendix C. Through an application of these techniques, we obtain the expression:

$$
\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\star \mathcal{H}_n \boldsymbol{\phi}} \rangle_q = \frac{(1-z)^2}{\text{Det}\left[\mathbb{1} - z\mathcal{T} e^{-i\epsilon \int_0^\tau dt\, \mathcal{U}_c^\dagger \mathcal{H}_n \mathcal{U}_c}\right]}. \quad (33)
$$

Inserting this expression into Eq. (32) and evaluating the integral over z, we obtain the result for the fidelity amplitude $\mathcal{A}_s$. Recalling that $\mathcal{A}_s$ are functionals of the control $\boldsymbol{\omega}(t)$ and the noise $\mathbf{n}(t)$, we finally obtain

$$
\mathcal{A}_s[\boldsymbol{\omega}(\cdot), \mathbf{n}(\cdot)] = \frac{1}{2s+1} \sum_{j=-s}^s e^{-2ji\cos^{-1}(\mathcal{A}_{1/2}[\boldsymbol{\omega}(\cdot), \mathbf{n}(\cdot)])}, \quad (34)
$$

where we are emphasizing the functional dependence of $\mathcal{A}_s$ on $\boldsymbol{\omega}(\cdot)$ and $\mathbf{n}(\cdot)$. The quantity $\mathcal{A}_{1/2}[\boldsymbol{\omega}(\cdot), \mathbf{n}(\cdot)]$, appearing in Eq. (34), is the fidelity amplitude for the spin-half case. It is given by the expression,

$$
\mathcal{A}_{1/2}[\boldsymbol{\omega}(\cdot), \mathbf{n}(\cdot)] = \frac{1}{2} \text{Tr} \mathcal{T} e^{-i\int_0^\tau dt\, \mathcal{U}_c^\dagger(t)\mathcal{H}_n(t)\mathcal{U}_c(t)}, \quad (35)
$$

in which $\mathcal{T}$ denotes the usual time-ordering operator, τ is the transit time, and

$$
\mathcal{U}_c(t) = \mathcal{T} e^{-i\int_0^t dt'\, \mathcal{H}_c(t')} \quad (36)
$$

is the evolution operator corresponding to the control Hamiltonian. The derivation of Eqs. (33) and (34) is given in Sec. 2 of Appendix C.

B. Connection with quaternions

Before continuing with our development, let us pause to streamline our notation. To this end, we observe a connection with quaternions that makes the physics more transparent. The advantage stems from the fact that quaternions provide a unified method for treating both geometrical 3-vectors and unitary evolution operators, all under the language of quaternion algebra. In what comes below, the reader may freely

replace pure quaternions (see Appendix A) by vectors in all settings, except when they occur in exponents. In that case, the correct interpretation is to replace pure quaternions by the corresponding vector dotted with $-i\boldsymbol{\sigma}$, where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is a vector of Pauli matrices. The quaternion wedge and dot products may be freely replaced by the vector cross and dot products, respectively, and unit quaternions may be replaced by the corresponding unitary matrices, as seen below. For a full review of all quaternion properties and operations used in this paper, see Appendix A.

Before returning to the task of finding the fidelity and optimizing it, let us define the quantities $\mathbf{E}_i(t)$, given by

$$\mathbf{E}_i(t) \equiv -i\,\mathcal{U}_c^\dagger(t)\,\sigma_i\,\mathcal{U}_c(t), \tag{37}$$

where, recalling that $\mathcal{U}_c(t)$ is the unitary corresponding to the control Hamiltonian, $\mathbf{E}_i(t)$ are the Pauli matrices in the rotating frame of the control fields. Note that the $\mathbf{E}_i(t)$ satisfy the quaternion algebra, *regardless* of the form of the control fields, for all times t, i.e.,

$$\mathbf{E}_i(t)\mathbf{E}_j(t) = -\delta_{ij}\mathbb{1} + \epsilon_{ij}{}^k\,\mathbf{E}_k(t), \tag{38}$$

where the indices i, j, and k take on the values 1, 2, and 3, corresponding, respectively, to x, y, and z, and repeated indices are summed over. The symbol $\mathbb{1}$ denotes the 2×2 unit matrix. Let us introduce $\mathbf{E}_0(t) \equiv \mathbb{1}$; then evidently we have

$$\mathbf{E}_0(t)\mathbf{E}_i(t) = \mathbf{E}_i(t)\mathbf{E}_0(t) = \mathbf{E}_i(t) \tag{39}$$

for all i. The isomorphism between the set of quantities $\mathbf{E}_i(t)$ and quaternions is given by identifying $\mathbf{E}_i(t)$ (for $i = 1, 2, 3$), respectively, with the quaternion imaginary units $\hat{\imath}, \hat{\jmath}, \hat{k}$, and $\mathbf{E}_0(t)$ with the real number 1. Quaternion addition is defined component-wise, whereas multiplication is defined through Eqs. (38) and (39); note that quaternion multiplication is not commutative.

Given this identification, for calculational purposes it is simpler to work with quaternion quantities. Recall that any quaternion Q can be expressed in terms of components as $Q = Q^i\,\mathbf{E}_i$ with $\{Q^i\}$ real and i summed over the values $0, 1, 2, 3$. The component $Q^0 E_0$ is referred to as the *real*, or *scalar*, part of the quaternion, and the remainder is the *imaginary* part, also called the *vector* part, of the quaternion. Quaternions having vanishing scalar part are called *pure* quaternions. In what follows we will almost exclusively work with pure quaternions. For additional properties of quaternions, other types of operations (specifically, the dot and wedge products) and the terminology associated with quaternions, we refer the reader to Appendix A.

We refer to the set of quaternions $\mathbf{E}_i(t)$, restricting i to be 1, 2, or 3 as the *rotating triad*, and we refer to the set $\mathbf{e}_i$ given by

$$\mathbf{e}_i = -i\sigma_i, \tag{40}$$

as the *static triad*, because these are fixed in the laboratory frame. For convenience we represent *pure quaternions* using a bold font, i.e., $\mathbf{p}$, and they can be represented by a restricted sum over the indices, i.e., $\mathbf{p} = \sum_{i=1}^{3} p^i\,\mathbf{E}_i$ in the rotating basis and $\mathbf{p} = \sum_{i=1}^{3} p^i\,\mathbf{e}_i$ in the static basis. Unless indicated otherwise, from now on we interpret repeated indices as a summation over the set $i = 1, 2, 3$.

It is straightforward to find the relation between $\mathbf{E}_i(t)$ and $\mathbf{e}_i$ within the quaternion language. Associated with the evolution operator $\mathcal{U}_c(t)$ is a *unit* quaternion; for simplicity, we indicate unit quaternions by means of a regular font. If we take the pure quaternion $\boldsymbol{\Omega}(t)$ to represent the control Hamiltonian in the quaternion language, $\boldsymbol{\Omega}(t) \equiv \omega^i(t)\,\mathbf{E}_i(t)$, then the associated unit quaternion, corresponding to the unitary in the vector language, is given by

$$u_c(t) = \mathcal{T} e^{\frac{1}{2}\int_0^t dt'\,\boldsymbol{\Omega}(t')}, \tag{41}$$

and the relation between the rotating triad $\mathbf{E}_i(t)$ and the static triad $\mathbf{e}_i$ is given by

$$\mathbf{E}_i(t) = \bar{u}_c(t)\,\mathbf{e}_i\,u_c(t), \tag{42}$$

where $\bar{u}_c(t) \equiv \mathcal{T} e^{-\frac{1}{2}\int_0^t dt'\,\boldsymbol{\Omega}(t')}$ is the quaternion conjugate of $u_c(t)$ [corresponds to $\mathcal{U}_c^\dagger(t)$ in matrix language; see Appendix A], and the operation on the right-hand side is just quaternion multiplication. In quaternion language, Eq. (42) tells us that the $\mathbf{e}_i$ and $\mathbf{E}_i(t)$ are related via a pure rotation—in other words the rotating triad corresponds to a rigid rotation of the static triad.

Let us now continue with the evaluation of the fidelity, but now making use of the quaternion language. Equation (35) can be rewritten in terms of quaternions as

$$\mathcal{A}_{1/2}[\mathbf{E}(\cdot), n_i(\cdot)] = \mathrm{Sc}\,\mathcal{T} e^{\frac{\epsilon}{2}\int_0^\tau dt\,\mathbf{n}(t)}, \tag{43}$$

where $\mathbf{n}(t) = n^i(t)\mathbf{E}_i(t)$ is now interpreted as the *pure quaternion* representing the stochastic field in the rotating frame, $\mathcal{A}_{1/2}$ is now interpreted to be a functional of the rotating triad [we take $\mathbf{E}(t)$ as shorthand for the set of $\mathbf{E}_i(t)$] as well as the stochastic fields $\{n_i(\cdot)\}$, and the Sc operation simply takes the *scalar* part of the expression following it.

In this formulation, the idea is to find the best rotating triad $\mathbf{E}(t)$, i.e., the one that *maximizes* the fidelity. It may seem that we have not gained much from this reformulation. However, working with $\mathbf{E}(t)$ is in fact a much simpler task than working with the control field $\boldsymbol{\Omega}(\cdot)$ directly, as $\boldsymbol{\Omega}(\cdot)$ is buried inside time-ordered exponentials [see Eqs. (35) and (36)]. As a result, in practice one usually resorts to approximate schemes. In contrast, the set $\mathbf{E}(t)$ appears at the same level as the stochastic fields, and thus can be treated exactly. Furthermore, once we have found the best triad history $\mathbf{E}(t)$, it is straightforward to recover the control fields $\boldsymbol{\Omega}(t)$:

$$\boldsymbol{\Omega}(t) = \tfrac{1}{2}(\partial_t\mathbf{E}_k(t))\mathbf{E}_k(t)$$
$$= \tfrac{1}{2}\epsilon^{ijk}\mathbf{E}_i(t)(\mathbf{E}_j(t)\cdot\partial_t\mathbf{E}_k(t)). \tag{44}$$

These are the control fields in the *rotating frame*. Ultimately, the objects we are interested in are the control field in the *laboratory frame*, $\boldsymbol{\omega}(t) = \omega^i(t)\,\mathbf{e}_i$, which is trivially found via

$$\boldsymbol{\omega}(t) = u_c(t)\,\boldsymbol{\Omega}(t)\,\bar{u}_c(t)$$
$$= \tfrac{1}{2}\epsilon^{ijk}u_c(t)\,\mathbf{E}_i(t)\,\bar{u}_c(t)(\mathbf{E}_j(t)\cdot\partial_t\mathbf{E}_k(t))$$
$$= \tfrac{1}{2}\epsilon^{ijk}\,\mathbf{e}_i(\mathbf{E}_j(t)\cdot\partial_t\mathbf{E}_k(t)). \tag{45}$$

In other words, $\omega^i(t) = \boldsymbol{\omega}(t)\cdot\mathbf{e}_i = \boldsymbol{\Omega}(t)\cdot\mathbf{E}_i(t)$. The laboratory frame components of the control field can be recovered directly from the rotating frame triad $\mathbf{E}(t)$, without the need to

take the intermediate step of computing either $\boldsymbol{\Omega}(t)$ or $u_c(t)$:

$$\omega^i(t) = \tfrac{1}{2}\epsilon^{ijk}\mathbf{E}_j(t) \cdot \partial_t \mathbf{E}_k(t). \tag{46}$$

Here and elsewhere, the dot denotes the quaternion dot product (see Appendix A), and repeated indices are summed over. Note that Eq. (46) is an *exact* relation between $\boldsymbol{\omega}(t)$ and $\mathbf{E}(t)$.

By taking the triad $\mathbf{E}(t)$ as the relevant degrees of freedom [as opposed to the control $\boldsymbol{\Omega}(t)$], we no longer have to worry about the time-ordered exponential associated with $\boldsymbol{\Omega}(t)$, but there still remains an overall time-ordering operator in front of the whole expression; see Eq. (35). For a pure quaternion $\mathbf{n}(t)$, one can always find *another* pure quaternion $\mathbf{m}_\epsilon(t)$ such that the relation,

$$\mathcal{T}e^{\frac{\epsilon}{2}\int_0^\tau dt\, \mathbf{n}(t)} = e^{\frac{\epsilon}{2}\mathbf{m}_\epsilon(\tau)}, \tag{47}$$

is satisfied, and we note that $\mathbf{m}_\epsilon(t)$ is a function of ϵ as well as time. The expression for $\mathbf{m}_\epsilon(t)$ can be found, as is commonly done in quantum mechanics from the Magnus expansion; i.e., as an expansion of $\log \mathcal{T}\exp(\epsilon X)$ (for some quantum operator X, i.e., see Ref. [21]) for a small parameter ϵ. By using the quaternion formulation, however, it is straightforward to determine how the two quantities $\mathbf{n}(t)$ and $\mathbf{m}_\epsilon(t)$ are related to each other *exactly*; this relationship comes in the form of a differential equation, viz.,

$$\frac{d\mathbf{m}_\epsilon(t)}{dt} = \mathbf{n}(t) - \frac{\epsilon}{2}\mathbf{m}_\epsilon(t) \wedge \mathbf{n}(t) + \left(1 - \frac{\epsilon m_\epsilon(t)}{2}\cot\frac{\epsilon m_\epsilon(t)}{2}\right)$$
$$\times \hat{\mathbf{m}}_\epsilon(t) \wedge (\hat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)), \tag{48}$$

where $m_\epsilon(t) \equiv |\mathbf{m}_\epsilon(t)|$ is the quaternion modulus, and $\hat{\mathbf{m}}_\epsilon(t) = \mathbf{m}_\epsilon(t)/|\mathbf{m}_\epsilon(t)|$ is a unit pure quaternion and, as such, it can be interpreted as the *direction* of $\mathbf{m}_\epsilon(t)$ (see Appendix B for a derivation of this result). Note that in the case of pure quaternions the quaternion wedge product acts just like the vector cross product. If Eq. (48) is solved perturbatively in ϵ, one recovers the Magnus expansion (see Appendix B for details), but the real power of Eq. (48) is that it is an exact relation: The solution of this differential equation is equivalent to the exact summation over *all* terms in the Magnus expansion.

After this lengthy detour, let us return to the task of analyzing the fidelity. Making use of the results in Eqs. (43) and (47), we find

$$\mathcal{A}_{1/2}[\mathbf{E}(\cdot),\mathbf{m}_\epsilon(\cdot)] = \cos\left(\frac{\epsilon m_\epsilon(\tau)}{2}\right), \tag{49}$$

which, in view of its simplicity relative to the expression in terms of $\{n_i(\cdot)\}$, we can use to find the exact expression for the fidelity amplitude for general spin s:

$$\mathcal{A}_s(\mathbf{E}(t),\mathbf{m}_\epsilon(t)) = \frac{1}{2s+1}\sum_{j=-s}^{s} e^{-ij\epsilon m_\epsilon(\tau)}. \tag{50}$$

This enables us to construct a simple expression for the fidelity $\mathcal{F}_s = \langle \mathcal{A}_s \rangle$, using Eqs. (50) and (21):

$$\mathcal{F}_s[\mathbf{E}(\cdot)] = \frac{1}{2s+1}\sum_{j=-s}^{s} e^{-(j\epsilon)^2 S[\mathbf{E}(\cdot)] + O(\epsilon^3)}, \tag{51}$$

where $S[\mathbf{E}(\cdot)]$ is a functional of the rotating triad, i.e.,

$$S[\mathbf{E}(\cdot)] = \frac{1}{2}\int_0^\tau \int_0^\tau dt\, dt'\, \mathcal{N}_{ij}(t,t')\,\mathbf{E}_i(t) \cdot \mathbf{E}_i(t'). \tag{52}$$

In the expression (51) for $\mathcal{F}_s$, we make the assumption that ϵ is small, and only consider the leading-order term in ϵ in the exponent. Higher order contributions can be easily computed; see Appendix B for details on how this is accomplished.

It is clear from Eq. (51) that in order to maximize the fidelity, we need only minimize the functional $S[\mathbf{E}(\cdot)]$—note that this condition is *independent* of the spin number s. In other words, the functional form of the rotating triad $\mathbf{E}_i(t)$ [and therefore the control field; see Eq. (46)] that maximizes the fidelity is the *same* for all values of s. The fidelity itself, however, *does* depend on s.

C. Extremals and constraints

The task that remains is to find the rotating triad $\mathbf{E}_i(t)$ that minimizes the functional $S[\mathbf{E}(\cdot)]$; cf. Eq. (52). We look for solutions in the form of extremals of the action, i.e., we seek $\mathbf{E}_i(t)$ such that the variation $\delta S[\mathbf{E}(\cdot)]$ vanishes. Note that the variations in the triad $\delta \mathbf{E}_i(t)$ are not fully arbitrary. They are *constrained* due to the fact that the triad has to rotate as a rigid body. Simply put, the variations are constrained to take the form,

$$\delta \mathbf{E}_i(t) = \delta \mathbf{A}(t) \wedge \mathbf{E}_i(t), \tag{53}$$

where $\delta \mathbf{A}(t)$ is now *any* arbitrary pure quaternion.

