Random Decoupling Schemes for Quantum Dynamical Control and Error Suppression

Lorenza Viola[1,*] and Emanuel Knill[2]

[1]Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, New Hampshire 03755, USA
[2]National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA

(Received 3 October 2004; published 17 February 2005)

We present a general control-theoretic framework for constructing and analyzing random decoupling schemes, applicable to quantum dynamical control of arbitrary finite-dimensional composite systems. The basic idea is to design the control propagator according to a random rather than deterministic path on a group. We characterize the performance of random decoupling protocols, and identify control scenarios where they can significantly weaken time scale requirements as compared to cyclic counterparts. Implications for reliable quantum computation are discussed.

DOI: 10.1103/PhysRevLett.94.060502 PACS numbers: 03.67.Pp, 03.65.Yz, 89.70.+c

Dynamical decoupling offers a versatile control toolbox for quantum dynamical engineering in both traditional settings like high-resolution spectroscopy [1] and quantum information science [2]. Decoupling schemes operate by subjecting the target system to a series of open-loop control transformations, in such a way that the net evolution is coherently modified to a desired one [3]. This combines intrinsic design simplicity with the ability to avoid auxiliary memory and measurement resources, while additionally enabling straightforward integration with other passive [4] or active [5] quantum control techniques. Applications of decoupling range from the removal of undesired couplings in interacting quantum subsystems to active decoherence control and symmetrization in open quantum systems [6]. In particular, the use of decoupling methods in conjunction with procedures for universal control [7] provides a route to noise-suppressed quantum computation based solely on unitary means. Remarkably, recent advances support the potential for highly fault-tolerant control architectures [8,9].

So far, general formulations of the decoupling problem have been restricted to *deterministic* control actions. In the simplest, so-called *bang-bang* setting, where the latter are instantaneous rotations drawn from a group $\mathcal{G}$, decoupling according to $\mathcal{G}$ is enforced by cycling the control propagator through *all* group elements, translating into pulse sequences with minimal length T_c proportional to the size of $\mathcal{G}$ [3]. This suffers from two main drawbacks. Because averaging requires traversing all of $\mathcal{G}$ in a suitable sense, decoupling becomes very inefficient for large groups, leading to unrealistically high control rates if the interactions to be removed have a short correlation time τ_c. Furthermore, it is not clear how to handle interactions which are themselves fluctuating on time scales which are short compared to the averaging period T_c. These limitations severely constrain the practicality of decoupling as a strategy for decoherence suppression in open systems.

In this Letter, we propose to overcome the above limitations by introducing a framework for *random dynamical decoupling*. Physically, our approach takes inspiration from a naturally occurring instance of a random decoupling process, that is, the self-averaging of intermolecular interactions in gases and isotropic liquids due to random translational and reorientational motions [1]. This intuition is cast in control-theoretic language by requesting that the control propagator follows a random but *known* path on $\mathcal{G}$ [10]. We show how random decoupling may be used to achieve a desired coherent averaging and obtain a bound on worst-case performance. By comparing to ordinary cyclic schemes, we find that, in the presence of rapidly fluctuating interactions and/or large control groups, randomized design may prove superior. From the point of view of decoherence suppression, this not only establishes in general the counterintuitive possibility to actively cancel noise using randomness, but it also opens new prospects for significantly mitigating time scale requirements in a wide class of control systems.

Random decoupling setting.—Let S be a quantum system with state space $\mathcal{H}_S$, $\dim(\mathcal{H}_S) = d < \infty$, evolving under an arbitrary, possibly time-dependent drift Hamiltonian $H_0(t)$. Without loss of generality, we assume $H_0(t)$ to be traceless for all t. We begin by constructing a random decoupling protocol for effectively switching off the evolution due to $H_0(t)$, under the assumption of perfect, unbounded control. Let the available control generate a discrete or continuous compact group $\mathcal{G}$, acting on $\mathcal{H}_S$ via a faithful, unitary, projective representation μ, $\mu(g) = \hat{g}$ for $g \in \mathcal{G}$, $\mu(\mathcal{G}) = \hat{\mathcal{G}}$. A *random decoupler* uses control in $\mathcal{G}$ in two ways: first, to establish a *logical* frame that is related to the physical one [where $H_0(t)$ is specified] by an element of $\hat{\mathcal{G}}$; second, to rotate the system according to $\mathcal{G}$ randomly over time, by following a random control path $U_c(t)$. Thus, both the past control operations and the times at which they are applied are known, but the *future* control path is random.

