Rigorous bounds for optimal dynamical decoupling

Götz S. Uhrig[1,*] and Daniel A. Lidar[2,†]
[1]*Lehrstuhl für Theoretische Physik I, Technische Universität Dortmund, Otto-Hahn Straße 4, D-44221 Dortmund, Germany*
[2]*Departments of Chemistry, Electrical Engineering, and Physics, Center for Quantum Information & Technology, University of Southern California, Los Angeles, California 90089, USA*

(Received 7 February 2010; published 2 July 2010)

We present rigorous performance bounds for the optimal dynamical decoupling pulse sequence protecting a quantum bit (qubit) against pure dephasing. Our bounds apply under the assumption of instantaneous pulses and of bounded perturbing environment and qubit-environment Hamiltonians such as those realized by baths of nuclear spins in quantum dots. We show that if the total sequence time is fixed the optimal sequence can be used to make the distance between the protected and unperturbed qubit states arbitrarily small in the number of applied pulses. If, on the other hand, the minimum pulse interval is fixed and the total sequence time is allowed to scale with the number of pulses, then longer sequences need not always be advantageous. The rigorous bound may serve as a testbed for approximate treatments of optimal decoupling in bounded models of decoherence.

DOI: 10.1103/PhysRevA.82.012301 PACS number(s): 03.67.Pp, 82.56.Jn, 76.60.Lz, 03.65.Yz

I. INTRODUCTION

Quantum systems tend to rapidly decohere due to the coupling to their environments, a process which is especially detrimental to quantum information processing and high-resolution spectroscopy [1]. Of the many methods which have been proposed in recent years to overcome the damage caused by decoherence, we focus here on dynamical decoupling (DD), a method for suppressing decoherence whose origins can be traced to the Hahn spin echo [2]. In DD, one applies a series of strong and frequent pulses to a system, designed to decouple it from its environment [3–6]. Recently, it was discovered how to optimally suppress decoherence of a single qubit using DD under the idealization of instantaneous pulses [7–9]. One of us found an optimal pulse sequence (later dubbed UDD) for suppressing pure dephasing (single-axis decoherence) of a qubit coupled to a boson bath with a hard frequency cutoff [7]. In UDD one applies a series of N instantaneous π pulses at instants t_j $(j \in \{1,2,\ldots,N\})$, with the instants given by $t_j = T\delta_j$ where T is the total time of the sequence and

$$\delta_j = \sin^2[j\pi/(2N+2)]. \tag{1}$$

By optimal it is meant that with each additional pulse the sequence suppresses dephasing in one additional order in an expansion in T [i.e., N pulses reduce dephasing to $O(T^{N+1})$]. The existence and convergence of an expansion in powers of T, at least as an asymptotic expansion, is a necessary assumption [9,10].

The UDD sequence was first conjectured [11,12] and then proven [10] to be universal in the sense that it applies to any bath causing pure dephasing of a qubit, not just bosonic baths. The performance of the UDD sequence was tested, and its advantages over other pulse sequences confirmed under appropriate circumstances, in a series of recent experiments [13–15]. Its limitations as a function of sharpness of the bath spectral-density high-frequency cutoff [16] and as a function of timing constraints [17] have also been discussed.

To suppress general (three-axis) decoherence on a qubit to all orders concatenated sequences can be used [18,19] whose efficiency can be improved by using UDD building blocks [20]. A near optimum suppression is achieved by quadratic dynamic decoupling (dubbed QDD). This scheme was proposed and numerically tested in Ref. [8] and analytically corroborated in Ref. [9]. In QDD, a sequence of $(N+1)^2$ pulses comprising two nested UDD sequences suppresses general qubit decoherence to $O(T^{N+1})$, which is known from brute-force symbolic algebra solutions for small N to be near optimal [8].

While rigorous performance bounds have been derived previously for periodic and concatenated DD pulse sequences [21–23], no such performance bounds have yet been derived for optimal decoupling pulse sequences, in particular UDD and QDD. In this work we focus on UDD and obtain rigorous performance bounds. We postpone the problem of finding rigorous QDD performance bounds to a future publication. Our main result here is an analytical upper bound for the distance between UDD-protected states subject to pure dephasing and unperturbed states as a function of the natural dimensionless parameters of the problem, namely the total evolution time T measured in units of the maximal intrabath energy J_0, and in units of the system-bath coupling strength J_z. The bound shows that this distance (technically, the trace-norm distance) can be made arbitrarily small as a function of the number of pulses N, as $(1/N!)(J_0T + J_zT)^N$. This presumes that the bounds J_α $(\alpha \in \{0,z\})$ are finite. Such a situation is realized in spin baths (i.e., in experimental setups where the decoherence is caused by coupling to a large number of spins). For instance, the electronic spin in a quantum dot couples via the hyperfine interaction with the nuclear spins, which is an experimentally relevant system [24]. The fact that the spin operators $\vec{S}$ are bounded implies that J_α $(\alpha \in \{0,z\})$ are finite for a finite number of spins.

The bounds J_α $(\alpha \in \{0,z\})$ can be infinite for unbounded baths, such as oscillator baths. In such cases, which includes the ubiquitous spin-boson model, our analysis does not apply. Alternative approaches, such as those based on correlation function bounds [23], are then required.

We begin by introducing the model of pure dephasing in the presence of instantaneous DD pulses in Sec. II. We then

*goetz.uhrig@tu-dortmund.de
†lidar@usc.edu

derive a general time-evolution bound in Sec. III, without reference to any particular pulse sequence. In Sec. IV we specialize this bound to the UDD sequence. Then, in Sec. V, we obtain our main result: an upper bound on the trace-norm distance between the UDD-protected and unperturbed states. In Sec. VI we analyze the implications of this bound in the more realistic setting when only a certain minimal interval between consecutive pulses can be attained. Certain technical details are presented in the Appendix, including the first complete universality proof of the UDD sequence, which does not rely on the interaction picture.

II. MODEL

We start from the general, uncontrolled, time-independent system-bath Hamiltonian for pure dephasing

$$H_{\rm unc} = I_{\rm S} \otimes B_0 + \sigma_z \otimes B_z, \tag{2}$$

where B_0 and B_z are bounded but otherwise arbitrary operators acting on the bath Hilbert space $\mathcal{H}_{\rm B}$, $I_{\rm S}$ is the identity operator on the system Hilbert space $\mathcal{H}_{\rm S}$, and σ_z is the diagonal Pauli matrix. The bath operator B_z need not be traceless (i.e., we allow for the possibility of a pure-system term $\sigma_z \otimes I_{\rm B}$ in the system-bath interaction term $H_{\rm SB} := \sigma_z \otimes B_z$). Such internal system dynamics will be removed by the DD pulse sequence we shall add next, along with the coupling to the bath. However, the assumption of pure dephasing means that we assume that the level splitting of the system, that is, any term proportional to $\sigma_\perp$ [with $\sigma_\perp$ being $\cos(\varphi)\sigma_x + \sin(\varphi)\sigma_y$ for arbitrary φ] acting on the system is fully controllable. Otherwise the model is one of general decoherence, and our methods require a modification along the lines of Refs. [8] and [9]. If the system described by Eq. (2) is subject to N instantaneous π pulses at the instants $\{t_j := T\delta_j\}_{j=1}^N$ about a spin axis perpendicular to the z axis, that is, if the Hamiltonian $H_{\rm DD}(t) = \frac{\pi}{2}\sum_{j=1}^N \delta(t-t_j)\sigma_\perp \otimes I_{\rm B}$ is added to $H_{\rm unc}$, the interaction picture ("toggling-frame") Hamiltonian $H_{\rm tog}(t) = U_{\rm DD}^\dagger(t) H_{\rm unc} U_{\rm DD}(t)$ reads

$$H_{\rm tog}(t) = I_{\rm S} \otimes B_0 + f(t)\sigma_z \otimes B_z, \tag{3}$$

where the unitary $U_{\rm DD}(t)$ alternates between $I_{\rm S} \otimes I_{\rm B}$ and $\sigma_\perp \otimes I_{\rm B}$ at the instants $\{t_j\}_{j=1}^N$, and consequently the "switching function" $f(t) = \pm 1$ changes sign at the same instants.