In addition to the rigidity constraints just stated, there are also *experimental* constraints present that one should account for, such as limitations on the frequency, amplitude, etc., that the control fields can take. Even in the absence of experimental considerations, it is natural to impose such constraints, if we are to make fair comparisons between different realizations of the control fields. For instance, if there is no limit to the amplitude of the control field, one may simply pick a large enough amplitude such that one can achieve the target operation over a time scale much smaller than any time scale associated with the noise $\mathcal{N}_{ij}(t,t')$. The greater the separation of time scales, the smaller will be the effective action S, and thus the higher will be the fidelity $\mathcal{F}$. In this case, there is no sense in comparing strong fields with weak ones, as strong fields always win.

In order to make a sensible comparison between different choices for the control fields, we have to put some constraint on the solution space in which we seek trajectories for the triad $\mathbf{E}(t)$. As an example, from a purely physical standpoint, one may choose to compare trajectories for which the *total energy output* associated with the control field is *prescribed*. This quantity is given by the time integral of the *square* of the control field, so one has the following constraint [22]:

$$E_{\text{out}} \equiv \int_0^\tau dt\, \frac{|\boldsymbol{\Omega}(t)|^2}{2} = \text{const.} \tag{54}$$

In what follows, we shall also prescribe the transit time τ. As we are fixing the energy output, if we were to make τ too small, there would not be enough time to achieve a given target. Effective values for τ should be bounded from below in an E_{out} dependent way to ensure that we have access to all

desired targets. We shall seek optimal controls in the space of fixed τ and E_{out}.

To determine the optimal controls in this constrained space we seek minima corresponding to the constrained functional S_c, determined via

$$S_c \equiv \frac{1}{2} \int_0^\tau dt \int_0^\tau dt' \{ \mathcal{N}_{ij}(t,t') \mathbf{E}_i(t) \cdot \mathbf{E}_j(t')$$
$$+ \lambda \, \delta(t - t') |\mathbf{\Omega}(t)|^2 \}, \tag{55}$$

where λ is the Lagrange multiplier associated with the energy output constraint, and the output power $|\mathbf{\Omega}(t)|^2$ is determined entirely in terms of the triad $\mathbf{E}(t)$, via

$$|\mathbf{\Omega}(t)|^2 = \frac{1}{8} \delta^{ik} \delta^{j\ell} (\mathbf{E}_i(t) \cdot \partial_t \mathbf{E}_j(t) - \mathbf{E}_j(t) \cdot \partial_t \mathbf{E}_i(t))$$
$$\times (\mathbf{E}_k(t) \cdot \partial_t \mathbf{E}_\ell(t) - \mathbf{E}_\ell(t) \cdot \partial_t \mathbf{E}_k(t))$$
$$= \frac{1}{2} \epsilon^{ijk} \mathbf{E}_i(t) \cdot (\partial_t \mathbf{E}_j(t) \wedge \partial_t \mathbf{E}_k(t)). \tag{56}$$

Thus we see that S_c is a functional of the triad $\mathbf{E}(t)$ and its first time derivatives only. To obtain the first line of Eq. (56) we use the expression for $\mathbf{\Omega}(t)$ in Eq. (44); the second line requires a little algebra. In practice it is simpler to work with the expression as given in the second line, because it is of lower order in the triad and its derivatives. Note that $\lambda |\mathbf{\Omega}(t)|^2/2$ has a natural interpretation as the *kinetic energy* associated with the triad $\mathbf{E}(t)$, with the Lagrange multiplier λ playing the role of inertia.

The constraint present in Eq. (55) is just one type of a constraint that we may impose on the system. We are free to impose other types. For instance, we can replace the Lagrange multiplier in Eq. (55) by a matrix, and even make that time dependent:

$$\frac{1}{2} \int_0^\tau dt \, \boldsymbol{\omega}(t) \cdot \lambda(t) \cdot \boldsymbol{\omega}(t).$$

This may be useful, e.g., in the situation where the control fields $\boldsymbol{\omega}(t)$ are constrained to lie in the xy plane. In this case, we simply let λ_{zz} tend to infinity. We are also free to impose constraints on the derivatives of $\boldsymbol{\omega}(t)$, i.e., to impose a frequency cutoff. Indeed, we have a lot of freedom on the types of constraints we may impose. An advantage of this method (based on extremals) is the ease with which one can implement them.

For illustration purposes, we work with with the case of scalar, time-independent λ in Eq. (55), imposing constraints on the total energy output. From the action S_c, we can obtain the Euler-Lagrange equation, which will give us the condition that the extremals must satisfy. For our case, we are specifically interested in the *minima*. Setting $\delta S_c / \delta \mathbf{A}(t) = 0$ [see Eq. (53)] we get the Euler-Lagrange equation for the triad,

$$\lambda \partial_t \mathbf{\Omega}(t) + \mathbf{E}_i(t) \wedge \left[\int_0^\tau dt' \mathcal{N}_{ij}(t,t') \mathbf{E}_j(t') \right] = 0, \tag{57}$$

where, in writing down Eq. (57), we used the fact that

$$\partial_t \mathbf{\Omega}(t) = \frac{1}{2} \epsilon^{ijk} \mathbf{E}_i(t) \left[\mathbf{E}_j(t) \cdot \partial_t^2 \mathbf{E}_k(t) \right]. \tag{58}$$

The form of Eq. (57) fits naturally with the fact that $\mathbf{\Omega}(t)$ is the angular velocity of the triad, and indicates that the interpretation of $\lambda |\mathbf{\Omega}(t)|^2/2$ as the kinetic energy, with λ playing the role of inertia, is correct.

Given the connection between $\mathbf{\Omega}(t)$ and the set $\mathbf{E}_i(t)$ [see Eq. (44)], we also have

$$\partial_t \mathbf{E}_i(t) = \mathbf{E}_i(t) \wedge \mathbf{\Omega}(t). \tag{59}$$

It is useful to define the *dual triad* $\mathbf{D}_i(t)$, via

$$\mathbf{D}_i(t) = \int_0^\tau dt' \mathcal{N}_{ij}(t,t') \mathbf{E}_j(t'). \tag{60}$$

The solution of the set of coupled equations,

$$\partial_t \mathbf{\Omega}(t) = -\frac{1}{\lambda} \mathbf{E}_i(t) \wedge \mathbf{D}_i(t),$$
$$\partial_t \mathbf{E}_i(t) = \mathbf{E}_i(t) \wedge \mathbf{\Omega}(t), \tag{61}$$

along with the appropriate set of boundary conditions, determines the optimal controls. Recall that we have a target operation U_T, which is to be satisfied at the transit time $t = \tau$. For our case, U_T is simply a rotation operator, and can therefore always be written as

$$U_T = e^{-i\frac{1}{2}\theta_T \hat{\mathbf{r}} \cdot \boldsymbol{\sigma}}, \tag{62}$$

where θ_T is the angle and $\hat{\mathbf{r}}$ is the axis of rotation. It is easy to translate this into quaternion language and find the corresponding unit quaternion u_T that gives the target rotation,

$$u_T = e^{\frac{1}{2}\mathbf{q}_T}, \tag{63}$$

where $\mathbf{q}_T = q^i \mathbf{e}_i$ is the pure quaternion corresponding to $-i\theta_T \hat{\mathbf{r}} \cdot \boldsymbol{\sigma}$, as determined by Eq. (40). The quaternion modulus $|\mathbf{q}_T|$ corresponds to the angle of rotation, and $\hat{\mathbf{q}}_T = \mathbf{q}_T/|\mathbf{q}_T|$ corresponds to the axis of rotation. The boundary conditions for the equation of motion Eq. (61) are then entirely determined from the target u_T, via

$$\mathbf{E}_i(0) = \mathbf{e}_i,$$
$$\mathbf{E}_i(\tau) = \bar{u}_T \mathbf{e}_i u_T. \tag{64}$$

The presence of boundary conditions specific to the target rotation is an inconvenience when investigating generic properties of extremals. It is easy to take care of *arbitrary* boundary conditions once and for all by transforming to a suitably *rotating frame*, which is described by some angular drift velocity $\mathbf{\Omega}_0(t)$, generally an arbitrary function of time. The rotating-frame triad $\widetilde{\mathbf{E}}_i(t)$ is then given by

$$\widetilde{\mathbf{E}}_i(t) = u_0(t) \mathbf{E}_i(t) \bar{u}_0(t), \tag{65}$$

where $u_0(t) \equiv \mathcal{T} \exp\left[\frac{1}{2} \int_0^t dt' \, \mathbf{\Omega}_0(t')\right]$. Likewise, for any rotating-frame quantity $\widetilde{\mathbf{X}}$ [e.g., the control fields $\widetilde{\mathbf{\Omega}}(t)$], we have $\widetilde{\mathbf{X}} = u_0(t) \mathbf{X} \bar{u}_0(t)$. We can also define the covariant derivative for the rotating frame $\widetilde{\partial}_t$, through the relation,

$$\widetilde{\partial}_t \mathbf{X}(t) = \partial_t \mathbf{X}(t) - \mathbf{\Omega}_0(t) \wedge \mathbf{X}(t); \tag{66}$$

likewise, for the noise kernel we have the covariant integral operator,

$$\int_0^\tau dt' \widetilde{\mathcal{N}}_{ij}(t,t') \mathbf{X}_j(t')$$
$$= \int_0^\tau dt' \mathcal{N}_{ij}(t,t') u_0(t,t') \mathbf{X}_j(t') \bar{u}_0(t,t'), \tag{67}$$

where $u_0(t,t') \equiv u_0(t) \bar{u}_0(t')$. In terms of the quantities carrying tildes, the Euler-Lagrange equations for S_c can be written in

a *generally covariant* form, valid for *arbitrary* time-dependent rotating frame $\boldsymbol{\Omega}_0(t)$:

$$\widetilde{\partial}_t \widetilde{\mathbf{E}}_i(t) = \widetilde{\mathbf{E}}_i(t) \wedge \widetilde{\boldsymbol{\Omega}}(t),$$

$$\widetilde{\partial}_t \widetilde{\boldsymbol{\Omega}}(t) = -\frac{1}{\lambda} \widetilde{\mathbf{E}}_i(t) \wedge \widetilde{\mathbf{D}}_i(t). \tag{68}$$

A consequence of the fact that the form of S_c expressed in terms of $\widetilde{\mathbf{E}}_i(t)$ is *invariant* under arbitrary $\boldsymbol{\Omega}_0(t)$, is that this general covariance carries over to all physical equations, i.e., the relation between $\widetilde{\boldsymbol{\Omega}}(t)$ and $\widetilde{\mathbf{E}}_i(t)$ takes the expected form,

$$\widetilde{\boldsymbol{\Omega}}(t) = \tfrac{1}{2} \epsilon^{ijk} \widetilde{\mathbf{E}}_i(t) [\widetilde{\mathbf{E}}_j(t) \cdot \widetilde{\partial}_t \widetilde{\mathbf{E}}_k(t)]. \tag{69}$$

We can take advantage of the freedom we have in choosing $\boldsymbol{\Omega}_0(t)$ to force the boundary conditions to take on a simple form,

$$\widetilde{\mathbf{E}}_i(0) = \widetilde{\mathbf{E}}_i(\tau) = \mathbf{e}_i, \tag{70}$$

the only requirement being that in order to satisfy the boundary condition in the laboratory frame, we simply need $(u_0(0), u_0(\tau)) = (1, u_T)$, with u_T being the target operation. The solutions $\widetilde{\mathbf{E}}_i(t)$ then satisfy periodic boundary conditions, i.e., each member of the triad forms a *closed loop* on the unit sphere. The only other constraint is that the triad rotates as a rigid body. The loops are not constrained in any other way, i.e., they are free to cross each other an arbitrary number of times.

Without loss of generality, we choose $\boldsymbol{\Omega}_0(t)$ to be a constant $\boldsymbol{\Omega}_D$, so that all nontrivial dynamical behavior is displayed by the triad $\widetilde{\mathbf{E}}_i(t)$. This particular rotating frame, with constant $\boldsymbol{\Omega}_D$, corresponds to eliminating the *drift* term whose effect is to take the triad $\mathbf{E}_i(t)$ through a *free geodesic path* on the unit sphere connecting the starting point at $t = 0$ with the target at $t = \tau$. This procedure simply accounts for this trivial part of the evolution, allowing us to focus on corrections due to fluctuations in the environment. For this reason, it is advantageous to solve the Euler-Lagrange equations in this rotating frame. We take a further step, and work with the *difference* $\delta\widetilde{\boldsymbol{\Omega}}(t) \equiv \widetilde{\boldsymbol{\Omega}}(t) - \boldsymbol{\Omega}_D$, and thus the set of equations,

$$\partial_t \delta\widetilde{\boldsymbol{\Omega}}(t) = \boldsymbol{\Omega}_D \wedge \delta\widetilde{\boldsymbol{\Omega}}(t) - \frac{1}{\lambda} \widetilde{\mathbf{E}}_i(t) \wedge \widetilde{\mathbf{D}}_i(t), \tag{71a}$$

$$\partial_t \widetilde{\mathbf{E}}_i(t) = \widetilde{\mathbf{E}}_i(t) \wedge \delta\widetilde{\boldsymbol{\Omega}}(t), \tag{71b}$$

satisfying boundary conditions given in Eq. (70). This set of variables holds the advantage that the trajectories $\delta\widetilde{\boldsymbol{\Omega}}(t), \widetilde{\mathbf{E}}_i(t)$, which describe deviations from the average drift, are *small* in the limit where the noise matrix $\mathcal{N}_{ij}(t,t')$ is small, and vanish in the limit of vanishing noise, making them a natural set of variables to work with.

Before proceeding with examples, we note that the trajectories as determined from Eqs. (71a) and (71b) can belong to *distinct topological sectors*. This is easy to see in the case where there is no noise, i.e., when $\boldsymbol{\Omega}(t) = \boldsymbol{\Omega}_D$, since the set of drift vectors $\boldsymbol{\Omega}_D^{(n)} = (f + 2\pi n)\,\hat{\mathbf{r}}$ (for constant f and integer n) give rise to trajectories for $\mathbf{E}_i(t)$ that satisfy the same boundary conditions (i.e., correspond to the same target), but differ in the number of times the triad $\mathbf{E}_i(t)$ winds around the axis determined by $\hat{\mathbf{r}}$. These *winding numbers n* then describe different topological sectors. The same situation arises in the case of nonzero noise; here one also obtains sets of trajectories, differing in winding number, but satisfying the same boundary

conditions. When seeking solutions to the equations of motion, one can thus also specify the topological sector of interest, which can without loss of generality be accounted for by the drift term $\boldsymbol{\Omega}_D^{(n)}$ (where there is no risk of confusion we omit the label n for brevity).

Finally, we mention that the *laboratory frame* components of the control field trajectories, the quantities we are ultimately interested in, can be determined entirely in terms of the rotating frame quantities $\omega_i(t) = \boldsymbol{\omega}(t) \cdot \mathbf{e}_i = (\boldsymbol{\Omega}_D + \delta\widetilde{\boldsymbol{\Omega}}(t)) \cdot \widetilde{\mathbf{E}}_i(t)$.

D. Full probability distribution function for the (in)fidelity

As we stated in Sec. II, for some applications one may require the use of the full probability distribution for the infidelity. As an example, we briefly discussed the possibility of looking into the effect of rare extreme events, which though unlikely to occur, can be unacceptably detrimental to the control task. Such questions are commonly addressed in *extreme value statistics*, and require more information than simply the average; usually a full account of the probability distribution is needed (see, e.g., Ref. [16]).

The formalism we shall be using here has been developed in Sec. II D, and we refer the reader there for details (see also Ref. [15]). In the present case, where we consider weak noise ($\epsilon \ll 1$), it is more convenient to work with the *infidelity amplitude*,

$$\mathcal{I}_s \equiv 1 - \mathcal{A}_s. \tag{72}$$

To illustrate the method, we expand this quantity to order ϵ^2, and all results obtained are good to this order. We have

$$\mathcal{I}_s \approx \epsilon_s^2 \int_0^\tau \int_0^\tau dt\, dt'\, n_i(t) n_j(t') \mathbf{E}_i(t) \cdot \mathbf{E}_j(t') \tag{73}$$

[see Eq. (50) for the relevant expression for $\mathcal{A}_s$ and Appendix B for details on the relationship between $\mathbf{m}_\epsilon(t)$ and $\mathbf{n}(t)$, especially Eq. (B10)], where we have defined the renormalized ϵ_s via

$$\epsilon_s = \sqrt{\frac{s(s+1)}{3!}}\epsilon. \tag{74}$$

Next, we give the result for the characteristic function:

$$\chi_s(k) = \langle e^{-ik\mathcal{I}_s} \rangle_n$$
$$= \frac{1}{\sqrt{\text{Det}(\mathbb{1} + ik\epsilon_s^2 \mathcal{M}[\mathbf{E}(\cdot)])}}, \tag{75}$$

where in obtaining $\chi_s(k)$ we have used the approximation for $\mathcal{I}_s$ given in Eq. (73), along with the assumption that the noise fields $n_i(t)$ are Gaussian correlated. In the second line of Eq. (75), $\mathbb{1}$ corresponds to a 3×3 unit matrix and $\mathcal{M}[\mathbf{E}(\cdot)]$ is a *matrix functional* of $\mathbf{E}_i(t)$ given by

$$\mathcal{M}[\mathbf{E}(\cdot)] = \frac{1}{2} \int_0^\tau \int_0^\tau dt\, dt'\, \mathcal{E}^T(t) \mathcal{N}(t,t') \mathcal{E}(t'), \tag{76}$$

in which $\mathcal{N}(t,t')$ is the noise correlation matrix with elements $\mathcal{N}_{ij}(t,t')$, and $\mathcal{E}(t)$ is a matrix, with transpose $\mathcal{E}^T(t)$, whose elements contain the *components* of the *rotating-frame* triad $\mathbf{E}_i(t)$ in the basis of the *laboratory frame* triad $\mathbf{e}_i$, viz.,

$$\mathcal{E}_{ij}(t) = \mathbf{E}_i(t) \cdot \mathbf{e}_j. \tag{77}$$

In the case of weak noise (i.e., small ϵ_s) considered here, it is sufficient to expand the determinant in Eq. (75) to leading order in ϵ_s, so that

$$\mathrm{Det}\big(\mathbb{1} + ik\epsilon_s^2 \mathcal{M}[\mathbf{E}(\cdot)]\big) \approx 1 + ik\epsilon_s^2 \mathrm{Tr}(\mathcal{M}[\mathbf{E}(\cdot)])$$
$$= 1 + ik\epsilon_s^2 S[\mathbf{E}(\cdot)], \qquad (78)$$

where $S[\mathbf{E}(\cdot)]$ is the effective "action" functional that we found in Sec. IV B [see Eq. (52)], which we rewrite here for convenience:

$$S[\mathbf{E}(\cdot)] = \frac{1}{2} \int_0^\tau \int_0^\tau dt\, dt'\, \mathcal{N}_{ij}(t,t')\, \mathbf{E}_i(t) \cdot \mathbf{E}_i(t'). \quad (79)$$

Recall that the triad $\mathbf{E}_i(t)$ that minimizes $S[\mathbf{E}(\cdot)]$ determines the optimal control field histories that maximize (minimize) the fidelity (infidelity). By using Eq. (79), we obtain the following expression for $\chi_s(k)$, which to this level of approximation depends on the triad $\mathbf{E}_i(t)$ (and therefore on the control field histories) *solely* through the functional $S[\mathbf{E}(\cdot)]$:

$$\chi_s(k) = \frac{1}{\sqrt{1 + ik\epsilon_s^2 S[\mathbf{E}(\cdot)]}}. \qquad (80)$$

By using Eq. (80), it is straightforward to calculate *all* the moments and cumulants in the limit of weak noise. For the moments, we find

$$\langle \mathcal{I}^m \rangle_n = (i\,\partial_k)^m \chi_s(k)|_{k=0}$$
$$= (2m-1)!! \left(\frac{\epsilon_s^2 S[\mathbf{E}(\cdot)]}{2} \right)^m, \qquad (81)$$

and for the cumulants we find

$$\langle\langle \mathcal{I}^m \rangle\rangle_n = (i\,\partial_k)^m \ln \chi_s(k)|_{k=0}$$
$$= (2m-2)!! \left(\frac{\epsilon_s^2 S[\mathbf{E}(\cdot)]}{2} \right)^m. \qquad (82)$$

Furthermore, it is straightforward to find $P(\mathcal{I}_s)$ in the weak-noise case under consideration:

$$P(\mathcal{I}_s) = \int_{-\infty}^\infty \frac{dk}{2\pi}\, \chi_s(k) e^{ik\mathcal{I}_s}$$
$$\approx \frac{\Theta(\mathcal{I}_s)}{\sqrt{\pi \mathcal{I}_s \epsilon_s^2 S[\mathbf{E}(\cdot)]}} \exp\left[-\frac{\mathcal{I}_s}{\epsilon_s^2 S[\mathbf{E}(\cdot)]} \right], \qquad (83)$$

where $\Theta(\mathcal{I}_s)$ is the Heaviside step function (which vanishes for $\mathcal{I}_s < 0$ and gives unity otherwise), and in evaluating Eq. (83) we have used the approximate expression for $\chi_s(k)$ given in Eq. (80). Note that although the expression obtained in Eq. (83) is valid only in the case of weak noise (i.e., $\epsilon_s \ll 1$), it has an essential singularity at $\epsilon_s = 0$, so the expression obtained here is nonperturbative in nature. For a given fixed ϵ_s, we see from Eq. (83) that in order to maximize the fidelity—along with all moments of the fidelity amplitude—one must employ control field histories that make $S[\mathbf{E}(\cdot)]$ as small as possible. We note, however, that Eq. (83) is not valid for vanishing S; in this case, higher-order contributions in ϵ_s from Eqs. (72) and (78) come into play, and there are quantities other than S that must vanish in order to recover perfect fidelity. Nevertheless, Eq. (83) is a useful result, that remains valid even for small infidelities, and this includes the examples considered in this paper.

E. Examples

A physically interesting situation arises when the noise has a $1/f$ character. In many systems, this accounts for the material-dependent noise source that leads to decoherence in quantum devices [18]. It is therefore of considerate importance for applications to mitigate this noise effectively. $1/f$ noise can be understood as arising from the collective effect of multiple sources of *telegraph noise* that are coupled to the quantum system of interest, with the weight of each source scaling as the inverse characteristic decay rate of that source [18,23].

In what follows, we focus on the case of mitigating noise along a fixed axis, bearing in mind that generalizations are straightforward. Without loss of generality, we take the noise to point along the x axis, i.e., the only nonvanishing component of the noise matrix is $\mathcal{N}_{xx}(t,t')$. For $1/f$ noise, this is obtained via the expression,

$$\mathcal{N}_{xx}(t,t') = \xi \int_{\gamma_1}^{\gamma_2} \frac{d\gamma}{\gamma} e^{-\gamma|t-t'|}, \qquad (84)$$

where ξ denotes the strength of the noise, and γ_1 (γ_2) is the lower (upper) decay-rate cutoff for the ensemble of telegraph processes that are coupled to the spin. For frequencies ν within the range $\gamma_1 \ll \nu \ll \gamma_2$, the noise correlator in the frequency domain $\mathcal{N}_{xx}(\nu) \sim \nu^{-1}$, i.e., is $1/f$ in character.

For the noise form under consideration, we seek solutions for $\mathbf{\Omega}(t), \mathbf{E}_i(t)$ that depend parametrically on the Lagrange multiplier λ. The case $\lambda^{-1} = 0$ corresponds to making energy output infinitely costly, and gives rise to solutions corresponding to the geodesic path on a sphere, i.e., we have $\mathbf{\Omega}(t) = \mathbf{\Omega}_D$. For increasing λ^{-1}, the energy cost decreases, so that there is nontrivial competition between keeping energy costs low and compensating for the noise. In the limiting case of $\lambda^{-1} \to \infty$, the energy costs become vanishingly small meaning that there are no constraints on the set of extremal control fields we get to choose from. In this case, we can simply take the control $\mathbf{\Omega}(t)$ to be an infinitely sharp pulse acting over a vanishingly small transit time τ, and this guarantees maximal fidelity. In actuality, as long as we ensure that the spectral content of the noise and the control do not overlap, which we are always free to do in the case where there are no constraints on the control, we guarantee maximal fidelity at leading order in ϵ; see Refs. [6,7].

As we have mentioned, one advantage of this approach is that we can naturally find families of solutions corresponding to a given $\mathbf{\Omega}_D$. Raising $\lambda^{-1} = 0$, we find a set of solutions that correspond to continuous deformations of the geodesic on a sphere. As we raise λ^{-1} and loosen the constraints on the energy output, we find solutions which improve the fidelity more and more.

In order to illustrate our approach, we take the drift $\mathbf{\Omega}_D$ to be $(2\pi, 0, 2\pi)/\tau$ so that we have a nontrivial winding-number (note that $\mathbf{\Omega}_D = (\sqrt{2} - 1)(2\pi, 0, 2\pi)/\tau$ would give us the same target rotation), and the target operation has both a nonvanishing component *parallel* to the noise and a nonvanishing component *orthogonal* to it. In this case, we expect to obtain nontrivial time dependence in both the amplitude of the control field, $|\mathbf{\Omega}(t)|$, and the direction $\widehat{\mathbf{\Omega}}(t)$. It is convenient to scale all parameters with respect to the transit time τ, i.e., we take $\xi = 8/\tau^2$ (the strength of the noise relative
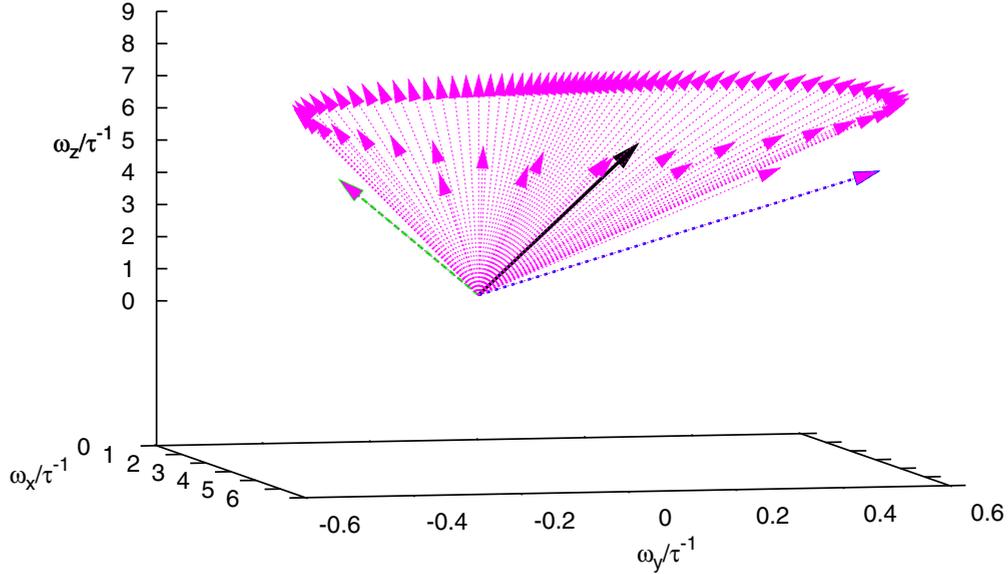
FIG. 1. Plot of $\boldsymbol{\omega}(t)/\tau^{-1}$ (dimensionless, with τ the total transit time) for $\lambda^{-1} = 50\,\tau^{-1}$ (light purple arrows), for the case of $1/f$ noise. The drift field $\boldsymbol{\Omega}_D = (2\pi, 0, 2\pi)/\tau$ (bold black arrow), which also corresponds to the optimal control for the case $\lambda^{-1} = 0$, is shown for comparison. The green dashed arrow (on the left side of the plot) corresponds to the initial value of $\boldsymbol{\omega}(t)$, while the blue dashed arrow (on the right side) corresponds to its final value. The other parameter values are $\gamma_1 = 20/\tau$, $\gamma_2 = 0.1/\tau$, and $\xi = 8/\tau^2$, as described in the text.

to the control is then $\sim \epsilon \xi$), $\gamma_1 = 0.1/\tau$, and $\gamma_2 = 20/\tau$. The values chosen for the cutoffs γ_1 and γ_2 give us a wide range of frequencies for which the noise $\mathcal{N}_{xx}$ is well approximated by $1/f$ noise.

Our interest lies in the case where λ^{-1} is not too large, so that there is nontrivial competition between minimizing energy output and maximizing the fidelity. The equations of motion (68) can be solved straightforwardly, numerically. The quantities we are ultimately interested in are the *control fields in the laboratory frame* $\boldsymbol{\omega}(t)$, whose components are given by

$$\omega_i(t) = \boldsymbol{\Omega}(t) \cdot \mathbf{E}_i(t)$$
$$= (\boldsymbol{\Omega}_D + \delta\widetilde{\boldsymbol{\Omega}}(t)) \cdot \widetilde{\mathbf{E}}_i(t), \tag{85}$$

where $\boldsymbol{\Omega}_D = (2\pi, 0, 2\pi)/\tau$ is the drift and $\delta\widetilde{\boldsymbol{\Omega}}(t) \equiv \widetilde{\boldsymbol{\Omega}}(t) - \boldsymbol{\Omega}_D$, with $\widetilde{\boldsymbol{\Omega}}(t)$ and $\widetilde{\mathbf{E}}_i(t)$ the control fields and the triad as given in the rotating frame; which we find by solving the rotating-frame Euler-Lagrange equations in the presence of the drift term $\boldsymbol{\Omega}_D$ [see Eqs. (71a) and (71b) and the accompanying

discussion]. The optimal control field history $\boldsymbol{\omega}(t)$ in this nontrivial regime (with parameters given in the caption) is shown in Fig. 1 (set of light purple arrows), where we have set $\lambda^{-1} = 50/\tau$. The drift field $\boldsymbol{\Omega}_D$ $((2\pi, 0, 2\pi)/\tau)$ is shown (bold black arrow) for comparison.

To get a better understanding of the results obtained in this example, let us take a closer look at how each component of $\boldsymbol{\omega}(t)$ behaves. In Fig. 2, we plot the results for $\delta\boldsymbol{\omega}(t) \equiv \boldsymbol{\omega}(t) - \boldsymbol{\Omega}_D$. The curves shown there correspond to several values of the Lagrange multiplier: $\lambda^{-1} = (0, 10, 20, 30, 50, 100)/\tau$; in each plot the solid curve corresponds to $\lambda^{-1} = 100/\tau$, and the dashed curves correspond to a sequence of smaller values.

Recalling the fact that in this example the noise lies along the x axis, the interpretation of the plots becomes clear: $\delta\omega_x(t)$ is *negative definite*, meaning that the overall amplitude of the control field is *reduced* relative to the drift $\boldsymbol{\Omega}_D$ for the *entire process*. This is reasonable because a driving field along the x axis does nothing to compensate for noise along the x direction: The driving and noise terms would commute in this
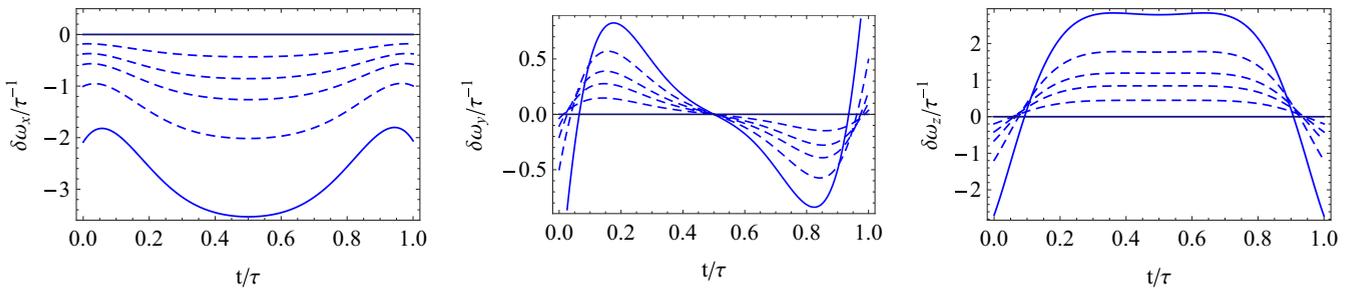

FIG. 2. Plot of the components of $\delta\boldsymbol{\omega}(t)$ (defined in text) vs time, for $\lambda^{-1} = 0, 10, 20, 30, 50, 100\,\tau^{-1}$. All plotted quantities are dimensionless, with τ, the total transit time, playing the role of the fundamental time unit. The solid lines correspond to $\lambda^{-1} = 100\tau^{-1}$ and the dashed lines correspond to intermediate values. As explained in the text, note that the x component of the control is reduced, relative to the noise-free case, at the expense of increasing the z component. This reduces the component of the control field along the noise (taken to be the x axis) in order to compensate more efficiently for the noise; see the text for further discussion.

case. To compensate for the reduced drive along the x axis, the drive along the z axis on average is increased. As far as the drive along the y direction is concerned, although its time average is zero, it takes on a nontrivial time dependence.

The behavior of the fields vary in concert in order to satisfy the boundary conditions while reducing the weight of the control field $\boldsymbol{\omega}(t)$ along the direction of the noise (in this case the x direction) as much as possible. As λ^{-1} is increased, more and more of the weight lies along the y and z axes. As λ^{-1} is increased, and thus the energy restrictions are lessened, we have at our disposal larger-amplitude control fields $\boldsymbol{\omega}(t)$, and therefore larger frequencies in the spectral content of $\mathbf{E}_i(t)$. In light of the fact that $1/f$ noise has higher weight at smaller frequencies, by increasing λ^{-1} we *reduce* the spectral overlap between the noise and the control, and hence *increase* the fidelity, which is what we set out to do. To recapitulate, for the fidelity to be as large as possible, one should put as much of the weight of the control field in a direction orthogonal to the noise for as long as possible. This all has to be done in such a way as to satisfy the boundary conditions (i.e., to obtain the sought after target operation) and the energy constraints. This is why we can only completely eliminate $\omega_x(t)$ in favor of $\omega_{y,z}(t)$ if we have access to arbitrarily large energy outputs.