We shall also need the magnitudes of the two parts of the Hamiltonian

$$J_0 := \|B_0\| < \infty, \quad J_z := \|B_z\| < \infty, \tag{4}$$

where $\|\cdot\|$ is the supoperator norm (see Appendix A). There are certainly situations where either J_0 or J_z can be divergent (e.g., J_0 in the case of oscillator baths).

III. TIME-EVOLUTION BOUNDS

We aim to bound certain parts of the time-evolution operator induced by $H_{\rm tog}(t)$

$$U(T) = \mathcal{T} \exp\left[-i\int_0^T H_{\rm tog}(t)\,dt\right], \tag{5}$$

where $\mathcal{T}$ is the time-ordering operator.

Standard time-dependent perturbation theory provides the following Dyson series for $U(T)$

$$U(T) = \sum_{n=0}^{\infty} (-iT)^n \sum_{\{\vec{\alpha};\dim(\vec{\alpha})=n\}} F_{\vec{\alpha}}\, \widehat{Q}_{\vec{\alpha}}, \tag{6a}$$

$$F_{\vec{\alpha}} := \int_0^1 ds_n f_{\alpha_n}(s_n) \int_0^{s_n} ds_{n-1} f_{\alpha_n}(s_n) \ldots \int_0^{s_3} ds_2 f_{\alpha_2}(s_2) \int_0^{s_2} ds_1 f_{\alpha_1}(s_1), \tag{6b}$$

$$\widehat{Q}_{\vec{\alpha}} := \sigma_{\alpha_n} B_{\alpha_n} \ldots \sigma_{\alpha_2} B_{\alpha_2} \sigma_{\alpha_1} B_{\alpha_1}, \tag{6c}$$

where $\dim(\vec{\alpha}) = n$ is the dimension of the vector $\vec{\alpha}$. The identity $I_{\rm S}$ in the Hilbert space of the qubit/spin is denoted by σ_0. In all sums over the vectors $\vec{\alpha}$ their components α_j take the values 0 or z. In this way, the summation includes all possible sequences of B_0 and B_z. The function $f_0(s)$ is constant and equal to 1 while $f_z(s) := f(sT)$ takes the values ± 1. We use the dimensionless relative time $s := t/T$ so that all dependence on T appears as a power in the prefactor. Note that the coefficients $F_{\vec{\alpha}}$ do not depend on T.

To find an upper bound on *each* term $F_{\vec{\alpha}}\, \widehat{Q}_{\vec{\alpha}}$ separately we proceed in two steps. First, we use $|f_\alpha| = 1$ to obtain

$$|F_{\vec{\alpha}}| \leqslant \int_0^1 ds_n \int_0^{s_n} ds_{n-1} \ldots \int_0^{s_3} ds_2 \int_0^{s_2} ds_1 = \frac{1}{n!}. \tag{7}$$

Second, we use Eq. (4) and $\|\sigma_\alpha\| = 1$ to arrive at

$$\|\widehat{Q}_{\vec{\alpha}}\| \leqslant \prod_{j=1}^n J_{\alpha_j} = J_0^{n-k(\vec{\alpha})} J_z^{k(\vec{\alpha})}, \tag{8}$$

where we used the submultiplicativity of the supoperator norm (see Appendix A). The number $k(\vec{\alpha})$ stands for the number of times that the factor J_z occurs. Standard combinatorics of binomial coefficients tells us that the term $J_0^{n-k} J_z^k$ occurs $n!/[k!(n-k)!]$ times in the sum over all the vectors $\vec{\alpha}$ of given dimensionality n in (6a). Hence each term of the time expansion of $U(T)$ is bounded by

$$\left\| \sum_{\{\vec{\alpha};\dim(\vec{\alpha})=n\}} F_{\vec{\alpha}}\, \widehat{Q}_{\vec{\alpha}} \right\| \leqslant \sum_{k=0}^n \frac{1}{k!(n-k)!} J_0^{n-k} J_z^k. \tag{9}$$

We therefore define the bounding series

$$S(J_0, J_z) := \sum_{n=0}^{\infty} T^n \sum_{k=0}^n \frac{1}{k!(n-k)!} J_0^{n-k} J_z^k \tag{10a}$$

$$= \exp[(J_0 + J_z)T]. \tag{10b}$$

It then follows from Eq. (9) that each multinomial in J_0 and J_z of the expansion of $S(J_0, J_z)$ is an upper bound on the norm of the sum of the corresponding multinomial in the operators B_0 and B_z of the expansion of $U(T)$ in Eq. (6a). This is the property which we will use in the sequel.

IV. BOUNDS FOR DEPHASING

From $\sigma_z^2 = I_S$ it is obvious that only the odd powers in B_z contribute to dephasing while the even ones do not. Hence we split $U(T)$ as

$$U(T) = I_S \otimes B_+(T) + \sigma_z \otimes B_-(T), \tag{11}$$

where the operators $B_\pm$ act only on the bath while I_S and σ_z act only on the qubit. The operator B_+ comprises all the terms with an even number k of $\sigma_z \otimes B_z$ [i.e., with an even number of J_z in the bounding series $S(J_0,J_z)$]. The operator B_- comprises all the terms with odd number k of $\sigma_z \otimes B_z$ [i.e., with an odd number of J_z in the bounding series $S(J_0,J_z)$]. Hence to bound the time series of $B_-(T)$ term by term we need the time series of the odd part of $S(J_0,J_z)$ in J_z. This, from (10b) is

$$S_-(J_0,J_z) = \exp(J_0T)\sinh(J_zT). \tag{12}$$

The time series of $S_-(J_0,J_z)$ provides a bounding series of $B_-(T)$ term by term. Hence we define

$$d_k := \frac{1}{k!}\frac{\partial^k}{\partial T^k} S_-(J_0,J_z)\bigg|_{T=0}, \tag{13}$$

such that $S_-(J_0,J_z) = \sum_{k=0}^\infty d_k T^k$.

We know from the proof of Yang and Liu [10] that in the B_0-interaction picture a UDD sequence with N pulses [which we denote by UDD(N)] should make the first N powers in T of $B_-(T)$ vanish [i.e., $B_-(T) = O(T^{N+1})$]. However, since the Yang-Liu proof does not directly apply to our discussion, we provide a complete version of this proof which avoids the B_0-interaction picture in Appendix B. The remaining powers are bounded by the corresponding coefficients d_k of S_-. Thus the expression

$$\Delta_N := \sum_{k=N+1}^\infty d_k T^k, \tag{14}$$

provides an upper bound for $B_-(T)$ if UDD(N) is applied

$$\|B_-(T)\| \leqslant \Delta_N. \tag{15}$$

Due to the obvious analyticity in the variable T of $S_-(J_0,J_z)$ as defined in (12) we know that the residual term vanishes for $N \to \infty$, that is,

$$\lim_{N\to\infty} \Delta_N = 0. \tag{16}$$

This statement holds true irrespective of the values of J_0 and J_z as long as they are finite.