Recall that although the optimal control scheme $(\boldsymbol{\Omega}(\cdot), \mathbf{E}_i(\cdot))$ is independent of the spin quantum number s, the fidelity itself, however, is not, although the dependence on s is elementary in the limit of weak noise:

$$\mathcal{F}_s[\mathbf{E}(\cdot)] = \frac{1}{2s+1} \sum_{j=-s}^{s} e^{-(j\epsilon)^2 S[\mathbf{E}(\cdot)] + O(\epsilon^3)}. \tag{86}$$

Thus, in the small-ϵ limit considered here, one does best by reducing $S = \frac{1}{2} \int_0^\tau dt \int_0^\tau dt' \mathcal{N}_{ij}(t,t') \mathbf{E}_i(t) \cdot \mathbf{E}_j(t')$ as much as the constraint in energy output will allow, obtaining an optimal control trajectory which is independent of ϵ. For a fixed value of ϵ (which we take to be 0.1), we plot S as a function of λ^{-1} in Fig. 3. Note that, in principle, there is no limit to how small
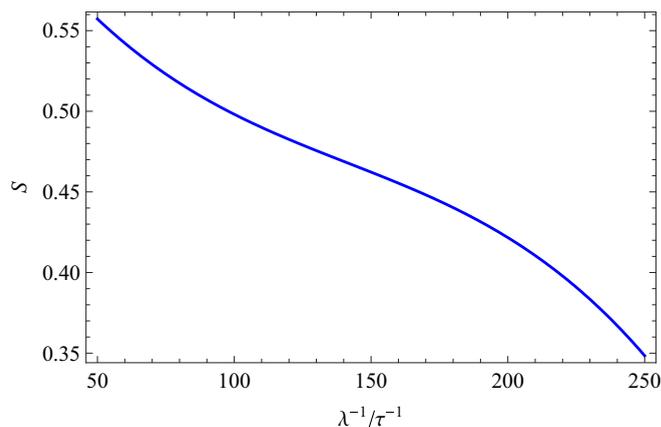

FIG. 3. Plot of the action S as a function of λ^{-1}/τ^{-1} (both quantities in plot are dimensionless), which determines the constraint on energy output for the control fields $\boldsymbol{\Omega}(t)$. Larger values of λ^{-1} correspond to less stringent constraints, i.e., larger energy outputs. As S approaches zero, the fidelity approaches unity [see Eq. (86)]. There is no limit to how small we can make S as long as we are willing to increase λ^{-1}.

S can be as long as we are willing to keep increasing λ^{-1}. For the case of $\lambda^{-1} = 250/\tau$, the maximum value present in Fig. 3, the fidelity for the $s = 1/2$ case is on the order of 0.999. In the weak noise limit (i.e., $\epsilon \ll 1$), the effective action integral, $S = \frac{1}{2} \int_0^\tau dt \int_0^\tau dt' \mathcal{N}_{ij}(t,t') \mathbf{E}_i(t) \cdot \mathbf{E}_j(t')$, not only tells us information regarding the fidelity [see Eq. (86)], but also gives us all the information needed to compute the full probability distribution of infidelity amplitudes, as seen by Eq. (83), which we rewrite here for convenience:

$$P(\mathcal{I}_s) \approx \frac{\Theta(\mathcal{I}_s)}{\sqrt{\pi \mathcal{I}_s \epsilon_s^2 S[\mathbf{E}(\cdot)]}} \exp\left[-\frac{\mathcal{I}_s}{\epsilon_s^2 S[\mathbf{E}(\cdot)]} \right]. \tag{87}$$

This expression tells us that by minimizing the infidelity (which amounts to minimizing S), we automatically minimize all higher moments and cumulants.

All results arrived at for this example can be generalized straightforwardly to the case of a general noise matrix. Greater fidelity usually results from putting as much of the weight of the control fields in a direction orthogonal to the *dominant* noise contributions, while choosing the amplitudes $|\boldsymbol{\omega}(t)|$ such that the triad $\mathbf{E}_i(t)$ has as little spectral overlap with the noise as possible, in other words, as much as is allowed by the constraints imposed upon the control fields.

IV. CONCLUSIONS

We have considered the problem of error mitigation in the control of quantum systems subject to time-dependent sources of noise, which in general includes environmental noise and noise inherent to the experimental control apparatus. We do this in terms of a fidelity metric, which measures how faithfully the time evolution matrix (determined by the controls, internal system dynamics, and the noise) reproduces a predetermined "target" unitary transformation at the end of a prescribed transit time τ. We have tackled the problem through a modification of Schwinger-Keldysh path-integral techniques, which enables us to account for the effects of general sources of noise in a unified manner. By analogy with the Martin-Siggia-Rose scheme we "integrate out" noise sources, thus arriving at an effective deterministic formulation, which we use to represent the fidelity.

Our methods yield the conditions obeyed by the control field history such that the effects of noise are optimally mitigated and hence, on average, the fidelity corresponding to the desired unitary transformation is as large as possible. These conditions are determined by solving equations of motion that are found by extremizing an effective action functional with the control fields playing the role of the degrees of freedom. Our method has the advantage that it admits a wide range of constraints through the use of Lagrange multipliers. These constraints may correspond to those naturally found in experiments, such as optimization of fidelity in the presence of fixed energy input, which is the main type of constraint we consider in the article. We also show how to compute the full probability distribution of the infidelity, which in principle can be used to address more complicated questions, such as those related to minimizing the likelihood of rare, detrimental events.

As an application of our methods, we consider a system composed of a single spin degree of freedom $\hat{\mathbf{S}}$ of arbitrary

spin quantum number s coupled to a noise source that is of the $1/f$ type, which is known to be a common source of quantum decoherence in many systems of interest. We address the limit of weak noise, and study the problem in the case where the total energy output is constrained at a fixed value. We use our methods to find the optimal control field histories subject to applied constraints, and interpret the results. In the case of weak noise, we find that the optimal control field histories are independent of the spin quantum number s, and that the fidelity depends on s in an elementary manner. Finally, we find a simple, analytical expression for the full probability distribution of the infidelity, and show that, in the weak-noise limit, it depends solely on the effective action functional.

Although we have studied the case of a single spin as a specific example, it is straightforward to generalize our methods, e.g., by applying them to chains of coupled spins, atomic systems, as well as general noise distributions. An intriguing avenue that merits exploration, and that we have not pursued in this paper, is the question of many body effects on the control task, where interactions may have nontrivial repercussions. This is a question that can be addressed within the formalism we have constructed here.

APPENDIX A: QUATERNION ALGEBRA

Here we briefly review quaternion algebra and explain the notation used in our paper. A quaternion q can be written in terms of its components as follows (using the summation convention over repeated indices μ):

$$\mathfrak{q} = q^\mu \mathbf{e}_\mu = q^0 \mathbf{e}_0 + q^1 \mathbf{e}_1 + q^2 \mathbf{e}_2 + q^3 \mathbf{e}_3, \qquad \text{(A1)}$$

where the *components* q^μ are real numbers, and the tetrad $\mathbf{e}_\mu$ are the *basis quaternions*. The tetrad $\mathbf{e}_\mu$ satisfy the quaternion multiplication table (i, j, and k are restricted to take nonzero values):

$$\begin{aligned} \mathbf{e}_0 \mathbf{e}_i &= \mathbf{e}_i \mathbf{e}_0 = \mathbf{e}_i, \\ \mathbf{e}_i \mathbf{e}_j &= -\delta_{ij} + \epsilon_{ijk} \mathbf{e}_k. \end{aligned} \qquad \text{(A2)}$$

We can without loss of generality take $\mathbf{e}_0 = 1$; the quaternion algebra can then be understood entirely in terms of the equation on the second line of Eq. (A2), i.e., we only need to consider the *triad* $\mathbf{e}_i$, where $i = 1, 2, 3$. We refer to the zeroth component q^0 as the *scalar part* (in the literature, it is often referred to as the *real part* in analogy with complex numbers), and we refer to $\mathbf{q} \equiv q^i \mathbf{e}_i$ as the *vector part* (it is also, in analogy with complex numbers, the *imaginary part*). In terms of its scalar and vector parts, we can write an arbitrary quaternion q as

$$q = q^0 + \mathbf{q}. \qquad \text{(A3)}$$

Taking the analogy with complex numbers another step, we also have the *quaternion conjugate* $\bar{q}$, where the conjugation operation is defined through its action on the triad $\mathbf{e}_i$. We have

$$\bar{\mathbf{e}}_i = -\mathbf{e}_i, \qquad \text{(A4)}$$

and in terms of an arbitrary quaternion written in terms of its scalar and vector parts, we then have

$$\bar{q} = q^0 - \mathbf{q}. \qquad \text{(A5)}$$

In addition to the *geometric product* between two quaternions $pq = r$, we can also define the quaternion dot product,

$$p \cdot q \equiv \tfrac{1}{2}(p\bar{q} + q\bar{p}), \qquad \text{(A6)}$$

and the quaternion wedge product,

$$p \wedge q \equiv \tfrac{1}{2}(pq - qp). \qquad \text{(A7)}$$

The *magnitude* of a quaternion is given by its *modulus* $|q|$, defined through the relation,

$$|q| = \sqrt{q\bar{q}} = \sqrt{(q^0)^2 + \mathbf{q} \cdot \mathbf{q}}. \qquad \text{(A8)}$$

Pure imaginary quaternions $\mathbf{q}$, referred to as *pure quaternions*, can naturally be interpreted as a geometric vector in $\mathbb{R}^3$. For pure quaternions, the quaternion dot and wedge products correspond to the usual vector dot and cross products, respectively,

$$\begin{aligned} \mathbf{p} \cdot \mathbf{q} &= \tfrac{1}{2} p^i q^j (\mathbf{e}_i \bar{\mathbf{e}}_j + \mathbf{e}_j \bar{\mathbf{e}}_i) = p^i q^j \delta_{ij}, \\ \mathbf{p} \wedge \mathbf{q} &= \tfrac{1}{2} p^i q^j (\mathbf{e}_i \mathbf{e}_j - \mathbf{e}_j \mathbf{e}_i) = p^i q^j {\epsilon_{ij}}^k \mathbf{e}_k, \end{aligned} \qquad \text{(A9)}$$

where we use the quaternion multiplication table in Eq. (A2) (repeated indices are to be summed over). The geometric quaternion product between two pure quaternions $\mathbf{pq}$ subsumes both the dot and wedge products,

$$\mathbf{pq} = -\mathbf{p} \cdot \mathbf{q} + \mathbf{p} \wedge \mathbf{q}. \qquad \text{(A10)}$$

Unit quaternions, defined as quaternions for which $|u| = 1$, can always be written in the form,

$$u = e^{\mathbf{p}}, \qquad \text{(A11)}$$

where $\mathbf{p}$ is a pure quaternion. Unit quaternions are used in the quaternion language to describe rotations. Given a unit quaternion $v = e^{\frac{1}{2}\boldsymbol{\theta}}$ and any pure quaternion $\mathbf{p}$ we can write down the rotated quaternion $\mathbf{p}'$ as

$$\begin{aligned} \mathbf{p}' = v\mathbf{p}\bar{v} &= e^{\frac{1}{2}\boldsymbol{\theta}} \mathbf{p} e^{-\frac{1}{2}\boldsymbol{\theta}} = \mathbf{p} \cos\theta + \widehat{\boldsymbol{\theta}} \wedge \mathbf{p} \sin\theta \\ &\quad + (\mathbf{p} - (\widehat{\boldsymbol{\theta}} \cdot \mathbf{p})\widehat{\boldsymbol{\theta}})(1 - \cos\theta), \end{aligned} \qquad \text{(A12)}$$

where $\mathbf{p}'$, $\mathbf{p}$, and $\boldsymbol{\theta}$ are all pure quaternions, and where we have $\theta \equiv |\boldsymbol{\theta}|$ and $\widehat{\boldsymbol{\theta}} \equiv \boldsymbol{\theta}/|\boldsymbol{\theta}|$. The third line of Eq. (A12) is arrived at by the rules of quaternion algebra.

The geometric interpretation of this relation is quite intuitive: $\boldsymbol{\theta}$ describes both the angle of rotation (through its modulus θ), and the axis of rotation (through $\widehat{\boldsymbol{\theta}}$). Note that the quaternions $\pm v$ corresponds to the same rotation—this is analogous to the 2 to 1 correspondence between $SU(2)$ and $SO(3)$. Indeed, there is an isomorphism between the description of rotations via unit quaternions and via $SU(2)$. An advantage of the quaternion description of rotations is that it does not run into issues of "gimbal lock" that afflict more conventional descriptions, such as Euler angles. This

robustness makes the quaternion description very attractive not just in a theoretical, but a practical point of view as well.

Finally, we note that composite rotations are also conveniently represented in terms of quaternions. It is straightforward to show that the quaternion product of two unit quaternions is also a unit quaternion, and so can also be used to represent a rotation. The quaternion corresponding to a rotation u_1 followed by a rotation u_2 is simply the quaternion product $u_3 = u_2 u_1$, which can be easily seen from the following:

$$\mathbf{p}' = u_3 \mathbf{p} \bar{u}_3 = u_2 u_1 \mathbf{p} \overline{u_2 u_1} = u_2(u_1 \mathbf{p} \bar{u}_1)\bar{u}_2, \qquad \text{(A13)}$$

where in going from the second to third lines, we used the following relation for the conjugate of the product of two quaternions $\overline{uv} = \bar{v}\bar{u}$.

APPENDIX B: RELATION BETWEEN $\mathbf{n}(t)$ AND $\mathbf{m}_\epsilon(t)$ FOR $\mathcal{T} e^{\frac{\epsilon}{2} \int_0^t dt' \, \mathbf{n}(t')} = e^{\frac{\epsilon}{2} \mathbf{m}_\epsilon(t)}$

Our goal is to find the relation between the pure quaternions $\mathbf{n}(t)$ and $\mathbf{m}_\epsilon(t)$ [they can equivalently be considered members of $\mathfrak{su}(2)$ due to the isomorphism between them]—we put a subscript ϵ in $\mathbf{m}_\epsilon(t)$ as a reminder that it is generally a function of ϵ. In order to find the relation between $\mathbf{n}(t)$ and $\mathbf{m}_\epsilon(t)$, we make use of the defining relation,

$$\mathcal{T} e^{\frac{\epsilon}{2} \int_0^t dt' \, \mathbf{n}(t')} = e^{\frac{\epsilon}{2} \mathbf{m}_\epsilon(t)}, \qquad \text{(B1)}$$

and the time derivative,

$$\frac{d}{dt} \mathcal{T} e^{\frac{\epsilon}{2} \int_0^t dt' \, \mathbf{n}(t')} = \frac{\epsilon}{2} \mathbf{n}(t) \mathcal{T} e^{\frac{\epsilon}{2} \int_0^t dt' \, \mathbf{n}(t')}, \qquad \text{(B2)}$$

in order to write the following expression,

$$\frac{\epsilon}{2} \mathbf{n}(t) = \left[\frac{d}{dt} e^{\frac{\epsilon}{2} \mathbf{m}_\epsilon(t)} \right] e^{-\frac{\epsilon}{2} \mathbf{m}_\epsilon(t)}. \qquad \text{(B3)}$$

To find the derivative of the exponential, we first rewrite it as an infinite product and get the result,

$$\frac{d}{dt} e^{\frac{\epsilon}{2} \mathbf{m}_\epsilon(t)} = \frac{d}{dt} \lim_{\substack{N \to \infty \\ \delta x \to 0 \\ N\delta x \to 1}} \prod_{i=0}^{N} e^{\frac{\epsilon}{2} \delta x \, \mathbf{m}_\epsilon(t)}$$

$$= \frac{\epsilon}{2} \lim_{\substack{N \to \infty \\ \delta x \to 0 \\ N\delta x \to 1}} \sum_{k=0}^{N} \delta x \left[\prod_{i=0}^{N-k} e^{\frac{\epsilon}{2} \delta x \, \mathbf{m}_\epsilon(t)} \right]$$

$$\times \frac{d\mathbf{m}_\epsilon(t)}{dt} \left[\prod_{j=k}^{N} e^{\frac{\epsilon}{2} \delta x \, \mathbf{m}_\epsilon(t)} \right]$$

$$= \frac{\epsilon}{2} \int_0^1 dx \, e^{\frac{\epsilon}{2} x \, \mathbf{m}_\epsilon(t)} \frac{d\mathbf{m}_\epsilon(t)}{dt} e^{\frac{\epsilon}{2}(1-x) \, \mathbf{m}_\epsilon(t)}. \qquad \text{(B4)}$$

Using this expression, we can rewrite Eq. (B3) as

$$\mathbf{n}(t) = \int_0^1 dx \, e^{\frac{\epsilon}{2} x \, \mathbf{m}_\epsilon(t)} \frac{d\mathbf{m}_\epsilon(t)}{dt} e^{-\frac{\epsilon}{2} x \, \mathbf{m}_\epsilon(t)}. \qquad \text{(B5)}$$

The geometric meaning of the integrand is simple: It corresponds to a rotation of $d\mathbf{m}_\epsilon(t)/dt$ by an angle $\epsilon x m_\epsilon(t)$ [where $m_\epsilon(t) \equiv |\mathbf{m}_\epsilon(t)|$] about the unit axis $\hat{\mathbf{m}}_\epsilon(t) \equiv \mathbf{m}_\epsilon(t)/m_\epsilon(t)$.