We can obtain a more explicit expression for Δ_N. Besides the dimensionless number of pulses N the bound Δ_N depends on J_0T and on J_zT. It is convenient to introduce the dimensionless parameters

$$\varepsilon := J_0T, \quad \eta := J_z/J_0, \tag{17}$$

instead. In terms of these parameters we have

$$S_-(\eta,\varepsilon) = \exp(\varepsilon)\sinh(\varepsilon\eta). \tag{18}$$

From the series

$$\exp(\varepsilon)\sinh(\varepsilon\eta) = \frac{1}{2}[e^{\varepsilon(1+\eta)} - e^{\varepsilon(1-\eta)}] \tag{19a}$$

$$= \sum_{l=0}^\infty \frac{\varepsilon^l}{2l!}[(1+\eta)^l - (1-\eta)^l] \tag{19b}$$

$$= \sum_{l=0}^\infty p_l(\eta)\varepsilon^l, \tag{19c}$$

with

$$p_l(\eta) := \frac{1}{2l!}[(1+\eta)^l - (1-\eta)^l], \tag{20}$$

we obtain

$$\Delta_N(\eta,\varepsilon) = \sum_{n=N+1}^\infty p_n(\eta)\varepsilon^n \tag{21a}$$

$$= p_{N+1}(\eta)\varepsilon^{N+1} + O(\varepsilon^{N+2}). \tag{21b}$$

This, together with the bound (15), is our key result: It captures how the "error" $\|B_-(T)\|$ is suppressed as a function of the relevant dimensionless parameters of the problem, η, ε, and N. Note that convergence for $N \to \infty$ is always ensured by the factorial in the denominator, irrespective of the values of ε and η as long as these are finite.

For practical purposes it is advantageous not to compute Δ_N by the infinite series in (21a), but by

$$\Delta_N(\eta,\varepsilon) = S_-(\eta,\varepsilon) - \sum_{n=0}^N p_n(\eta)\varepsilon^n, \tag{22}$$

which can easily be computed by computer algebra programs. Figures 1 and 2 depict the results of this computation. Consider first Fig. 1. Each curve shows $\Delta_N(\eta,\varepsilon)$ as a function of ε, at fixed η and N. The error $\|B_-(T)\|$ always lies under the corresponding curve. Clearly, the bound becomes tighter as ε decreases. Moreover, the more pulses are applied (the different

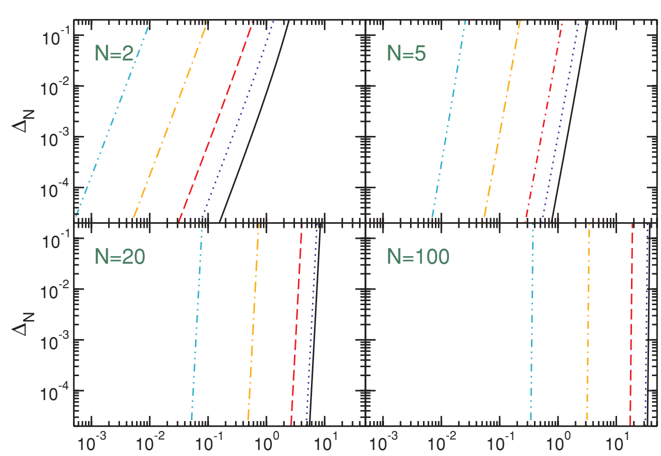

FIG. 1. (Color online) The bounding function Δ_N as a function of $\varepsilon = J_0T$, as given in Eq. (14) for various numbers of pulses N and various values of the parameter $\eta = J_z/J_0 \in \{0.01, 0.1, 1, 10, 100\}$, with η increasing from the rightmost curve to the leftmost curve in each panel.

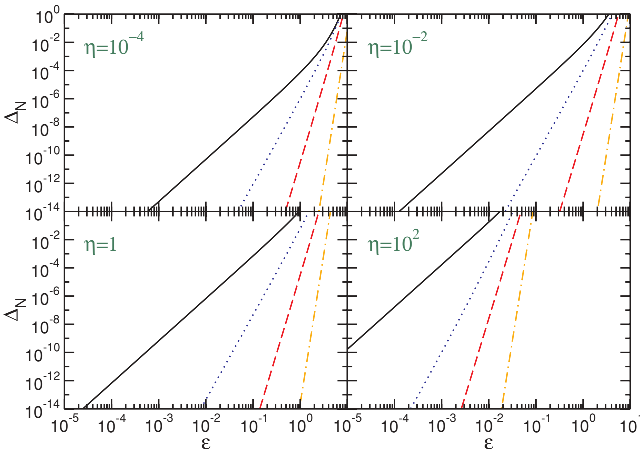

FIG. 2. (Color online) The bounding function Δ_N as a function of $\varepsilon = J_0 T$, as given in Eq. (21a) for various numbers of pulses $N \in \{2,5,10,20\}$, at fixed values of η. In each panel the curves become steeper as N increases.

panels) the higher the power in ε and thus the steeper the curve. Additionally, the curves are shifted to the right as N increases. Clearly then, a larger number of pulses improves the error bound significantly at fixed ε and η. This effect is even more conspicuous in Fig. 2 where η is fixed in each of the two panels, and the different curves correspond to different values of N. The vertical line intersects the bounding function at progressively lower points as N is increased, showing how the bound becomes tighter.

V. DISTANCE BOUND

Intuitively, we expect the bound on $\|B_-(T)\|$ derived in the previous section to be sufficient to bound the effect of dephasing. However, to make this rigorous we need a bound on the trace-norm distance $D[\rho_S(T),\rho_S^0(T)]$ between the "actual qubit" state

$$\rho_S(T) := \mathrm{tr}_B[\rho_{SB}(T)], \tag{23}$$

and the "ideal qubit" state

$$\rho_S^0(T) := \mathrm{tr}_B\big[\rho_{SB}^0(T)\big], \tag{24}$$

where $\rho_{SB}^0(T)$ is the time-evolved state without coupling between qubit and bath. The partial trace over the bath degrees of freedom is a map from the joint system-bath Hilbert space to the system-only Hilbert space (see Appendix A), and is denoted by tr_B. As we shall see, the term $I_S \otimes B_+(T)$ in Eq. (11) indeed has a small, and in fact, essentially negligible effect.

To obtain the desired distance bound we consider a factorized initial state $\rho_{SB}^0(0) = |\psi\rangle\langle\psi| \otimes \rho_B$, which evolves to $\rho_{SB}(T) = U(T)\rho_{SB}^0(0)U^\dagger(T)$ when the system-bath interaction is on (the "actual" state), or to $\rho_{SB}^0(T) = I_S \otimes U_B(T)\rho_{SB}^0(0) I_S \otimes U_B^\dagger(T)$ when the interaction is off (the "ideal" state). The unitary bath time-evolution operator without coupling reads

$$U_B(T) := \exp(-iTB_0), \tag{25}$$

where B_0 is the pure-bath term in Eq. (2). The initial bath state ρ_B is arbitrary (e.g., a mixed thermal equilibrium state), while the initial system state is pure. Let us define the correlation functions

$$b_{\alpha\beta}(T) := \mathrm{tr}[B_\alpha(T)\rho_B B_\beta^\dagger(T)], \tag{26}$$

where $\alpha,\beta \in \{+,-\}$, and where all operators under the trace act only on the bath Hilbert space. Explicit computation (see Appendix C) then yields

$$D\big[\rho_S(T),\rho_S^0(T)\big] \leqslant \tfrac{1}{2}[|b_{++}(T)-1| + |b_{+-}(T)| + |b_{-+}(T)| + |b_{--}(T)|]. \tag{27}$$

We will show that b_{++} is very close to 1 while the other $b_{\alpha\beta}$ quantities are small in the sense that they are bounded by Eq. (15).