Using quaternion algebra [or equivalently the $\mathfrak{su}(2)$ algebra], we find

$$e^{\frac{\epsilon}{2} x \, \mathbf{m}_\epsilon(t)} \frac{d\mathbf{m}_\epsilon(t)}{dt} e^{-\frac{\epsilon}{2} x \, \mathbf{m}_\epsilon(t)}$$

$$= \frac{d\mathbf{m}_\epsilon(t)}{dt} + \hat{\mathbf{m}}_\epsilon(t) \wedge \frac{d\mathbf{m}_\epsilon(t)}{dt} \, \sin x\epsilon m_\epsilon(t)$$

$$+ \hat{\mathbf{m}}_\epsilon(t) \wedge \left(\hat{\mathbf{m}}_\epsilon(t) \wedge \frac{d\mathbf{m}_\epsilon(t)}{dt} \right)(1 - \cos x\epsilon m_\epsilon(t)). \qquad \text{(B6)}$$

With this, it is now trivial to carry out the integral in x in Eq. (B5). In order to simplify the expression further, we make use of the relation,

$$\frac{d\mathbf{m}_\epsilon(t)}{dt} = \hat{\mathbf{m}}_\epsilon(t) m_\epsilon(t) + \hat{\mathbf{m}}_\epsilon(t)\dot{m}_\epsilon(t), \qquad \text{(B7)}$$

where the overhead dot denotes a time derivative, along with the fact that $\hat{\mathbf{m}}_\epsilon(t)$ and $\dot{\hat{\mathbf{m}}}_\epsilon(t)$ are orthogonal since $\hat{\mathbf{m}}_\epsilon(t)$ is a unit pure quaternion. We obtain the expression,

$$\mathbf{n}(t) = \dot{m}_\epsilon(t)\hat{\mathbf{m}}_\epsilon(t) + \dot{\hat{\mathbf{m}}}_\epsilon(t)\frac{\sin \epsilon m_\epsilon(t)}{\epsilon}$$

$$+ \hat{\mathbf{m}}_\epsilon(t) \wedge \dot{\hat{\mathbf{m}}}_\epsilon(t)\frac{1 - \cos \epsilon m_\epsilon(t)}{\epsilon}. \qquad \text{(B8)}$$

This equation is very useful, as it gives us an exact expression for $\mathbf{n}(t)$ in terms of $\mathbf{m}_\epsilon(t)$ and its time derivatives. Fortunately it is possible to invert Eq. (B8) and unambiguously solve for both $\dot{m}_\epsilon(t)$ and $\dot{\hat{\mathbf{m}}}_\epsilon(t)$ in terms of $\mathbf{n}(t)$, $m_\epsilon(t)$, and $\hat{\mathbf{m}}_\epsilon(t)$, where we find after some algebra,

$$\dot{m}_\epsilon(t) = \mathbf{n}(t) \cdot \hat{\mathbf{m}}_\epsilon(t),$$

$$\dot{\hat{\mathbf{m}}}_\epsilon(t) = -\frac{\epsilon}{2} \hat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)$$

$$- \left(\frac{\epsilon}{2} \cot \frac{\epsilon m_\epsilon(t)}{2} \right) \hat{\mathbf{m}}_\epsilon(t) \wedge (\hat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)). \qquad \text{(B9)}$$

Using Eq. (B7), we can combine both equations in Eq. (B9) to find a single differential equation unambiguously relating $\mathbf{m}_\epsilon(t)$ and $\mathbf{n}(t)$. After a little algebra we find

$$\frac{d\mathbf{m}_\epsilon(t)}{dt} = \mathbf{n}(t) - \frac{\epsilon}{2} \mathbf{m}_\epsilon(t) \wedge \mathbf{n}(t) + \left(1 - \frac{\epsilon m_\epsilon(t)}{2} \cot \frac{\epsilon m_\epsilon(t)}{2} \right)$$

$$\times \hat{\mathbf{m}}_\epsilon(t) \wedge (\hat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)). \qquad \text{(B10)}$$

Note that Eq. (B10) is an exact relation between $\mathbf{m}_\epsilon(t)$ and $\mathbf{n}(t)$. Solving it is equivalent to summing up all terms in the Magnus expansion for the case where $\mathbf{n}(t)$ is an arbitrary time-dependent linear combination of elements of $\mathfrak{su}(2)$, or equivalently, any pure quaternion since the two are isomorphic. For cases where ϵ is large, Eq. (B10) can be used to give us the exact solution—it can also be used to generate a diagrammatic expansion in ϵ. By summing up certain classes of diagrams, we can obtain useful results even for moderately large ϵ.

For small ϵ, we can solve Eq. (B10) perturbatively in a fairly straightforward to arbitrary order. This is useful if one is interested in finding a series expansion for higher order contributions for the fidelity functional. We seek a solution of

the form,

$$\mathbf{m}_\epsilon(t) = \sum_{j=0}^{\infty} \epsilon^j \mathbf{m}^{(j)}(t), \tag{B11}$$

where $\mathbf{m}^{(j)}(t)$ is the jth order term in the perturbation theory [we drop the subscript ϵ for the perturbation terms $\mathbf{m}^{(j)}(t)$ since the ϵ dependence is accounted for in the prefactor ϵ^j]. Let us rewrite Eq. (B10) in a way that is more amenable to this perturbative treatment. We make use of the following series expansion [24]:

$$1 - \frac{x}{2}\cot\frac{x}{2} = \sum_{j=1}^{\infty} \frac{(-1)^{j+1} B_{2j} x^{2j}}{(2j)!}, \tag{B12}$$

where B_j are the first Bernoulli numbers. The first Bernoulli numbers can be obtained from the Bernoulli polynomials, defined through the generating function,

$$\frac{x e^{tx}}{e^x - 1} = \sum_{j=0}^{\infty} B_j(t) \frac{x^j}{j!}, \tag{B13}$$

where we have $B_j \equiv B_j(0)$, and are given explicitly by

$$B_j = \sum_{k=0}^{j} \sum_{\ell=0}^{k} \frac{(-1)^\ell k! \ell^j}{\ell!(k-\ell)!(k+1)}. \tag{B14}$$

Using Eq. (B12), we then see that the third term in Eq. (B10) only contains positive even powers of $\epsilon m_\epsilon(t)$, where in particular the leading order term is quadratic in ϵ. In a series expansion in ϵ, in the right-hand side of Eq. (B10), we get terms of the form,

$$(-1)^{j+1} m_\epsilon^{2j} \widehat{\mathbf{m}}_\epsilon(t) \wedge (\widehat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)).$$

These terms can be rewritten entirely in terms of $\mathbf{m}_\epsilon(t)$ *only*, simplifying the perturbative expansion since we do not have to consider the amplitude $m_\epsilon(t)$ and the unit axis $\widehat{\mathbf{m}}_\epsilon(t)$ separately. To do this, we introduce the *nested* wedge product, which can be defined recursively through the relation,

$$(\mathbf{a}\wedge)^j \mathbf{b} \equiv \mathbf{a} \wedge [(\mathbf{a}\wedge)^{j-1} \mathbf{b}], \tag{B15}$$

where j is an integer greater than or equal to zero, and where we use the convention $(\mathbf{a}\wedge)^0 \mathbf{b} = \mathbf{b}$. Using this definition for the nested wedge product, along with some quaternion algebra, we can show that

$$(-1)^{j+1} m_\epsilon^{2j} \widehat{\mathbf{m}}_\epsilon(t) \wedge (\widehat{\mathbf{m}}_\epsilon(t) \wedge \mathbf{n}(t)) = (\mathbf{m}_\epsilon(t)\wedge)^{2j} \mathbf{n}(t). \tag{B16}$$

Next, we take advantage of the fact that the odd Bernoulli numbers, B_{2j+1}, vanish for all integers $j > 0$ so that we can include the odd order nested products in the series expansion without changing its value. Using this fact along with the result $B_0 = 1$ and $B_1 = -1/2$ and the nested wedge product defined in Eq. (B16), we can rewrite Eq. (B10) in a simple unified way that is completely amenable to the perturbative expansion as given in Eq. (B11):

$$\frac{d\mathbf{m}_\epsilon(t)}{dt} = \sum_{j=0}^{\infty} \frac{\epsilon^j B_j}{j!} (\mathbf{m}_\epsilon(t)\wedge)^j \mathbf{n}(t). \tag{B17}$$

We now define the nested wedge product for N generally distinguishable factors, which can be, for example, factors

$\mathbf{m}^{(j_1)}(t), \mathbf{m}^{(j_2)}(t), \ldots, \mathbf{m}^{(j_N)}(t)$, appearing in the perturbative expansion. It is defined recursively as

$$\mathbf{m}^{\{j_1 j_2 \cdots j_N\}}(t) \wedge \mathbf{n}(t) \equiv \mathbf{m}^{(j_1)}(t) \wedge [\mathbf{m}^{\{j_2 \cdots j_N\}}(t) \wedge \mathbf{n}(t)]. \tag{B18}$$

For the special case where the number of factors N vanishes, i.e., $\{j_1 \ldots j_N\} = \{\emptyset\}$, we define $\mathbf{m}^{\{\emptyset\}}(t) \wedge \mathbf{n}(t) \equiv \mathbf{n}(t)$—with this convention all other cases are uniquely defined through Eq. (B18).

Using Eqs. (B11), (B17), and (B18), it is straightforward to find the expression for arbitrary $\mathbf{m}^{(j)}(t)$ in the perturbative expansion. The zeroth order $j = 0$ term is given by the expression,

$$\mathbf{m}^{(0)}(t) = \int_0^t dt' \, \mathbf{n}(t'), \tag{B19}$$

while for all other values $j > 0$ we find the result,

$$\mathbf{m}^{(j)}(t) = \sum_{k=1}^{j} \frac{B_k}{k!} \left\{ \sum_{\sum_{\ell=1}^{k} i_\ell = j-k} \int_0^t dt' \, \mathbf{m}^{\{i_1 i_2 \cdots i_N\}}(t') \wedge \mathbf{n}(t') \right\}, \tag{B20}$$

where the second sum over the set $i_1, i_2, \ldots, i_k$ in Eq. (B20) satisfies the constraint that $\sum_{\ell=1}^{k} i_\ell = j - k$. One can easily check that with this constraint, only $\mathbf{m}^{(k)}(t)$ for $0 \leqslant k \leqslant j-1$ contribute to the right-hand side of Eq. (B20), so that this equation gives us an explicit result for the jth order perturbation term entirely in terms of lower order perturbation terms $\mathbf{m}^{(k \leqslant j-1)}(t)$ and $\mathbf{n}(t)$. Equations (B19) and (B20) reproduce what is expected from the Magnus expansion [21], which is not surprising since the same approximation scheme has been used here. It can be easily shown that our perturbative result, Eqs. (B19) and (B20), readily generalizes to *all* Lie algebras, provided one replaces the nested wedge products with nested commutators (Lie brackets)—in particular the quaternion wedge product corresponds exactly to the commutator for $\mathfrak{su}(2)$. For the case of quaternions [also $\mathfrak{su}(2)$ due to the isomorphism], we have used quaternion methods to derive an *exact* differential equation Eq. (B10) in a straightforward manner, which when solved is equivalent to summing up all terms in the Magnus expansion. The result given in Eqs. (B19) and (B20), along with Eq. (B11), gives us a perturbative solution to the differential equation in powers of ϵ—this reproduces the Magnus expansion term by term.

Using Eqs. (B19) and (B20) we find explicit expressions for the first few terms in the perturbative expansion, given here,

$$\mathbf{m}^{(0)}(t) = \int_0^t dt_1 \, \mathbf{n}(t_1),$$

$$\mathbf{m}^{(1)}(t) = \frac{1}{2} \int_0^t dt_1 \int_0^{t_1} dt_2 \, \mathbf{n}(t_1) \wedge \mathbf{n}(t_2),$$

$$\mathbf{m}^{(2)}(t) = \frac{1}{6} \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \{\mathbf{n}(t_1) \wedge [\mathbf{n}(t_2) \wedge \mathbf{n}(t_3)]$$
$$+ \, \mathbf{n}(t_3) \wedge [\mathbf{n}(t_2) \wedge \mathbf{n}(t_1)]\}, \quad \text{etc.} \tag{B21}$$

The form of Eq. (B10) also suggests a different way of approximating the solution which is better than the Magnus expansion in the sense that we include contributions from *all* orders of ϵ at each iteration (with the exception of zeroth

order which coincides with the expression obtained from the Magnus expansion). Because of this, we then obtain a much better approximation which works well even in the case where ϵ is large. We denote the [j]th order approximation with a superscript [n] using brackets instead of parentheses to distinguish this from the Magnus expansion (note that we

now include the subscript ϵ since each term now *does* depend explicitly on ϵ). For the leading order we have

$$\mathbf{m}_\epsilon^{[0]}(t) = \int_0^t dt' \, \mathbf{n}(t'), \tag{B22}$$

and all higher order approximations $j > 0$ are given by

$$\mathbf{m}_\epsilon^{[j]}(t) = \int_0^t dt' \left\{ \mathbf{n}(t') - \frac{\epsilon}{2} \mathbf{m}_\epsilon^{[j-1]}(t') \wedge \mathbf{n}(t') + \left(1 - \frac{\epsilon m_\epsilon^{[j-1]}(t')}{2} \cot \frac{\epsilon m_\epsilon^{[j-1]}(t')}{2} \right) \widehat{\mathbf{m}}_\epsilon^{[j-1]}(t') \wedge (\widehat{\mathbf{m}}_\epsilon^{[j-1]}(t') \wedge \mathbf{n}(t')) \right\}. \tag{B23}$$

The approximate nth order expression is simply given by the nth iterate of this procedure,

$$\mathbf{m}_\epsilon(t) \simeq \mathbf{m}_\epsilon^{[n]}(t). \tag{B24}$$

This procedure is equivalent to solving the exact equation Eq. (B10) by iteration—when it is carried out to infinite order, it gives us the *exact* solution (as long as the solution obtained this way is *unique*, then it is unquestionably the solution), i.e., we have

$$\mathbf{m}_\epsilon(t) = \lim_{j \to \infty} \mathbf{m}_\epsilon^{[j]}(t). \tag{B25}$$

In practice we find that even for fairly large values of ϵ, this procedure converges very quickly to a unique answer, i.e., there is some finite order N upon which $\mathbf{m}_\epsilon^{[N]}(t)$ is practically indistinguishable from the exact solution. For extremely large values of ϵ this procedure may not converge—i.e., the iterates may jump back and forth between two or more different values—in which case one must modify the approximation to obtain a unique answer. One can always check whether this unique answer is the solution by plugging it back into Eq. (B10).

APPENDIX C: EVALUATION OF THE PATH INTEGRAL EXPRESSION FOR THE FIDELITY AMPLITUDE

Here we show how to evaluate the path integral expression for the fidelity amplitude [given in Eq. (26) in the main paper]. We rewrite it here for convenience [in this appendix, unlike the main paper, we write explicitly the normalization prefactor, $(2s + 1)^{-1}$, corresponding to the dimension of the Hilbert space],

$$\mathcal{A}_s = \frac{1}{2s+1} \text{Tr} \int D^2\boldsymbol{\alpha}_f D^2\boldsymbol{\alpha}_b \times |\boldsymbol{\alpha}_b(0)\rangle\langle\boldsymbol{\alpha}_f(0)|$$
$$\times e^{-(|\boldsymbol{\alpha}_b(\tau)|^2 + |\boldsymbol{\alpha}_f(0)|^2) + \boldsymbol{\alpha}_b^\star(\tau)\cdot\boldsymbol{\alpha}_f(\tau)}$$
$$\times e^{i\int_0^\tau dt[\boldsymbol{\alpha}_f^\star\cdot(i\partial_t - \mathcal{H}_\epsilon)\cdot\boldsymbol{\alpha}_f - \boldsymbol{\alpha}_b^\star\cdot(i\partial_t - \mathcal{H}_c)\cdot\boldsymbol{\alpha}_b]}, \tag{C1}$$

where $|\boldsymbol{\alpha}_{f,b}(0)\rangle$ are two-component *coherent states* corresponding to the two-mode operator $\hat{\mathbf{a}}^\dagger \equiv (\hat{a}_1^\dagger, \hat{a}_2^\dagger)$. The trace is taken over the complete set of states $|\psi\rangle$ satisfying the constraint $\langle\psi|\frac{1}{2}\hat{\mathbf{a}}^\dagger\hat{\mathbf{a}}|\psi\rangle = s$, fixing the quantum spin number s. The integral measure $D^2\boldsymbol{\alpha}_i$ for $i = f,b$ is given by the following expression:

$$D^2\boldsymbol{\alpha}_i \equiv \lim_{N\to\infty} \prod_{n=0}^N d^2\boldsymbol{\alpha}_i(t_n), \tag{C2}$$

where $t_n = n\tau/N$ and τ denotes the total transit time. The local (in time) measure $d^2\boldsymbol{\alpha}_i(t_n)$ is given by

$$d^2\boldsymbol{\alpha}_i(t_n) \equiv d^2\alpha_i^{(1)}(t_n) \, d^2\alpha_i^{(2)}(t_n), \tag{C3}$$

where the superscripts (1) and (2) denote the components of the two-component field $\boldsymbol{\alpha}_i$, and we have

$$d^2\alpha_i^{(j)}(t_n) \equiv \frac{d\alpha_i^{(j)}(t_n) \, d\alpha_i^{\star(j)}(t_n)}{2\pi i}$$
$$= \frac{d\text{Re}\alpha_i^{(j)}(t_n) \, d\text{Im}\alpha_i^{(j)}(t_n)}{\pi}, \tag{C4}$$

where $\text{Re}\alpha_i^{(j)}(t_n)$ and $\text{Im}\alpha_i^{(j)}(t_n)$ denote the real and imaginary parts of $\alpha_i^{(j)}(t_n)$, respectively. The first and second lines of Eq. (C4) are equivalent, and it is a matter of convenience as to which representation we use.

To evaluate Eq. (C1), we introduce the *generating functional* $G_s[\mathbf{J}]$:

$$G_s[\mathbf{J}] = \frac{1}{2s+1} \text{Tr} \int D^2\boldsymbol{\alpha}_f D^2\boldsymbol{\alpha}_b \times |\boldsymbol{\alpha}_b(0)\rangle\langle\boldsymbol{\alpha}_f(0)|$$
$$\times e^{-(|\boldsymbol{\alpha}_b(\tau)|^2 + |\boldsymbol{\alpha}_f(0)|^2) + \boldsymbol{\alpha}_b^\star(\tau)\cdot\boldsymbol{\alpha}_f(\tau)}$$
$$\times e^{i\int_0^\tau dt[\boldsymbol{\alpha}_f^\star\cdot(i\partial_t - \mathcal{H}_c)\cdot\boldsymbol{\alpha}_f - \boldsymbol{\alpha}_b^\star\cdot(i\partial_t - \mathcal{H}_c)\cdot\boldsymbol{\alpha}_b]}$$
$$\times e^{\int_0^\tau dt[\mathbf{J}^\star\cdot\boldsymbol{\alpha}_f + \boldsymbol{\alpha}_f^\star\cdot\mathbf{J}]}, \tag{C5}$$

where $\mathbf{J}$ and $\mathbf{J}^\star$ are two-component *source fields* which couple only to the *forward branch* of the path integral (due to the fact that the noise field only couples to this branch—see discussion in main paper). We can evaluate $\mathcal{A}_s$ (which includes the noise contribution) via the following expression:

$$\mathcal{A}_s = \exp\left\{ -i\epsilon \int_0^\tau dt \, \frac{\delta}{\delta\mathbf{J}(t)} \mathcal{H}_n(t) \frac{\delta}{\delta\mathbf{J}^\star(t)} \right\} G_s[\mathbf{J}] \bigg|_{\substack{\mathbf{J}(t)=0 \\ \mathbf{J}^\star(t)=0}}. \tag{C6}$$

Note that $G_s[\mathbf{J}]$ depends only on the noise-free term in the Hamiltonian (which shows up symmetrically in both forward and backward branches). For this reason, it is a simpler task to calculate $G_s[\mathbf{J}]$ first, and then use Eq. (C6) to obtain $\mathcal{A}_s$, as opposed to calculating $\mathcal{A}_s$ directly. Taking the generating functional route also holds the advantage that the physics is more transparent.

Our first goal is to evaluate $G_s[\mathbf{J}]$ exactly, which we show in the following subsection. Once we have this quantity we will use it to obtain an exact expression for the fidelity amplitude $\mathcal{A}_s$.

1. Evaluation of the generating functional $G_s[\mathbf{J}]$

In what follows, we shall evaluate the path integral expression for $G_s[\mathbf{J}]$ [defined in Eq. (C5)] *exactly*. We start with a change of variables (a *Keldysh rotation*) defined by

$$\boldsymbol{\psi}(t) = \frac{\boldsymbol{\alpha}_f(t) + \boldsymbol{\alpha}_b(t)}{2},$$

$$\boldsymbol{\eta}(t) = \boldsymbol{\alpha}_f(t) - \boldsymbol{\alpha}_b(t). \qquad (C7)$$

It is easy to check that the Jacobian associated with this change of variables is exactly unity. The description of the path integral in terms of these fields has a nice physical interpretation. The symmetric combination $\boldsymbol{\psi}$ is usually referred to as the "classical" component in the literature (it is the only component that survives in the classical limit, giving rise to a unique classical trajectory corresponding to the saddle point of $G_s[\mathbf{J}]$). The antisymmetric combination $\boldsymbol{\eta}$ is referred to as the "quantum" component since it accounts for deviations from the classical trajectory [12]. The evaluation of the path integral in terms of this set of variables simplifies greatly as we shall see below.

Let us rewrite the expression for $G_s[\mathbf{J}]$ in terms of this new set of variables. After some algebra (and integration by parts in order to move time derivatives from $\boldsymbol{\eta}$'s to $\boldsymbol{\psi}$'s) we obtain the following expression:

$$G_s[\mathbf{J}] = \int \mathrm{D}^2\boldsymbol{\psi}\,\mathrm{D}^2\boldsymbol{\eta}\; W(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star, \boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star)\, e^{-\frac{1}{2}|\boldsymbol{\eta}_\tau|^2}$$
$$\times\, e^{\int_0^\tau dt[i\boldsymbol{\eta}^\star\cdot(i\partial_t\boldsymbol{\psi}-\mathcal{H}_c\cdot\boldsymbol{\psi}-i\mathbf{J}/2)+\boldsymbol{\psi}^\star\cdot\mathbf{J}]}$$
$$\times\, e^{\int_0^\tau dt[(-i\partial_t\boldsymbol{\psi}^\star-\boldsymbol{\psi}^\star\cdot\mathcal{H}_c-i\mathbf{J}^\star/2)\cdot i\boldsymbol{\eta}+\mathbf{J}^\star\cdot\boldsymbol{\psi}]}, \qquad (C8)$$

where we use the shorthand notation $\boldsymbol{\psi}_0 \equiv \boldsymbol{\psi}(0)$, $\boldsymbol{\eta}_0 \equiv \boldsymbol{\eta}(0)$, and $\boldsymbol{\eta}_\tau \equiv \boldsymbol{\eta}(\tau)$. The function $W(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star, \boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star)$ in the integrand of Eq. (C8) is given by the expression,

$$W(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star, \boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star) = e^{\frac{1}{2}(\boldsymbol{\psi}_0^\star\cdot\boldsymbol{\eta}_0 - \boldsymbol{\eta}_0^\star\cdot\boldsymbol{\psi}_0)} \frac{1}{2s+1}$$
$$\times \mathrm{Tr}\left|\boldsymbol{\psi}_0 - \frac{\boldsymbol{\eta}_0}{2}\right\rangle\left\langle\boldsymbol{\psi}_0 + \frac{\boldsymbol{\eta}_0}{2}\right|, \qquad (C9)$$

where the states $|\boldsymbol{\psi}_0 \pm \frac{\boldsymbol{\eta}_0}{2}\rangle$ are coherent states, and we recall that the trace in Eq. (C9) is constrained to be over states for which the spin quantum number s is *fixed*, i.e., the constrained subspace contains the two-mode Fock states $|n_1, n_2\rangle$ for which $n_1 + n_2 = 2s$.

Before evaluating the path integral expression for $G_s[\mathbf{J}]$ in Eq. (C8), we take a moment to write down explicitly the form taken by the integral measures in the new set of variables, $\boldsymbol{\psi}$ and $\boldsymbol{\eta}$. We have

$$\mathrm{D}^2\boldsymbol{\psi} \equiv \lim_{N\to\infty} \prod_{n=0}^{N} d^2\boldsymbol{\psi}(t_n), \quad \mathrm{D}^2\boldsymbol{\eta} \equiv \lim_{N\to\infty} \prod_{n=0}^{N} d^2\boldsymbol{\eta}(t_n), \quad (C10)$$

with the local measures $d^2\boldsymbol{\psi}(t_n)$ and $d^2\boldsymbol{\eta}(t_n)$ taking the expected form,

$$d^2\boldsymbol{\psi}(t_n) \equiv d^2\psi^{(1)}(t_n)\, d^2\psi^{(2)}(t_n),$$
$$d^2\boldsymbol{\eta}(t_n) \equiv d^2\eta^{(1)}(t_n)\, d^2\eta^{(2)}(t_n), \qquad (C11)$$

where as usual, the superscripts denote the components of the fields. Importantly, the expression for the measures $d^2\psi^{(j)}(t_n)$

and $d^2\eta^{(j)}(t_n)$ slightly differ. Suppressing time arguments for brevity, we have

$$d^2\psi^{(j)} \equiv \frac{d\psi^{(j)}\,d\psi^{\star(j)}}{\pi i} = \frac{2\,d\mathrm{Re}\psi^{(j)}\,d\mathrm{Im}\psi^{(j)}}{\pi},$$
$$d^2\eta^{(j)} \equiv \frac{d\eta^{(j)}\,d\eta^{\star(j)}}{4\pi i} = \frac{d\mathrm{Re}\eta^{(j)}\,d\mathrm{Im}\eta^{(j)}}{2\pi}, \qquad (C12)$$

where we see that the measure for $\boldsymbol{\psi}$ is a factor of four larger than the measure for $\boldsymbol{\eta}$. Note that the expressions in the second and third columns are entirely equivalent representations, and one can simply choose whichever representation is most convenient when performing calculations.

Going back to the expression for $G_s[\mathbf{J}]$ in Eq. (C8), we can immediately evaluate the integrals over $\boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star, \boldsymbol{\eta}_\tau$, and $\boldsymbol{\eta}_\tau^\star$. The only contribution to the integrand for the fields $\boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star$ that survives the time-continuum limit of the path integral comes from the term $W(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star, \boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star)$ given in Eq. (C9), giving us the result,

$$\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star) \equiv \int d^2\boldsymbol{\eta}_0\, W(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star, \boldsymbol{\eta}_0, \boldsymbol{\eta}_0^\star)$$
$$= \frac{(-1)^{2s}}{2s+1} L_{2s}^{(1)}(4|\boldsymbol{\psi}_0|^2) e^{-2|\boldsymbol{\psi}_0|^2}, \qquad (C13)$$

where $L_{2s}^{(1)}(x)$ is an associated Laguerre polynomial, which we express here in terms of its Rodrigues representation,

$$L_n^{(k)}(x) = \frac{e^x x^{-k}}{n!} \frac{d^n}{dx^n}(e^{-x} x^{k+n})$$
$$= \sum_{j=0}^{n} \frac{(k+n)!}{(n-j)!(k+j)!} \frac{(-x)^j}{j!}. \qquad (C14)$$

In the process of evaluating Eq. (C13), we made use of the following expression giving the overlap between the two-component coherent states $|\boldsymbol{\psi}_0 \pm \frac{1}{2}\boldsymbol{\eta}_0\rangle$ and the two-component Fock states $|n, 2s - n\rangle$:

$$\left\langle n, 2s-n \middle| \boldsymbol{\psi}_0 \pm \frac{1}{2}\boldsymbol{\eta}_0 \right\rangle = \frac{e^{-\frac{1}{2}|\boldsymbol{\psi}_0 \pm \frac{1}{2}\boldsymbol{\eta}_0|^2}}{\sqrt{n!(2s-n)!}} \left(\psi_0^{(1)} \pm \frac{1}{2}\eta_0^{(1)}\right)^n$$
$$\times \left(\psi_0^{(2)} \pm \frac{1}{2}\eta_0^{(2)}\right)^{2s-n}, \qquad (C15)$$

where the superscripts (1) and (2) on the right-hand side of Eq. (C15) denote the components of the respective fields. The function $\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star)$ given in Eq. (C13), referred to as the *Wigner function* in the context of quantum dynamics, can be interpreted as a distribution over the fields $\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star$. Note that this distribution function, as described by $\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star)$, is generally *not* positive definite, and in our case it shows oscillatory behavior. As far as the integral over $\boldsymbol{\eta}_\tau, \boldsymbol{\eta}_\tau^\star$ is concerned, the only contribution that survives the time-continuum limit is given by

$$\int d^2\boldsymbol{\eta}_\tau\, e^{-\frac{1}{2}|\boldsymbol{\eta}_\tau|^2} = 1. \qquad (C16)$$

We now proceed to evaluate the integrals over the fields $\boldsymbol{\eta}(t), \boldsymbol{\eta}^\star(t)$ in the bulk [i.e., for all $t > 0$ in the path integral given in Eq. (C8)]. The change of variables introduced in Eq. (C8) really simplifies things here. The integrals are readily

evaluated and result in a product of *Dirac delta functions*,

$$\lim_{N \to \infty} \prod_{n=0}^{N-1} [\delta(\boldsymbol{\psi}(t_{n+1}) + i\delta t[\mathcal{H}_c(t_n) \cdot \boldsymbol{\psi}(t_n) + i\mathbf{J}(t_n)/2])$$

$$\times \delta(\boldsymbol{\psi}^\star(t_{n+1}) - i\delta t[\boldsymbol{\psi}^\star(t_n) \cdot \mathcal{H}_c(t_n) + i\mathbf{J}^\star(t_n)/2])(2\pi)^2], \tag{C17}$$

where $\delta t \equiv \tau/N$. This result makes the evaluation of the integrals over $\boldsymbol{\psi}(t)$ and $\boldsymbol{\psi}^\star(t)$ (for $t > 0$) simple as well. The delta functions constrain the fields $\boldsymbol{\psi}(t), \boldsymbol{\psi}^\dagger(t)$ (for $t > 0$) to obey the following equations of motion,

$$i\partial_t \boldsymbol{\psi} = \mathcal{H}_c \cdot \boldsymbol{\psi} + \frac{i\mathbf{J}}{2},$$
$$i\partial_t \boldsymbol{\psi}^\star = -\boldsymbol{\psi}^\star \cdot \mathcal{H}_c - \frac{i\mathbf{J}^\star}{2}, \tag{C18}$$

where we have suppressed time arguments for brevity. Assuming general initial conditions $\boldsymbol{\psi}(0) = \boldsymbol{\psi}_0$ we can write down the exact solution for all $t > 0$,

$$\boldsymbol{\psi}(t) = \mathcal{U}_c(t,0) \cdot \boldsymbol{\psi}_0 + \frac{1}{2} \int_0^t dt' \, \mathcal{U}_c(t,t') \cdot \mathbf{J}(t'),$$
$$\boldsymbol{\psi}^\star(t) = \boldsymbol{\psi}_0^\star \cdot \mathcal{U}_c(0,t) - \frac{1}{2} \int_0^t dt' \, \mathbf{J}^\star(t') \cdot \mathcal{U}_c(t',t), \tag{C19}$$

where we have

$$\mathcal{U}_c(t,t') = \mathcal{T} \exp\left[-i \int_{t'}^t \mathcal{H}_c(t'') dt''\right]. \tag{C20}$$

Note that in general we have $\boldsymbol{\psi}^\star(t) \neq [\boldsymbol{\psi}(t)]^\star$, as one can see from Eq. (C19), and these fields have to be treated as being independent of one another. This also applies to the fields $\boldsymbol{\psi}_0^\star$ and $\boldsymbol{\psi}_0$.