First note from the unitarity of Eq. (11) that

$$\begin{aligned} I &= U^\dagger U \\ &= I_S \otimes (B_+^\dagger B_+ + B_-^\dagger B_-) + \sigma_z \otimes (B_-^\dagger B_+ + B_+^\dagger B_-), \end{aligned} \tag{28}$$

where we omitted the time dependence T to lighten the notation. Hence we have

$$I = B_+^\dagger B_+ + B_-^\dagger B_-, \tag{29a}$$

$$0 = B_+^\dagger B_- + B_-^\dagger B_+. \tag{29b}$$

It follows that $\langle i|B_+^\dagger B_+|i\rangle = \|B_+|i\rangle\|^2 \leqslant 1$ for all normalized states $|i\rangle$ because $\langle i|B_- B_-^\dagger|i\rangle = \|B_-|i\rangle\|^2$ is nonnegative. Thus in particular $\max_{\||i\rangle\|=1}\|B_+|i\rangle\| \leqslant 1$, and we can conclude that

$$\|B_+\| \leqslant 1. \tag{30}$$

Cyclic invariance of the trace in $b_{\alpha\beta}$ together with Eq. (29a) and the normalization $\mathrm{tr}[\rho_B] = 1$ immediately yields $b_{++} + b_{--} = 1$, while the combination with Eq. (29b) implies $b_{+-} + b_{-+} = 0$. Hence Eq. (27) can be simplified to

$$D\big[\rho_S(T),\rho_S^0(T)\big] \leqslant |b_{+-}(T)| + |b_{--}(T)|. \tag{31}$$

To obtain a bound on the correlation functions $b_{\alpha\beta}$ we use the following general correlation function inequality (for a proof see Appendix D):

$$|\mathrm{tr}[Q\rho_B Q']| \leqslant \|Q'\|\|Q\|, \tag{32}$$

which holds for arbitrary bounded bath operators Q,Q'. Applying Eq. (32) to Eq. (26) yields

$$|b_{--}(T)| \leqslant \|B_-(T)\|^2, \tag{33a}$$

$$|b_{+-}(T)| \leqslant \|B_+(T)\|\|B_-(T)\| \tag{33b}$$

$$\leqslant \|B_-(T)\|, \tag{33c}$$

where in the last inequality we used Eq. (30).

Summarizing, together with Eqs. (15) and (31) we have obtained the following rigorous upper bound for the trace-norm distance

$$D\big[\rho_S(T),\rho_S^0(T)\big] \leqslant \min\big[1,\Delta_N(\eta,\varepsilon) + \Delta_N^2(\eta,\varepsilon)\big]. \tag{34}$$

This upper bound completes our main result. Since as we saw in Eq. (21b) $\Delta_N(\eta,\varepsilon) = p_{N+1}(\eta)\varepsilon^{N+1} + O(\varepsilon^{N+2})$, the appearance of the squared term in Eq. (34) [whose origin is

$|b_{--}(T)|]$ is not relevant in the sense that even in the presence of this term the bound

$$D\big[\rho_S(T),\rho_S^0(T)\big] \leqslant p_{N+1}(\eta)\varepsilon^{N+1} + O(\varepsilon^{N+2}), \tag{35}$$

holds. Hence the result of Eq. (21a) depicted in Figs. 1 and 2 provides the desired result. Ignoring the Δ_N^2 term in Eq. (34), we note that Figs. 1 and 2 also reveal the limitations of our bound when ε or η are too large for a given value of N: For any pair of states it is always the case that $D \leqslant 1$, so that as soon as $\Delta_N = 1$ the bound no longer provides any useful information.

Note further that the results shown in Fig. 1 are qualitatively similar to the results obtained for the analytically solvable spin-boson model for pure dephasing [7]. Heuristically, the necessary identification is $J_0 = \omega_D$ where ω_D is the hard cutoff of the spectral function and $\eta \propto \alpha$ where α is the dimensionless coupling constant for Ohmic noise. We stress that the advantage of Eq. (34) compared to the analytically exact results in Ref. [7] is that it holds rigorously for a large class of pure dephasing models, namely those of bounded Hamiltonians which comprise the experimentally relevant class of spin baths.

VI. ANALYSIS FOR FINITE MINIMUM PULSE INTERVAL

So far we have essentially treated the total time T and the number of pulses N as independent parameters. This is possible when there is no lower limit on the pulse intervals. However, in reality this is never the case and in this section we analyze what happens when there is such a lower limit. Note that it follows from Eq. (1) that the smallest pulse interval is the first: $t_1 = T\sin^2[\pi/(2N+2)]$. Let us assume that t_1 is fixed, so that, given t_1 and N, the total time is

$$T(N) = t_1 q(N), \tag{36a}$$

$$q(N) := \csc^2\left(\frac{\pi}{2N+2}\right). \tag{36b}$$

For large N we can expand the $\csc^2$ function to first order in its small argument, yielding

$$q(N) = \left(\frac{2N+2}{\pi}\right)^2 + \frac{1}{3} + O(N^{-2}), \tag{37}$$

which shows how the total time grows as a function of N at fixed minimum pulse interval t_1. Along with η, the relevant dimensionless parameter is now

$$\varepsilon_1 := J_0 t_1, \tag{38}$$

instead of $\varepsilon = q(N)\varepsilon_1$. We can then rewrite the bounding function (21a) in terms of these quantities as

$$\Delta_N(\eta,\varepsilon_1) = \sum_{n=N+1}^{\infty} p_n(\eta) q^n(n)\varepsilon_1^n. \tag{39}$$

Considering now the large N limit of the first term in this sum, we have

$$p_{N+1}(\eta)q^{N+1}(N+1)\varepsilon_1^{N+1} \approx \frac{1}{2N!}\left(\frac{2}{\pi}N\right)^{2N}[(1+\eta)\varepsilon_1]^N \approx (cN)^N, \tag{40}$$

where we kept only the leading order terms and neglected all additive constants relative to N, and in Eq. (40) used Stirling's approximation $n! \approx (n/e)^n$. The constant c is $\frac{1}{2}(\frac{2}{\pi})^2 e(1+\eta)\varepsilon_1$. We thus see clearly that for fixed t_1 it becomes counterproductive to make N too large since no matter how small c is, for large enough N the factor N^N will eventually dominate. This reflects the competition between the gains due to higher-order pulse sequences and the losses due to the increased coupling time to the qubit allotted to the environment. Similar conclusions, delineating regimes where increasingly long DD sequences become disadvantageous, have been reported for periodic [19,21] and concatenated [19,23,25] DD pulse sequences, as well as for the QDD sequence [8].

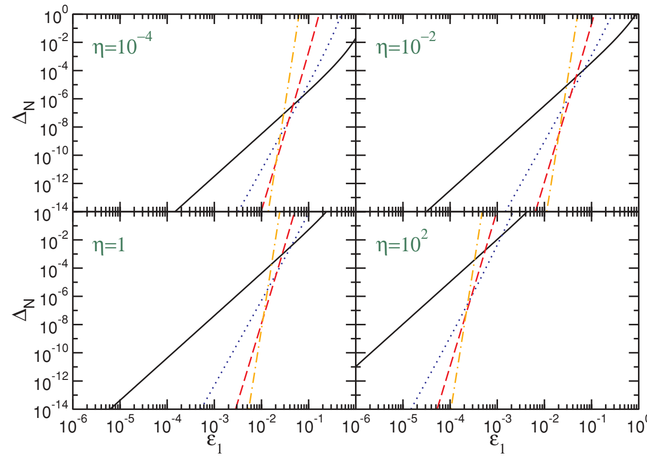

FIG. 3. (Color online) The bounding function Δ_N as a function of $\varepsilon_1 = t_1 J_0$ (where t_1 is the smallest pulse interval), at fixed values of η. The number of pulses $N \in \{2,5,10,20\}$ is varied from curve to curve. In each panel the curves become steeper as N increases.