Applying the results obtained in Eqs. (C13), (C16), (C17), and (C19) to Eq. (C8), we get the following expression for $G_s[\mathbf{J}]$:

$$G_s[\mathbf{J}] = e^{\frac{1}{2} \int_0^\tau dt \int_0^t dt' [\mathbf{J}^\star(t) \cdot \mathcal{U}_c(t,t') \cdot \mathbf{J}(t') - \mathbf{J}^\star(t') \cdot \mathcal{U}_c(t',t) \cdot \mathbf{J}(t)]} \int d^2\boldsymbol{\psi}_0$$

$$\times \mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star) \, e^{\int_0^\tau dt [\mathbf{J}^\star(t) \cdot \mathcal{U}_c(t,0) \cdot \boldsymbol{\psi}_0 + \boldsymbol{\psi}_0^\star \cdot \mathcal{U}_c(0,t) \cdot \mathbf{J}(t)]}. \tag{C21}$$

All that remains to be done in order to obtain the final result for $G_s[\mathbf{J}]$ is the evaluation of ordinary integrals over the fields $\boldsymbol{\psi}_0$ and $\boldsymbol{\psi}_0^\star$. This is a Gaussian integral with a polynomial prefactor in the integrand [see Eq. (C13)], and it can therefore be evaluated exactly. Before doing so we rewrite the expression for $\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star)$ by making use of the following integral representation for the associated Laguerre polynomial,

$$L_n^{(k)}(x) = \oint \frac{dz}{2\pi i} \frac{(-1)^k e^{\frac{xz}{1+z}}}{(1+z)^{k+1} z^{n+1}}, \tag{C22}$$

where the contour runs counterclockwise and encloses the pole at $z = 0$, but not the pole at $z = 1$. We can use the expression in Eq. (C22) to rewrite $\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star)$ as follows:

$$\mathcal{W}(\boldsymbol{\psi}_0, \boldsymbol{\psi}_0^\star) = \frac{1}{2s+1} \oint \frac{dz}{2\pi i} \frac{e^{-2\left(\frac{1-z}{1+z}\right)|\boldsymbol{\psi}_0|^2}}{(1+z)^2 z^{2s+1}}. \tag{C23}$$

This greatly simplifies the evaluation of the integrals over $\boldsymbol{\psi}_0$ and $\boldsymbol{\psi}_0^\star$ in Eq. (C21), and we obtain the following result:

$$\int d^2\boldsymbol{\psi}_0 \cdots = \oint \frac{dz}{2\pi i} \frac{e^{\frac{1}{2}\left(\frac{1+z}{1-z}\right) \int_0^\tau \int_0^\tau dt \, dt' \, \mathbf{J}^\star(t) \cdot \mathcal{U}_c(t,t') \cdot \mathbf{J}(t')}}{(2s+1)(1-z)^2 z^{2s+1}}. \tag{C24}$$

It is a simple matter now to combine the expression in Eq. (C24) with the prefactor in Eq. (C21) to obtain the expression for $G_s[\mathbf{J}]$. After some algebra, we finally get

$$G_s[\mathbf{J}] = \oint \frac{dz}{2\pi i} \frac{e^{\int_0^\tau \int_0^\tau dt \, dt' \, \mathbf{J}^\star(t) \cdot \mathcal{G}_c(z,t,t') \cdot \mathbf{J}(t')}}{(2s+1)(1-z)^2 z^{2s+1}}, \tag{C25}$$

where $\mathcal{G}_c(z,t,t')$ (which plays a role similar to that of a *Green's function*) is given by the expression,

$$\mathcal{G}_c(z,t,t') = \left(\frac{z}{1-z} + \Theta(t-t')\right) \mathcal{U}_c(t,t'). \tag{C26}$$

The quantity $\Theta(t - t')$ denotes the Heaviside (unit step) function, defined as

$$\Theta(t - t') = \begin{cases} 1 & \text{for } t > t' \\ 0 & \text{for } t < t' \end{cases}. \tag{C27}$$

It is straightforward to evaluate the integral over z in Eq. (C25) if we wish, but the expression for $G_s[\mathbf{J}]$ is much more useful for doing calculations as is, and for this reason this is the expression used in the main paper [see Eq. (29)]. The (continuous) variable z can be thought of as a *conjugate variable* to the (discrete) variable s, and the right-hand side of Eq. (C25) can be interpreted as an *integral transform* between the z representation and the s representation, i.e.,

$$G_s[\mathbf{J}] = \oint \frac{dz}{2\pi i} \frac{\widetilde{G}_z[\mathbf{J}]}{(2s+1)(1-z)^2 z^{2s+1}}, \tag{C28}$$

where we have

$$\widetilde{G}_z[\mathbf{J}] \equiv e^{\int_0^\tau \int_0^\tau dt \, dt' \, \mathbf{J}^\star(t) \cdot \mathcal{G}_c(z,t,t') \cdot \mathbf{J}(t')}. \tag{C29}$$

In practice, it is easier to compute quantities using the continuous z representation generating functional $\tilde{G}_z[\mathbf{J}]$ (where the expression is a simple Gaussian) and take the transform back into the s representation as a final step. In the next subsection of this appendix, we use $G_s[\mathbf{J}]$ to calculate the fidelity amplitude $\mathcal{A}_s$.

2. Using $G_s[\mathbf{J}]$ to calculate the fidelity amplitude $\mathcal{A}_s$

With the expression for $G_s[\mathbf{J}]$ [see Eq. (C25)], we can make use of Eq. (C6) to calculate $\mathcal{A}_s$ by taking functional derivatives of $G_s[\mathbf{J}]$. Because $G_s[\mathbf{J}]$ is given by Gaussian, the right-hand side of Eq. (C6) can be rewritten in an entirely equivalent representation by way of introducing *fictitious two-component quantum fields* $\boldsymbol{\phi}(t)^\dagger \equiv (\phi_1^\star(t), \phi_2^\star(t))$ that are completely described in terms of the expectation value,

$$\langle \boldsymbol{\phi}(t) \rangle_q = \langle \boldsymbol{\phi}^\star(t) \rangle_q = \mathbf{0},$$
$$\langle \boldsymbol{\phi}(t) \boldsymbol{\phi}^\dagger(t) \rangle_q = \mathcal{G}_c(z,t,t'), \tag{C30}$$

where the Green's function $\mathcal{G}_c(z,t,t')$ is the same one given in Eq. (C26), and $\langle \cdot \rangle_q$ denotes a quantum expectation value over the fields $\boldsymbol{\phi}(t)$ with all higher order expectation values entirely determined through the use of Wick's theorem. With the aid

of Eq. (C30), it is straightforward to show that the expression (suppressing time arguments for brevity),

$$\mathcal{A}_s = \oint \frac{dz}{2\pi i} \frac{\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\dagger \mathcal{H}_n \boldsymbol{\phi}} \rangle_q}{(2s+1)(1-z)^2 z^{2s+1}}, \qquad (C31)$$

is equivalent to the defining relation for $\mathcal{A}_s$ given in Eq. (C6). This formulation, given in terms of the dynamics of the fictitious fields $\boldsymbol{\phi}$, has the advantage of being physically more intuitive.

We shall now evaluate the quantum expectation value by making use of the following cumulant expansion:

$$\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\dagger \mathcal{H}_n \boldsymbol{\phi}} \rangle_q = \exp \left\{ \sum_{m=1}^\infty \mathcal{X}_m \right\}, \qquad (C32)$$

where the quantities $\mathcal{X}_m$ are defined by the relation,

$$\mathcal{X}_m \equiv \frac{1}{m!} \left\langle\!\!\left\langle \left(-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\dagger \mathcal{H}_n \boldsymbol{\phi} \right)^m \right\rangle\!\!\right\rangle_q . \qquad (C33)$$

The double brackets $\langle\!\langle \cdot \rangle\!\rangle_q$ on the right-hand side of Eq. (C33) denote cumulant averages. It is a well-known fact that in a diagrammatic expansion only *connected* diagrams contribute to cumulant averages. In what follows, we shall use the diagrammatic expansion to *exactly* evaluate the right-hand side of Eq. (C32), i.e., we will perform an exact resummation of all connected diagrams.

Given the form of Eq. (C32), we see that each diagram contributing to $\mathcal{X}_m$ contains *exactly m* vertices and m propagators. Considering this together with the fact that only connected diagrams contribute (since we are taking cumulant averages) places a severe restriction on the form allowed for the diagrams: All contributing diagrams have the topology of a *single* non-self-crossing *closed loop*. For convenience, we assign a definite orientation to this loop by placing arrows on the propagators (but note that different orientations are not distinct and should not be counted as such).

The Feynman rules are easily determined, and we state them here. In a given diagram, each vertex (with its time label t_i) corresponds to a factor $\mathcal{H}_n(t_i)$, and each propagator with starting point t_i and ending point t_j (as determined from the direction of the arrow) corresponds to the Green's function $\mathcal{G}_c(z, t_i, t_j)$. To obtain the value of an mth order diagram (i.e., one contributing to $\mathcal{X}_m$), start at an arbitrary vertex and write down all of the factors corresponding to each vertex and propagator in the order determined by the direction of the arrows (taking care not to count the starting point twice). Then simply integrate over all time variables, take the matrix trace, and multiply by an overall prefactor $(-i\epsilon)^m / m$. That this prefactor is $(m-1)!$ times larger than expected from Eq. (C33) simply comes from the fact that for a given order m there are $(m-1)!$ diagrams which differ only in their labels. Since all labels are dummy variables, all $(m-1)!$ diagrams give the same numerical contribution, hence the overall prefactor. To obtain the entire contribution to $\mathcal{X}_m$, it suffices to evaluate a single representative diagram, so, i.e., in Fig. 4(c) we only need to evaluate either the left or right diagram, but not both.

The diagrams corresponding to $\mathcal{X}_m$ for $m = 1,2,3$ are given in Figs. 4(a), 4(b), and 4(c), respectively. The integer labels displayed ($i = 1,2,\ldots$ etc.) correspond the time labels t_i. We
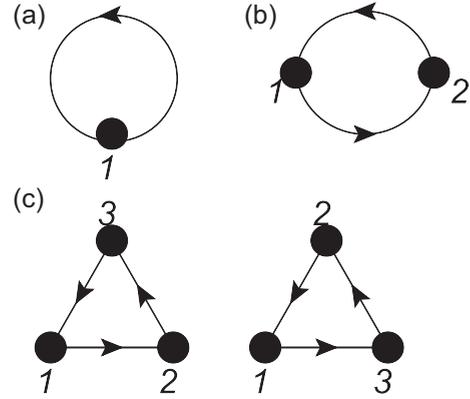

FIG. 4. (a) Diagram corresponding to entire contribution for $\mathcal{X}_1$. (b) Diagram corresponding to entire contribution to $\mathcal{X}_2$. (c) The 2! diagrams corresponding to the entire contribution for $\mathcal{X}_3$. Both diagrams give the same contribution, since all variables are internal, so in practice one only need evaluate a single representative diagram (see discussion in text). Likewise, for a given order m, all $(m-1)!$ diagrams give the exact same contribution. The expressions corresponding to the diagrams given in (a), (b), and (c) are given in Eqs. (C34a), (C34b), and (C34c), respectively (see text for Feynman rules). For all diagrams, the integer labels $i = 1,2,\ldots$, etc. next to the points refer to the time labels t_i, which are all dummy variables to be integrated over. Though we have chosen a counterclockwise orientation for the diagrams, one is free to choose whichever orientation one wishes (i.e., diagrams with different orientations are the exact same diagram).

have arbitrarily chosen to use a counterclockwise orientation when drawing the loop diagrams, though as we noted earlier different orientations correspond the the same diagram. As we have stated earlier, both diagrams shown in Fig. 4(c) give the same contribution, since all labels correspond to internal variables. As an example, we evaluate the diagrams shown in Fig. 4, giving expressions corresponding to $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$. We find

$$\mathcal{X}_1 = \frac{(-i\epsilon)^1}{1} \int_0^\tau dt_1 \, \mathrm{Tr}\, \mathcal{H}_n(t_1) \mathcal{G}_c(z, t_1, t_1), \qquad (C34a)$$

$$\mathcal{X}_2 = \frac{(-i\epsilon)^2}{2} \int_0^\tau dt_1 \int_0^\tau dt_2 \, \mathrm{Tr}[\mathcal{H}_n(t_1) \mathcal{G}_c(z, t_1, t_2) \\ \times \mathcal{H}_n(t_2) \mathcal{G}_c(z, t_2, t_1)], \qquad (C34b)$$

$$\mathcal{X}_3 = \frac{(-i\epsilon)^3}{3} \int_0^\tau dt_1 \int_0^\tau dt_2 \int_0^\tau dt_3 \, \mathrm{Tr}\,[\mathcal{H}_n(t_1) \mathcal{G}_c(z, t_1, t_2) \\ \times \mathcal{H}_n(t_2) \mathcal{G}_c(z, t_2, t_3) \mathcal{H}_n(t_3) \mathcal{G}_c(z, t_3, t_1)], \qquad (C34c)$$

and it is straightforward to see to how the expression generalizes for arbitrary $\mathcal{X}_m$.

Now that we know the expression for all $\mathcal{X}_m$, all that remains is to evaluate the sum $\sum_{m=1}^\infty \mathcal{X}_m$. Let us first introduce the shorthand notation,

$$\zeta \equiv \frac{z}{1-z},$$
$$\Theta_{ij} \equiv \Theta(t_i - t_j), \qquad (C35)$$

and rewrite the expression for $\mathcal{X}_m$ as follows:

$$
\begin{aligned}
\mathcal{X}_m &= \frac{(-i\epsilon)^m}{m} \int_0^\tau dt_1 \int_0^\tau dt_2 \cdots \int_0^\tau dt_m \, \mathrm{Tr}[\mathcal{H}_n(t_1)\mathcal{G}_c(z,t_1,t_2)\mathcal{H}_n(t_2)\mathcal{G}_c(z,t_2,t_3)\cdots\mathcal{H}_n(t_m)\mathcal{G}_c(z,t_m,t_1)] \\
&= \frac{(-i\epsilon)^m}{m} \int_0^\tau dt_1 \int_0^\tau dt_2 \cdots \int_0^\tau dt_m \, \mathrm{Tr}[\mathcal{H}_n(t_1)\mathcal{U}_c(t_1,t_2)\mathcal{H}_n(t_2)\mathcal{U}_c(t_2,t_3)\cdots\mathcal{H}_n(t_m)\mathcal{U}_c(t_m,t_1)] \\
&\quad \times (\zeta + \Theta_{12})(\zeta + \Theta_{23})\cdots(\zeta + \Theta_{m1}) \\
&= \frac{(-i\epsilon)^m}{m} \int_0^\tau dt_1 \int_0^\tau dt_2 \cdots \int_0^\tau dt_m \, \mathrm{Tr}[(\mathcal{U}_c^\dagger(t_1,0)\mathcal{H}_n(t_1)\mathcal{U}_c(t_1,0))\cdots(\mathcal{U}_c^\dagger(t_m,0)\mathcal{H}_n(t_m)\mathcal{U}_c(t_m,0))] \\
&\quad \times (\zeta + \Theta_{12})(\zeta + \Theta_{23})\cdots(\zeta + \Theta_{m1}) \\
&= \frac{(-i\epsilon)^m}{m} \int_0^\tau dt_1 \cdots \int_0^\tau dt_m \, \mathrm{Tr}[\widetilde{\mathcal{H}}_n(t_1)\cdots\widetilde{\mathcal{H}}_n(t_m)](\zeta + \Theta_{12})\cdots(\zeta + \Theta_{m1}),
\end{aligned}
\tag{C36}
$$

where in the second line of Eq. (C36) we simply use the definition of $\mathcal{G}_c(z,t,t')$ [see Eq. (C26)], in the third line we use the property $\mathcal{U}_c(t,t') = \mathcal{U}_c(t,0)\mathcal{U}_c(0,t') = \mathcal{U}_c(t,0)\mathcal{U}_c^\dagger(t',0)$ along with the cyclic property of traces $\mathrm{Tr}\,AB\cdots YZ = \mathrm{Tr}\,ZAB\cdots Y$, and in the fourth line the quantity $\widetilde{\mathcal{H}}_n(t)$ is defined as

$$
\widetilde{\mathcal{H}}_n(t) = \mathcal{U}_c(t,0)^\dagger \mathcal{H}_n(t)\mathcal{U}_c(t,0).
\tag{C37}
$$

Note that $\widetilde{\mathcal{H}}_n(t)$ is simply the noise Hamiltonian in the rotating frame, or equivalently, in the interaction picture [with respect to the control Hamiltonian $\mathcal{H}_c(t,0)$].

At first sight, it may seem that the sum $\sum_m \mathcal{X}_m$ cannot be explicitly carried out since the expression for $\mathcal{X}_m$, as given in Eq. (C36), does not factorize (the factors of Θ_{ij} make it impossible to factorize the time integrals). This situation can be remedied by rearranging $\sum_m \mathcal{X}_m$ into a power series in ζ,

$$
\sum_{m=1}^\infty \mathcal{X}_m = \sum_{m=1}^\infty \mathcal{Y}_m \zeta^m.
\tag{C38}
$$

Unlike $\mathcal{X}_m$, the coefficients $\mathcal{Y}_m$ *do* factorize, making this rearrangement advantageous. They take the form,

$$
\mathcal{Y}_m = \mathrm{Tr}\,\frac{\Delta^m}{m},
\tag{C39}
$$

where Δ is a 2×2 matrix to be determined below. With the expression in Eq. (C39), we can evaluate the sum in Eq. (C38) and obtain the result,

$$
\sum_{m=1}^\infty \mathcal{X}_m = \mathrm{Tr}\log(\mathbb{1} - \zeta\Delta)^{-1},
\tag{C40}
$$

where $\mathbb{1}$ is the 2×2 unit matrix.

We will now show that Eq. (C39) is true, finding the matrix Δ in the process. We note that an *infinite set* of $\mathcal{X}_q$ contribute to $\mathcal{Y}_m$, since *all* $\mathcal{X}_q$ with $q \geq m$ contain terms proportional to ζ^m. Starting with the first term, $\mathcal{Y}_0$ (i.e., the term in the sum $\sum_m \mathcal{X}_m$ proportional to ζ^0), we find that it exactly vanishes. This is due to the fact that the contribution coming from each $\mathcal{X}_m$ is proportional to $\Theta_{12}\Theta_{23}\cdots\Theta_{m1}$, and in order for this to be nonvanishing we need $t_1 > t_2 \cdots > t_{m-1} > t_m > t_1$, an impossibility.

Next, we seek the expression for $\mathcal{Y}_1$ by collecting all terms linear in ζ from $\sum_m \mathcal{X}_m$. It is easy to see from Eq. (C36) that a given $\mathcal{X}_m$ contributes exactly m terms, furthermore all of these

terms give the exact same contribution since they only differ in the labeling of internal variables. It then suffices to take a single representative and multiply the result my m. Let us take as the representative the term proportional to $\Theta_{12}\Theta_{23}\cdots\Theta_{m-1,m}$. The effect of this factor is to simply cut off the limits in the time integrals so that $t_1 > t_2 \cdots > t_m$, and we are left with a *time ordered* sequence of factors $\widetilde{\mathcal{H}}_n(t_1)\widetilde{\mathcal{H}}_n(t_2)\cdots\widetilde{\mathcal{H}}_n(t_m)$ which can be easily factorized via the help of the time ordering operator $\mathcal{T}$. The (linear in ζ) contribution from $\mathcal{X}_m$ is given explicitly by

$$
\begin{aligned}
\mathcal{X}_m &\to \zeta(-i\epsilon)^m \mathrm{Tr}\int_0^\tau dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{m-1}} dt_m \, \widetilde{\mathcal{H}}_n(t_1) \\
&\quad \times \widetilde{\mathcal{H}}_n(t_2)\cdots\widetilde{\mathcal{H}}_n(t_m) \\
&= \zeta\,\mathrm{Tr}\,\frac{1}{m!}\mathcal{T}\left[-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_1(t)\right]^m.
\end{aligned}
\tag{C41}
$$

Collecting all terms in $\sum_m \mathcal{X}_m$ proportional to ζ, we obtain

$$
\begin{aligned}
\mathcal{Y}_1 &= \mathrm{Tr}\,\mathcal{T}\left\{\sum_{m=1}^\infty \frac{1}{m!}\left[-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right]^m\right\} \\
&= \mathrm{Tr}\left[\mathcal{T}e^{-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_1(t)} - \mathbb{1}\right] \\
&\equiv \mathrm{Tr}\,\Delta,
\end{aligned}
\tag{C42}
$$

giving us the sought for expression for the matrix Δ.

Now we proceed to find $\mathcal{Y}_2$, showing that the general result given in Eq. (C39) holds. Looking at Eq. (C36), we see that all $\mathcal{X}_m$ for $m \geq 2$ contain a term proportional to ζ^2 and therefore contribute to $\mathcal{Y}_2$. In order to more easily understand the pattern that emerges, let us consider the contributions from the first few $\mathcal{X}_m$. The lowest order term to contribute to $\mathcal{Y}_2$ is $\mathcal{X}_2$, from which we get the expression,

$$
\begin{aligned}
\mathcal{X}_2 &\to \zeta^2\frac{(-i\epsilon)^2}{2}\mathrm{Tr}\int_0^\tau dt_1 \int_0^\tau dt_2 \, \widetilde{\mathcal{H}}_n(t_1)\widetilde{\mathcal{H}}_n(t_2) \\
&= \frac{\zeta^2}{2}\mathrm{Tr}\left[\frac{1}{1!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)\right. \\
&\quad \left. \times \frac{1}{1!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)\right],
\end{aligned}
\tag{C43}
$$

where in the second line of Eq. (C43) we have just rewritten the expression in a way that suggests what the pattern will be for higher order terms.

In order to more clearly see the pattern that emerges, let us work out the ζ^2 order contribution coming from $\mathcal{X}_3$. We have

$$
\begin{aligned}
\mathcal{X}_3 \to {} & \zeta^2 \frac{(-i\epsilon)^3}{3} \mathrm{Tr}\left[\int_0^\tau dt_1 \int_0^\tau dt_2 \int_0^\tau dt_3\, \widetilde{\mathcal{H}}_n(t_1)\widetilde{\mathcal{H}}_n(t_2)\widetilde{\mathcal{H}}_n(t_3)(\Theta_{12} + \Theta_{23} + \Theta_{31})\right] \\
= {} & \zeta^2(-i\epsilon)^3 \mathrm{Tr}\left[\int_0^\tau dt_1 \int_0^\tau dt_2 \int_0^\tau dt_3\, \widetilde{\mathcal{H}}_n(t_1)\widetilde{\mathcal{H}}_n(t_2)\widetilde{\mathcal{H}}_n(t_3)\,\Theta_{12}\right] \\
= {} & \zeta^2(-i\epsilon)^3 \mathrm{Tr}\left[\left(\int_0^\tau dt_1 \int_0^{t_1} dt_2\, \widetilde{\mathcal{H}}_n(t_1)\widetilde{\mathcal{H}}_n(t_2)\right)\left(\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)\right] \\
= {} & \frac{\zeta^2}{2}\mathrm{Tr}\left[\frac{1}{2!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^2 \frac{1}{1!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right) + \frac{1}{1!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)\frac{1}{2!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^2\right] \\
= {} & \frac{\zeta^2}{2}\,\mathrm{Tr}\,\sum_{j=1}^{2} \frac{1}{j!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^j \frac{1}{(2-j)!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^{2-j}.
\end{aligned} \tag{C44}
$$

In the second line we have used the fact that the terms proportional to Θ_{23} and Θ_{31} are identical to the one proportional to Θ_{12}. This can be seen by simply relabeling variables along with taking cyclic permutations of $\widetilde{\mathcal{H}}_n(t_i)$, since the trace operation remains invariant under this. In the third line, we simply apply Θ_{12} to cut off the integral over t_2, and note that the expression factorizes as the parentheses suggest. In the third line we rewrite the expression as a sum of two terms, via the use of time-ordering operators $\mathcal{T}$, and by once again taking advantage of the invariance of the trace operation under cyclic permutations of matrices. In the final line, we rewrite the sum in a way that is suggestive of how higher order $\mathcal{X}_m$ contribute to $\mathcal{Y}_2$.

As one can easily guess from the expression in Eq. (C44) (we do not show the proof explicitly here, though it is easy to prove by induction) the contribution to $\mathcal{Y}_2$ coming from $\mathcal{X}_m$ for general m is given by the expression,

$$
\begin{aligned}
\mathcal{X}_m \to {} & \frac{\zeta^2}{2}\,\mathrm{Tr}\left[\sum_{j=1}^{m-1} \frac{1}{j!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^j\right. \\
& \left.\times \frac{1}{(m-j)!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^{m-j}\right].
\end{aligned} \tag{C45}
$$

We are now ready to write down the entire contribution to $\mathcal{Y}_2$, coming from all $\mathcal{X}_m$, where we find

$$
\begin{aligned}
\mathcal{Y}_2 = {} & \frac{1}{2}\sum_{m=2}^{\infty}\mathrm{Tr}\left[\sum_{j=1}^{m-1}\frac{1}{j!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^j\right. \\
& \left.\times \frac{1}{(m-j)!}\mathcal{T}\left(-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)\right)^{m-j}\right] \\
= {} & \mathrm{Tr}\left[\frac{1}{2}\left(\mathcal{T}e^{-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)} - \mathbb{1}\right)^2\right] = \mathrm{Tr}\,\frac{\Delta^2}{2},
\end{aligned} \tag{C46}
$$

where we see that in the last line of Eq. (C46), the matrix Δ is the same matrix appearing in Eq. (C42).

Continuing with the same procedure used above to determine $\mathcal{Y}_1$ and $\mathcal{Y}_2$, it is simple to show (though we do not explicitly show the proof here) that for general order m we get the expression,

$$
\mathcal{Y}_m = \mathrm{Tr}\,\frac{\Delta^m}{m}, \tag{C47}
$$

reproducing the expression given in Eq. (C39), which is what we intended to show. Using this expression, along with Eqs. (C38) and (C40), we obtain the exact result for the quantum average in Eq. (C32). We have

$$
\begin{aligned}
\left\langle e^{-i\epsilon\int_0^\tau dt\,\boldsymbol{\phi}^\dagger \mathcal{H}_n\boldsymbol{\phi}}\right\rangle_q &= \exp\left\{\sum_{m=1}^{\infty}\mathcal{Y}_m\zeta^m\right\} \\
&= \exp\{\mathrm{Tr}\log(\mathbb{1} - \zeta\Delta)^{-1}\} \\
&= \exp\{\log \mathrm{Det}(\mathbb{1} - \zeta\Delta)^{-1}\} \\
&= \frac{1}{\mathrm{Det}(\mathbb{1} - \zeta\Delta)}.
\end{aligned} \tag{C48}
$$

With this expression, all that remains is to evaluate the integral over z in Eq. (C31) in order to obtain the expression for $\mathcal{A}_s$.

Before doing so, let us first define the matrix δ

$$
\begin{aligned}
\delta &\equiv \mathbb{1} + \Delta \\
&= \mathcal{T}e^{-i\epsilon\int_0^\tau dt\,\widetilde{\mathcal{H}}_n(t)},
\end{aligned} \tag{C49}
$$

in terms of which the analysis to follow simplifies. Recalling that $\zeta \equiv z/(1-z)$, we see that

$$
\frac{1}{\mathrm{Det}(\mathbb{1} - \zeta\Delta)} = \frac{(1-z)^2}{\mathrm{Det}(\mathbb{1} - z\delta)}, \tag{C50}
$$

where we used the fact that for any 2×2 matrix A and c-number c we have $\mathrm{Det}(cA) = c^2\mathrm{Det}\,A$. This is a very convenient representation since the factor of $(1-z)^2$ in Eq. (C50) cancels a similar factor appearing in Eq. (C31) which simplifies the evaluation of the integral over z. It also turns out to be very convenient to reexpress the determinant

on the right-hand side of Eq. (C50) as

$$
\begin{aligned}
[\mathrm{Det}(\mathbb{1} - z\delta)]^{-1} &= [1 - z\,\mathrm{Tr}\,\delta + z^2\,\mathrm{Det}\,\delta]^{-1} \\
&= [1 - z\,\mathrm{Tr}\,\delta + z^2]^{-1},
\end{aligned}
\tag{C51}
$$

where the first line is an *exact* relation for *any* 2×2 matrix δ, and in the second line we used the fact that δ is a $SU(2)$ matrix [see Eq. (C49)] with unit determinant.

As a final step before evaluating the integral over z in Eq. (C31) to find $\mathcal{A}_s$, we take advantage of the fact that the right-hand side of Eq. (C51) takes the form of a *generating function* for the *Chebyshev polynomials of the second kind* [24], which play an important role in the development of spherical harmonics in four dimensions. We have

$$
\frac{1}{1 - z\,\mathrm{Tr}\,\delta + z^2} = \sum_{j=0}^{\infty} V_j\!\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right) z^j,
\tag{C52}
$$

where $V_j(x)$ denotes the jth order Chebyshev polynomial of the second kind. Equation (C52) is valid as long as the conditions $|z| < 1$ and $|\frac{1}{2}\,\mathrm{Tr}\,\delta| \leqslant 1$ apply, which is always the case for our problem. Using Eqs. (C50), (C51), and (C52) we find a very useful representation for Eq. (C48):

$$
\left\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\dagger \mathcal{H}_n \boldsymbol{\phi}} \right\rangle_q = (1 - z)^2 \sum_{j=0}^{\infty} V_j\!\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right) z^j.
\tag{C53}
$$

We are now ready to evaluate Eq. (C31) to find $\mathcal{A}_s$. Using Eq. (C53), we find

$$
\begin{aligned}
\mathcal{A}_s &= \oint \frac{dz}{2\pi i} \frac{\left\langle e^{-i\epsilon \int_0^\tau dt\, \boldsymbol{\phi}^\dagger \mathcal{H}_n \boldsymbol{\phi}} \right\rangle_q}{(2s+1)(1-z)^2 z^{2s+1}} \\
&= \frac{1}{2s+1} \sum_{j=0}^{\infty} \oint \frac{dz}{2\pi i} \frac{V_j\!\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)}{z^{2s+1-j}} \\
&= \frac{1}{2s+1} V_{2s}\!\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right),
\end{aligned}
\tag{C54}
$$

where in going from the second to the third line, we use Cauchy's residue theorem. In order to arrive at the expression for A_s shown in Eq. (26) in the main paper, we make use of the following relation:

$$
V_j(\cos\theta) = \frac{\sin\,(j+1)\theta}{\sin\theta}.
\tag{C55}
$$

We then have

$$
\begin{aligned}
\mathcal{A}_s &= \frac{1}{2s+1} \frac{\sin\left[(2s+1)\cos^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)\right]}{\sin\left[\cos^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)\right]} \\
&= \frac{1}{2s+1} \sum_{j=-s}^{s} e^{-2ji\cos^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)},
\end{aligned}
\tag{C56}
$$

where the expression given in the second line turns out to be more convenient for calculations. By carrying out the sum in the second line, one arrives at the expression given in the right-hand side of the first line, showing that the two expressions are entirely equivalent. Taking the simplest case, $s = 1/2$, we get the expression,

$$
\begin{aligned}
\mathcal{A}_{1/2} &= \tfrac{1}{2}\left[e^{i\cos^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)} + e^{-i\cos^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\,\delta\right)}\right] \\
&= \tfrac{1}{2}\,\mathrm{Tr}\,\delta \\
&= \tfrac{1}{2}\,\mathrm{Tr}\,\mathcal{T} e^{-i\epsilon \int_0^\tau dt\, \widetilde{\mathcal{H}}_n(t)},
\end{aligned}
\tag{C57}
$$

where in the third line we made use of Eq. (C49). Making use of Eq. (C57), we finally arrive at the expression given in Eq. (34) in the main paper:

$$
\mathcal{A}_s = \frac{1}{2s+1} \sum_{j=-s}^{s} e^{-2ji\cos^{-1}\mathcal{A}_{1/2}}.
\tag{C58}
$$

This relation is remarkable in the fact that it essentially shows us we can understand the behavior of the spin s system entirely in terms of quantities associated with the spin $1/2$ system. The dynamics of a spin s system is entirely contained in the dynamics of a spin-half system.

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Communication* (Cambridge University Press, Cambridge, 2000).

[2] D. d'Alessandro, *Introduction to Quantum Control and Dynamics* (CRC Press, Boca Raton, 2007).

[3] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).

[4] C. H. Bennett and D. P. DiVincenzo, Nature (London) **404**, 247 (2000).

[5] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).

[6] A. G. Kofman and G. Kurizki, Phys. Rev. Lett. **87**, 270405 (2001).

[7] A. G. Kofman and G. Kurizki, Phys. Rev. Lett. **93**, 130406 (2004).

[8] T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk, New J. Phys. **15**, 095004 (2013).

[9] C. Kabytayev, T. J. Green, K. Khodjasteh, M. J. Biercuk, L. Viola, and K. R. Brown, Phys. Rev. A **90**, 012316 (2014).

[10] J. Schwinger, J. Math. Phys. **2**, 407 (1960).

[11] L. V. Keldysh, Sov. Phys. JETP **20**, 1018 (1965).

[12] A. Kamenev, *Field Theory of Non-equilibrium Systems* (Cambridge University Press, Cambridge, 2011).

[13] A. Polkovnikov, Ann. Phys. **325**, 1790 (2010).

[14] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).

[15] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer, New York, 2003).

[16] J.-P. Bouchaud and M. Mezard, J. Phys. A: Math. Gen. **30**, 7997 (1997).

[17] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Rev. Mod. Phys. **23**, 863 (2012).

[18] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, Rev. Mod. Phys. **86**, 361 (2014).

[19] U. Weiss, *Quantum Dissipative Systems* (World Scientific, Singapore, 1999).

[20] J. Klauder and B. Skagerstam, *Coherent States: Applications in Physics and Mathematical Physics* (World Scientific, Singapore, 1985).

[21] S. Blanes, F. Casas, J. A. Oteo, and J. Ros, Phys. Rep. **470**, 151 (2008).

[22] Strictly speaking, one should use the *lab frame* quantity $|\boldsymbol{\omega}(t)|^2$, but in our system due to rotational invariance $|\boldsymbol{\Omega}(t)|^2 = |\boldsymbol{\omega}(t)|^2$ *for any and all* realizations of the control field.

[23] J. Bergli, Y. M. Galperin, and B. L. Altshuler, New J. Phys. **11**, 025002 (2009).

[24] E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis*, 4th ed. (Cambridge University Press, Cambridge, 1990).