These conclusions are further illustrated in Fig. 3, where we plot the bound $\Delta_N(\eta,\varepsilon_1)$ by replacing ε with $\varepsilon_1 q(N)$ in Eq. (21a). This figure should be contrasted with Fig. 2. The most notable change is that increasing N now no longer uniformly improves performance. Whereas in Fig. 2 the curves for different values of N all tend to converge at high values of ε, in Fig. 3 a high N value results in a steeper slope, but also moves the curve to the left. Thus, for a fixed value of ε_1 it can be advantageous to use a small value of N (e.g., for $\eta = 0.01$ and $\varepsilon_1 = 0.1$ the $N = 2$ curve provides the tightest bound).

VII. CONCLUSION

We have derived rigorous performance bounds for the UDD sequence protecting a qubit against pure dephasing. The derivation is based on the existence of finite bounds for the relevant parts of the Hamiltonian, captured in the dimensionless parameters ε and η. Under this assumption the bounds show rigorously that dephasing is suppressed to leading order as $(1/N!)[\varepsilon(1+\eta)]^N$. We consider it a vital step to know that irrespective of any details of the bath, except for the existence of finite bounds, a large number N of pulses is always advantageous at fixed T—at least under the idealized assumption of perfect and instantaneous pulses. The requirement of finite bounds is fulfilled for the experimentally relevant class of spin baths such as the nuclear spins in quantum dots [24].

An immediate corollary of our results is that identical bounds apply for the case of the UDD sequence protecting a qubit against longitudinal relaxation. This is the case when the uncontrolled Hamiltonian (2) is replaced by $H_{\rm unc} = I_{\rm S} \otimes H_{\rm B} + \sigma_\perp \otimes B$, and the UDD pulse sequence consists of rotations about the spin-z axis. A practical implication is that the bounds found here can be used to check numerical and approximate calculations. Such calculations must obey our mathematically rigorous bounds so that a testbed is provided.

Furthermore, a number of interesting generalizations and extensions of our results readily suggest themselves. One is to consider rigorous bounds for finite pulse-width UDD sequences. It is already known how to construct such sequences with pulse-width errors which appear only to third order in the value of the pulse width [26], but no rigorous bounds have been found. Another important generalization, as mentioned previously, is to the QDD sequence for general decoherence [8,9]. We expect that techniques similar to the ones we introduced here will apply to both of these open problems. Yet another direction, which will require different techniques, is to find rigorous UDD performance bounds for unbounded baths, such as oscillator baths. It is likely that a correlation function analysis similar to that performed in Ref. [23] for periodic and concatenated DD sequences will prove useful in this case.

ACKNOWLEDGMENTS

G.S.U. is supported under DFG Grant No. UH 90/5-1. D.A.L. is supported under NSF Grant Nos. CHE-924318 and PHY-802678, and by the United States Department of Defense.

APPENDIX A: NORMS AND DISTANCES

A. Trace and partial trace

We deal only with linear trace-class bounded operators that map between separable Hilbert spaces in this work. A Hilbert space $\mathcal{H}$ is separable if and only if it admits a countable orthonormal basis. A bounded linear operator $A : \mathcal{H} \mapsto \mathcal{H}$, where $\mathcal{H}$ is separable, is said to be in the trace class if for some (and hence all) orthonormal bases $\{|k\rangle\}_k$ of $\mathcal{H}$ the sum of positive terms $\sum_k \langle k|\sqrt{A^\dagger A}|k\rangle$ is finite. In this case, the sum $\sum_k \langle k|A|k\rangle$ is absolutely convergent and is independent of the choice of the orthonormal basis. This value is called the trace of A, denoted by tr(A). Whenever we use the symbol tr in this work we mean the trace over the full Hilbert space the operator the trace is taken over is acting on.

Now consider two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ and let $A : \mathcal{H} \mapsto \mathcal{H}$ denote a linear trace-class bounded operator acting on the tensor product Hilbert space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $\{|k_i\rangle\}_k$ denote an orthonormal basis for $\mathcal{H}_i$, where $i \in 1,2$. The partial trace operation over the first (second) Hilbert space is a map from $\mathcal{H}$ to the second (first) Hilbert space, and has the operational definition $\mathrm{tr}_i(A) := \sum_{k_i} \langle k_i|A|k_i\rangle$. When A is decomposed in terms of the two orthonormal bases as $A = \sum_{k_1,k_1',l_2,l_2'} \langle k_1 l_2|A|k_1' l_2'\rangle |k_1 l_2\rangle\langle k_1' l_2'|$, where $|k_1 l_2\rangle := |k_1\rangle \otimes |l_2\rangle$, and so on, the partial trace over $\mathcal{H}_2$ can be written as $\mathrm{tr}_2(A) = \sum_{k_1,k_1',l_2} \langle k_1 l_2|A|k_1' l_2\rangle |k_1\rangle\langle k_1'|$. This makes it clear that $\mathrm{tr}_2(A)$ is an operator that acts on $\mathcal{H}_1$.

B. Supoperator norm and trace norm

We make frequent use of two matrix norms [27] in this work. The first is the supoperator norm

$$\|A\|_\infty := \sup_{\||v\rangle\|=1} \|A|v\rangle\| = \sup_{|v\rangle} \sqrt{\langle v|A^\dagger A|v\rangle}/\sqrt{\langle v|v\rangle}. \quad \text{(A1)}$$

The supoperator norm of A is the largest eigenvalue of $|A| := \sqrt{A^\dagger A}$ (i.e., the largest singular value of A). Since we use it often we denote $\|A\|_\infty$ for simplicity by $\|A\|$, and context should make it clear whether we are taking the norm of an operator or simply the Euclidean norm $\||v\rangle\| := \sqrt{\langle v|v\rangle}$ of a vector $|v\rangle$. Note that if A is normal ($A^\dagger A = AA^\dagger$, it can be unitarily diagonalized, so that $A^\dagger A = VD^\dagger DV^\dagger$ where V is unitary and D is the diagonal matrix of eigenvalues of A), the largest singular value coincides with the largest absolute value of the eigenvalues of A (i.e., $\|A\| = \sup_{\||v\rangle\|=1} |\langle v|A|v\rangle|$).

The trace norm

$$\|A\|_1 := \mathrm{tr}\sqrt{A^\dagger A}, \quad \text{(A2)}$$

is the sum of the eigenvalues of $|A|$ (i.e., the sum of the singular values of A). Therefore $\|A\| \leqslant \|A\|_1$. Both norms are unitarily invariant ($\|VAW\|_{\rm ui} = \|A\|_{\rm ui}$ for any pair of unitaries V and W) and therefore submultiplicative ($\|AB\|_{\rm ui} \leqslant \|A\|_{\rm ui}\|B\|_{\rm ui}$) [27]. In this work we make frequent use of both properties. In addition unitarily invariant norms are invariant under Hermitian conjugation (i.e., $\|A\|_{\rm ui} = \|A^\dagger\|_{\rm ui}$). This follows from the singular value decomposition: $A = V\Sigma W^\dagger$, where V and W are unitaries and Σ is the diagonal matrix of singular values of A. Since the singular values are all positive we have $A^\dagger = W\Sigma V^\dagger$ and hence $\|A^\dagger\|_{\rm ui} = \|\Sigma\|_{\rm ui} = \|A\|_{\rm ui}$.

C. Trace-norm distance and fidelity

The trace-norm distance between two mixed states described by the density operators ρ_1 and ρ_2 is defined as

$$D[\rho_1,\rho_2] := \tfrac{1}{2}\|\rho_1 - \rho_2\|_1. \quad \text{(A3)}$$

It is bounded between 0 and 1, vanishes if and only if $\rho_1 = \rho_2$, and is 1 if and only if ρ_1 are ρ_2 are orthogonal [i.e., $\mathrm{tr}(\rho_1\rho_2) = 0$].

The trace-norm distance is a standard and useful measure of distinguishability between states. The reason is this: Assume that we perform a generalized measurement (POVM—positive operator valued measurement) E with corresponding measurement operators $\{E_i\}$ satisfying the normalization condition $\sum_i E_i = I$. The measurement outcomes are described by the the measurement probabilities $p_i = \mathrm{tr}[\rho_1 E_i]$ and $q_i = \mathrm{tr}[\rho_2 E_i]$. The Kolmogorov distance between the two probability distributions produced by these measurements is $K_E(p,q) = \frac{1}{2}\sum_i |p_i - q_i|$, and it can be shown that $D[\rho_1,\rho_2] = \max_E K_E$ (i.e., the trace-norm distance equals the maximum over all possible generalized measurements of the Kolmogorov distance between the probability distributions resulting from measuring ρ_1 and ρ_2 [28]). The trace-norm distance is related to the Uhlman fidelity

$$F[\rho_1,\rho_2] := \|\sqrt{\rho_1}\sqrt{\rho_2}\|_1 = \mathrm{tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}, \quad \text{(A4)}$$

via

$$1 - D \leqslant F \leqslant \sqrt{1 - D^2}, \quad \text{(A5)}$$

so that one bounds the other [29].

APPENDIX B: PROOF OF THE VANISHING ORDERS IN UDD

The paper by Yang and Liu [10] sketched a proof of the universality of UDD in the interaction picture. In this appendix we provide the first comprehensive proof of the Yang-Liu universality result. Our proof is done in the toggling frame rather than the interaction picture.

We shall prove that all powers $n \leqslant N$ vanish in the expansion of the time-evolution operator in (III), which have an *odd* number of $\sigma_z B_z$ in $\widehat{Q}_{\vec{\alpha}}$, thus also an odd number of f_z in $F_{\vec{\alpha}}$. This is equivalent to showing that the first N powers in T of $B_-(T)$ vanish (i.e., that dephasing occurs only in order T^{N+1} or higher). Henceforth we use the shorthand $\bar{N} := N + 1$.

The substitution $s = \sin^2(\theta/2)$ suggests itself based on the UDD choice for the $\{\delta_j\}$ in (1) because it renders

$$\tilde{f}_\alpha(\theta) := f_\alpha\big(\sin^2(\theta/2)\big), \tag{B1}$$

particularly simple if the $\{\delta_j\}$ are chosen according to Eq. (1):

$$\tilde{f}_z(\theta) = (-1)^j, \tag{B2}$$

holds for $\theta \in \big(j\pi/\bar{N},(j+1)\pi/\bar{N}\big)$ with $j \in \{0,\ldots,N\}$. For simplicity, we will omit the tilde on the functions f_α from now on because only the argument θ will appear henceforth.

Since $f(\theta)$ enters the nested integrals only with an argument in $[0,\pi]$ it does not matter what we assume about $f(\theta)$ outside the limits of these integral, and we release the constraint on j, allowing $j \in \mathbb{Z}$. The function $f_z(\theta)$ then becomes an odd function with antiperiod $\pi/\bar{N}$. Thus its Fourier series

$$f_z(\theta) = \sum_{k=0}^{\infty} c_{2k+1} \sin[(2k+1)\bar{N}\theta], \tag{B3}$$

contains only harmonics $\sin(r\bar{N}\theta)$ with r an odd integer. The precise coefficients c_{2k+1} do not matter, a fact which can be exploited for other purposes (e.g., to deal with pulses of finite duration [26]).

Under the substitution $s = \sin^2(\theta/2)$ the infinitesimal element ds becomes $ds \to \frac{1}{2}\sin(\theta)\,d\theta$, converting (6b) into

$$\begin{aligned} F_{\vec{\alpha}} = &\int_0^{\pi} \sin(\theta_n)\,d\theta_n\, f_{\alpha_n}(\theta_n) \int_0^{\theta_n} \sin(\theta_{n-1})\,d\theta_{n-1}\, f_{\alpha_{n-1}}(\theta_{n-1}) \\ &\ldots \int_0^{\theta_3} \sin(\theta_2)\,d\theta_2\, f_{\alpha_2}(\theta_2) \int_0^{\theta_2} \sin(\theta_1)\,d\theta_1\, f_{\alpha_1}(\theta_1), \end{aligned} \tag{B4}$$

where we absorbed the $1/2$ factors coming from the infinitesimal elements into the coefficients c_{2k+1}.

What happens if we perform the successive integrations in Eq. (B4)? Replacing $f_z(\theta)$ by its Fourier series (B3) we deal with integrands which are products of trigonometric functions. The substitution gave rise to the factor $\sin\theta$. The Fourier series gives rise to additional factors $\sin(r_o\bar{N}\theta)$, where r_o is an odd integer. Recall the elementary trigonometric identities

$$\sin a \sin b = \tfrac{1}{2}[\cos(a-b) - \cos(a+b)], \tag{B5a}$$

$$\cos a \sin b = \tfrac{1}{2}[\sin(a+b) - \sin(a-b)], \tag{B5b}$$

$$\cos a \cos b = \tfrac{1}{2}[\cos(a+b) + \cos(a-b)]. \tag{B5c}$$

Using these, the most general trigonometric factor to occur in the course of the integrations in Eq. (B4) can be written as either $\sin[(q + r\bar{N})\theta]$ or $\cos[(q + r\bar{N})\theta]$ where r and q are integers. Since we are only concerned with values of n such that $n < \bar{N}$, the absolute value of q always remains below $\bar{N}$, so that the representation of the integer factor $(q + r\bar{N})$ in the arguments of the trigonometric functions is unique.

We now consider a complete set of four different cases which can occur in the course of the evaluation of each $F_{\vec{\alpha}}$. The first two cases are associated with the occurrence of $f_0 = 1$ in one or more of the nested integrals. Suppose for concreteness that this happens in nested integral number j. Then the factor $\sin(r_o\bar{N}\theta_j)$ does not occur in this integral since this factor arises exclusively due to the presence of $f_z(\theta_j)$. The two cases are now distinguished by whether a summand in the integrand of this jth integral, after a complete expansion of trigonometric products *excluding the* $\sin(\theta_j)$ *term*, into sums using Eqs. (B5a) through (B5c), involves the factor $\cos[(q + r\bar{N})\theta]$ (whence we call the integrand "cosine-type") or the factor $\sin[(q + r\bar{N})\theta]$ (whence we call the integrand "sine-type"). The third and fourth cases are associated with the occurrence of f_z in integrand number j. Then the factor $\sin(r_o\bar{N}\theta_j)$ does occur in this integrand, and again we distinguish two cases according to the presence of $\cos[(q + r\bar{N})\theta]$ ("cosine-type") or $\sin[(q + r\bar{N})\theta]$ ("sine-type") arising from a complete expansion of trigonometric products *excluding the* $\sin(\theta_j)$ *term and also the* $\sin(r_o\bar{N}\theta_j)$ *term*. Here then are the four cases in detail.

(i) Assume that one of the nested integrals contains $f_0 = 1$ and the factor $\cos[(q + r\bar{N})\theta]$. As we shall see in item (4) below this case occurs when r is even. Then this integral reads

$$\begin{aligned} 2\int &\cos[(q + r\bar{N})\theta]\sin(\theta)\,d\theta \\ &= \frac{\cos[(-1 + q + r\bar{N})\theta]}{-1 + q + r\bar{N}} - \frac{\cos[(1 + q + r\bar{N})\theta]}{1 + q + r\bar{N}}. \end{aligned} \tag{B6}$$

In writing Eq. (B6) we have assumed that the denominators do not vanish. The denominators may in fact vanish because r may be zero. When $r = 0$ the case $|q| = 1$ is special and yields

$$2\int \cos(\pm\theta)\sin(\theta)\,d\theta = -\frac{1}{2}\cos(2\theta). \tag{B7}$$

The important point is that both Eqs. (B6) and (B7) have only cosine terms on the right-hand side.

(ii) Assume that one of the nested integrals contains $f_0 = 1$ and the factor $\sin[(q + r\bar{N})\theta]$. As we shall see in item (4) this case occurs when r is odd. Then this integral reads

$$\begin{aligned} 2\int &\sin[(q + r\bar{N})\theta]\sin(\theta)\,d\theta \\ &= \frac{\sin[(-1 + q + r\bar{N})\theta]}{-1 + q + r\bar{N}} - \frac{\sin[(1 + q + r\bar{N})\theta]}{1 + q + r\bar{N}}. \end{aligned} \tag{B8}$$

No denominator can vanish because, as we shall see in item (2), $|q| < N$. The important point here is that Eq. (B8) has only sine terms on the right-hand side.

(iii) Assume that one of the nested integrals contains f_z and the factor $\cos[(q + r\bar{N})\theta]$. As we shall see in item (4) this case occurs when r is even. Then this integral reads

$$4\int \cos[(q+r\bar{N})\theta]\sin(r_o\bar{N}\theta)\sin(\theta)\,d\theta = \frac{\sin\{[-1+q+(r+r_o)\bar{N}]\theta\}}{-1+q+(r+r_o)\bar{N}} - \frac{\sin\{[1+q+(r+r_o)\bar{N}]\theta\}}{1+q+(r+r_o)\bar{N}} - \frac{\sin\{[-1+q+(r-r_o)\bar{N}]\theta\}}{-1+q+(r-r_o)\bar{N}} + \frac{\sin\{[1+q+(r-r_o)\bar{N}]\theta\}}{1+q+(r-r_o)\bar{N}}. \quad \text{(B9)}$$

Since $r \pm r_o$ is odd none of the denominators can vanish as long as $|q| < N$. Again, the important point here is that Eq. (B9) has only sine terms on the right-hand side.

(iv) Finally, assume that one of the nested integrals contains f_z and the factor $\sin[(q+r\bar{N})\theta]$. As we shall see in item (4) this case occurs when r is odd. Then this integral reads

$$4\int \sin[(q+r\bar{N})\theta]\sin(r_o\bar{N}\theta)\sin(\theta)\,d\theta = \frac{\cos\{[-1+q+(r-r_o)\bar{N}]\theta\}}{-1+q+(r-r_o)\bar{N}} - \frac{\cos\{[1+q+(r-r_o)\bar{N}]\theta\}}{1+q+(r-r_o)\bar{N}} - \frac{\cos\{[-1+q+(r+r_o)\bar{N}]\theta\}}{-1+q+(r+r_o)\bar{N}} + \frac{\cos\{[1+q+(r+r_o)\bar{N}]\theta\}}{1+q+(r+r_o)\bar{N}}. \quad \text{(B10)}$$

In writing Eq. (B10) we have assumed that the denominators do not vanish. A denominator can vanish only when $|q| = 1$, which leads to the two special cases $r = \pm r_o$. In analyzing these two cases we can assume without loss of generality that $q = 1$ and $r = r_o$. Otherwise we multiply the argument of the first and/or the second sine-function in the integrand by -1. This yields

$$4\int \sin[(1+r\bar{N})\theta]\sin(r\bar{N}\theta)\sin(\theta)\,d\theta = -\frac{\cos(2\theta)}{2} - \frac{\cos(2r\bar{N}\theta)}{2r\bar{N}} + \frac{\cos[(2+2r\bar{N})\theta]}{2+2r\bar{N}}. \quad \text{(B11)}$$

The important point here is that in Eq. (B10) only cosine terms appear on the right-hand side.

The number of possible terms proliferates in the course of the successive integrations. Therefore, in the sequel we discuss only the common features of the resulting summands. It is always understood that sums with varying sets of q and r are considered. We present a series of observations which leads to the desired proof of the cancellation of the first N powers in T of $B_-(T)$. The key to doing so will be to show that after integrating with an odd number of f_z factors we always end up with a sine-type integrand.

(1) Recall that we call an integrand summand "cosine-type" or "sine-type" if after complete expansion of all trigonometric products, excluding the $\sin(\theta)$ term arising from the change of variables and of the f_z term if it is there, the trigonometric factor is $\cos[(q+r\bar{N})\theta]$ or $\sin[(q+r\bar{N})\theta]$, respectively. Cases (i) and (iii) above are cosine-type, while cases (ii) and (iv) are sine-type. It is clear that the first integrand in (B4) is cosine-type with $r = 0$ and $q = 0$.

(2) We track which values of q may occur in each integration. The first integrand in (B4) is either $\sin(\theta_1)$ or $\sin(\theta_1)\sin(r\bar{N}\theta_1)$ (i.e., it has $q = 0$) so that this is our starting point. It follows from Eqs. (B6) through (B11) that each integration increments the possible maximum of $|q|$ by unity. The highest power in T studied is T^N so that there are $n \leqslant N$ integrations. This implies that the final value q_{final} before the very last integration obeys $|q_{\text{final}}| < N$.

(3) We track whether even or odd values of r occur at each integration. As mentioned in item (1), $r = 0$ holds in the first integration, so that our starting point is an even value. Each integration involving f_0 leaves r unchanged. Each integration involving f_z adds or subtracts r_o, so that r changes from even to odd or vice versa. If we combine this with the results of cases (i) through (iv) [Eqs. (B6)–(B11) this reveals the input-output table in Eq. (B12)]. The integrands (i.e., the inputs) are indicated by the case number in the table entries, while the values of the integrals (i.e., the outputs) are the types indicated in corresponding table entries. Also indicated is the transformation undergone by r from input to output.

$\times$	cosine-type	sine-type
f_0	case (i): cosine-type, $r \mapsto r$	case (ii): sine-type, $r \mapsto r$
f_z	case (iii): sine-type, $r \mapsto r \pm r_o$	case (iv): cosine-type, $r \mapsto r \pm r_o$.

(B12)

(4) Consider the output of the table as the input into the next integration and focus on the f_z row. Note that cases (iii) and (iv) alternate along with a change in parity of r (i.e., cosine-type changes into sine-type and vice versa) while odd r changes to even r and vice versa. *Therefore if we start with a cosine-type integrand and perform an odd number of f_z integrations, we will end up with a sine-type output* and a change in parity of r. For the same reason since the first integrand in (B4) is cosine-type with even r, and case (i) can only be arrived at after an even number of f_z integrations (the number of f_0 integrations is arbitrary), case (i) always involves even r. Repeating this reasoning explains why case (iii) also has even r, while cases (ii) and (iv) have odd r.

(5) Dephasing results only from the terms which comprise an *odd* number of f_z integrations. Considering that as noted in item (1) we start from a cosine-type integral and with $r = 0$, it follows from item (4) that *the last integration provides a sine-type* result. This integral can therefore be written as a sum over terms all of which are of the form

$$\sin[(q_{\text{final}} + 1 + r_{\text{final}}\bar{N})\theta]\big|_0^{\pi} = 0. \quad \text{(B13)}$$

Recall that the operator B_- in Eq. (11) comprises all the terms with odd number of $\sigma_z \otimes B_z$. Hence we have proven that the first N powers in T of $B_-(T)$ vanish [i.e., $B_-(T) = O(T^{N+1})$]. This is what we set out to show and concludes the derivation.

A remark concerning the result of Yang and Liu obtained in the interaction picture [10] is in order. They showed that $\exp(iTB_0)U(T)$ comprises only odd powers in σ_z which are of order T^{N+1} or higher. Since $\exp(\pm iTB_0)$ does not contain any term proportional to σ_z the Yang-Liu result implies our result and vice versa.

For time-dependent Hamiltonian, the proof in the interaction picture [9] is more convenient because powers in time occur anyway. Therefore we stress that the statement that only odd powers in σ_z of order T^{N+1} or higher occur is independent of the choice of reference frame (i.e., the description in the interaction picture or in the toggling frame).

APPENDIX C: DISTANCE BOUND CALCULATION

We prove the trace-norm distance bound Eq. (27)

$$\begin{aligned}
&2D\big[\rho_S(T),\rho_S^0(T)\big]\\
&\quad= \big\|\mathrm{tr}_B\big[\rho_{SB}(T)-\rho_{SB}^0(T)\big]\big\|_1\\
&\quad= \big\|\mathrm{tr}_B\big[U(T)\rho_{SB}^0(0)U^\dagger(T)\big]-\mathrm{tr}_B\big[\rho_{SB}^0(0)\big]\big\|_1\\
&\quad= \|\mathrm{tr}_B\{[I\otimes B_+(T)+\sigma_z\otimes B_-(T)](|\psi\rangle\langle\psi|\\
&\qquad\otimes\rho_B[I\otimes B_+^\dagger(T)+\sigma_z\otimes B_-^\dagger(T)]\}-|\psi\rangle\langle\psi|\,\|_1\\
&\quad= \|[b_{++}(T)-1]|\psi\rangle\langle\psi|+b_{+-}(T)|\psi\rangle\langle\psi|\sigma_z\\
&\qquad+b_{-+}(T)\sigma_z|\psi\rangle\langle\psi|+b_{--}(T)\sigma_z|\psi\rangle\langle\psi|\sigma_z\|_1. \qquad \text{(C1)}
\end{aligned}$$

We used the definition of $b_{\alpha\beta}(T)$ [Eq. (26)] in the last equality. Next, we use the triangle inequality, and finally the unitary invariance of the trace norm along with the normalization of $|\psi\rangle$

$$\begin{aligned}
&2D\big[\rho_S(T),\rho_S^0(T)\big]\\
&\quad\leqslant |b_{++}(T)-1|\,\||\psi\rangle\langle\psi|\,\|_1+|b_{+-}(T)|\,\||\psi\rangle\langle\psi|\sigma_z\|_1\\
&\qquad+|b_{-+}(T)|\,\|\sigma_z|\psi\rangle\langle\psi|\,\|_1+|b_{--}(T)|\,\|\sigma_z|\psi\rangle\langle\psi|\sigma_z\|_1\\
&\quad= |b_{++}(T)-1|+|b_{+-}(T)|+|b_{-+}(T)|+|b_{--}(T)|. \qquad \text{(C2)}
\end{aligned}$$

APPENDIX D: CORRELATION FUNCTION INEQUALITY

We prove the correlation function inequality (32). Consider the spectral decomposition of the bath density operator: $\rho_B=\sum_i\lambda_i|i\rangle\langle i|$, where $\{|i\rangle\}$ are normalized eigenstates, $\lambda_i\geqslant 0$ are the eigenvalues, and $\sum_i\lambda_i=1$. Defining $|v_i\rangle := Q|i\rangle$ and $|v_i'\rangle := (Q')^\dagger|i\rangle$, we have in this eigenbasis of ρ_B

$$\begin{aligned}
|\mathrm{tr}[Q\rho_B Q']| &= |\mathrm{tr}[Q'Q\rho_B]| = \Big|\sum_i\langle i|Q'Q|i\rangle\lambda_i\Big|\\
&= \Big|\sum_i\langle v_i'|v_i\rangle\lambda_i\Big| \leqslant \sum_i|\langle v_i'|v_i\rangle|\lambda_i\\
&\leqslant \sum_i\||v_i'\rangle\|\||v_i\rangle\|\lambda_i \leqslant \sum_i\|Q'\|\|Q\|\lambda_i\\
&= \|Q'\|\|Q\|, \qquad \text{(D1)}
\end{aligned}$$

where we used the triangle inequality, followed by the Cauchy-Schwartz inequality, and then the bounds $\||v_i\rangle\|=\|Q|i\rangle\|\leqslant\|Q\|$ and $\||v_i'\rangle\|=\|(Q')^\dagger|i\rangle\|\leqslant\|(Q')^\dagger\|=\|Q'\|$, which follow from the definition and properties of the supoperator norm (see Appendix A).

[1] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).
[2] E. Hahn, Phys. Rev. **80**, 580 (1950).
[3] L. Viola and S. Lloyd, Phys. Rev. A **58**, 2733 (1998).
[4] M. Ban, J. Mod. Opt. **45**, 2315 (1998).
[5] P. Zanardi, Phys. Lett. A **258**, 77 (1999).
[6] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).
[7] G. S. Uhrig, Phys. Rev. Lett. **98**, 100504 (2007).
[8] J. R. West, B. H. Fong, and D. A. Lidar, Phys. Rev. Lett. **104**, 130501 (2010).
[9] S. Pasini and G. S. Uhrig, J. Phys. A: Math. Theo. **43**, 132001 (2010).
[10] W. Yang and R.-B. Liu, Phys. Rev. Lett. **101**(18), 180403 (2008).
[11] B. Lee, W. M. Witzel, and S. Das Sarma, Phys. Rev. Lett. **100**, 160505 (2008).
[12] G. S. Uhrig, New J. Phys. **10**, 083024 (2008).
[13] M. Biercuk, H. Uys, A. VanDevender, N. Shiga, W. Itano, and J. Bollinger, Nature (London) **458**, 996 (2009).
[14] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Phys. Rev. A **79**, 062324 (2009).
[15] J. Du, X. Rong, N. Zhao, Y. Wang, J. Yang, and R. B. Liu, Nature (London) **461**, 1265 (2009).
[16] S. Pasini and G. S. Uhrig, Phys. Rev. A **81**, 012309 (2010).
[17] T. E. Hodgson, L. Viola, and I. D'Amico, Phys. Rev. A **81**, 062321 (2010).
[18] K. Khodjasteh and D. A. Lidar, Phys. Rev. Lett. **95**, 180501 (2005).
[19] K. Khodjasteh and D. A. Lidar, Phys. Rev. A **75**, 062310 (2007).
[20] G. S. Uhrig, Phys. Rev. Lett. **102**, 120502 (2009).
[21] K. Khodjasteh and D. A. Lidar, Phys. Rev. A **78**, 012355 (2008).
[22] D. A. Lidar, P. Zanardi, and K. Khodjasteh, Phys. Rev. A **78**, 012308 (2008).
[23] H.-K. Ng, D. A. Lidar, and J. P. Preskill (2009), e-print arXiv:0911.3202.
[24] J. Schliemann, A. Khaetskii, and D. Loss, J. Phys.: Condens. Matter **15**, R1809 (2003).
[25] J. R. West, D. A. Lidar, B. H. Fong, M. F. Gyure, X. Peng, and D. Suter (2009), e-print arXiv:0911.2398.
[26] G. S. Uhrig and S. Pasini, New J. Phys. **12**, 045001 (2010).
[27] R. Bhatia, *Matrix Analysis*, no. 169 in Graduate Texts in Mathematics (Springer-Verlag, New York, 1997).
[28] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[29] C. A. Fuchs and J. van de Graaf, IEEE Trans. Inf. Theory **45**, 1216 (1999).