The essence of the random decoupling approach is to directly depict the evolution of the system in the logical frame that continuously follows the applied control. Let $\rho_S(t) = U(t)\rho_S(0)U^\dagger(t)$ describe the state of the system in the physical frame, evolving under the action of both the internal Hamiltonian and the controller, and let

$\tilde{\rho}_S(t) = U_c^\dagger(t)\rho_S(t)U_c(t)$ denote the corresponding logical state, with $\tilde{\rho}_S(0) = \rho_S(0)$. Then the evolution in the logical frame is fully specified by a propagator ($\hbar = 1$)

$$\tilde{U}(t) = U_c^\dagger(t)U(t) = \mathcal{T}\exp\left\{-i\int_0^t du\tilde{H}(u)\right\},$$

where $\tilde{H}(t) = U_c^\dagger(t)H_0(t)U_c(t)$. Under the usual cyclicity assumption of deterministic decoupling, $U_c(t + T_c) = U_c(t)$ for $T_c > 0$, the physical and logical frames stroboscopically coincide at times $t_N = NT_c$, $N \in \mathbb{N}$. By contrast, random decoupling is intrinsically *acyclic*, and the control path almost never returns the system to the physical frame. However, the available information about the past control trajectory may be exploited to bring the state of the system back to the physical frame if desired.

Error bounds.—To determine whether and how well random decoupling succeeds at suppressing the dynamics due to $H_0(t)$, it is necessary to compare the evolution under the propagator $\tilde{U}(T)$ over a time interval T to the identity evolution, up to a global phase. A natural measure is provided by the error probability for an arbitrary pure initial state $P_S = |\psi\rangle\langle\psi|$ of S. With respect to the random nature of the control path, the *a priori* error probability can be expressed as an expectation

$$\begin{aligned}\epsilon_T(P_S) &= \mathbb{E}\{\mathrm{tr}_S[P_S^\perp\tilde{\rho}_S(T)]\}\\ &= \mathbb{E}\{\mathrm{tr}_S[P_S^\perp\tilde{U}(T)P_S\tilde{U}(T)^\dagger]\},\end{aligned} \tag{1}$$

where $P_S^\perp = \mathbb{1}_S - |\psi\rangle\langle\psi|$ is the orthogonal complement of P_S and $\mathbb{E}$ denotes ensemble average. Then a *worst-case pure state error probability* may be defined as

$$\epsilon_T = \max_{P_S}\{\epsilon_T(P_S)\}. \tag{2}$$

A quantitative bound for ϵ_T is contained in the following:

Theorem 1: *Suppose that (i) $\mathcal{G}$ acts irreducibly on $\mathcal{H}_S$. (ii) $U_c(t)$ is uniformly random for each t. (iii) For any $t, s > 0$, $U_c(t)$ and $U_c(t + s)$ are independent for $s > \Delta t$. (iv) $||H_0(t)||_2$ is uniformly bounded in time by $k > 0$. Then*

$$\epsilon_T = O(T\Delta t k^2) \quad \text{for } T\Delta t k^2 \ll 1. \tag{3}$$

Here, $||A||_2 = \max|\mathrm{eig}(\sqrt{A^\dagger A})|$, and uniformly random is intended relative to the invariant Haar measure $\nu_\mathcal{G}$ on $\mathcal{G}$, normalized such that $\nu_\mathcal{G}(\mathcal{G}) = 1$ [11]. While a rigorous proof of the above Theorem 1 is rather lengthy [12], an outline of the underlying strategy suffices for gaining physical insight. The key step is to realize that, in each of the integrals involved in the Dyson series expansion of the time-ordered exponentials defining $\tilde{U}(T)$ and $\tilde{U}^\dagger(T)$ in Eq. (1), the independence assumption (iii) effectively partitions the integration domain in two separate regions: a volume $W_1(\Delta t)$, where none of the integration variables is more than Δt away from all the remaining ones; and the complement $W_2(\Delta t)$, where this condition is violated by at least one variable. The expectation relative to such a variable may be taken separately, leading, under the uniformity assumption (ii), to a contribution of the form

$$\mathbb{E}\{U_c^\dagger(t)H_0(t)U_c(t)\} = \int_\mathcal{G} d\nu_\mathcal{G}\,\hat{g}^\dagger H_0(t)\hat{g},$$

$t \in (0, \Delta t)$.

Notice, as the result of such an *ensemble* average, the appearance of the same dynamical $\mathcal{G}$ symmetrization which, in standard deterministic schemes, is achieved through the *time* average over a cycle [3,13]. In particular, the irreducibility assumption (i) implies maximal projection in the set of scalars. That is, for X traceless,

$$\int_\mathcal{G} d\nu_\mathcal{G}\,\hat{g}^\dagger X\hat{g} = \frac{\mathrm{tr}(X)}{d}\mathbb{1}_S = 0. \tag{4}$$

As a consequence, all terms originating from $W_2(\Delta t)$ vanish, and the desired upper bound to $\epsilon_T(P_S)$ may be determined by estimating the volume of $W_1(\Delta t)$. The irreducibility assumption can of course be weakened. As it turns out, the final result (3) for ϵ_T has a simple intuitive explanation, which we defer until after we describe the corresponding error bound for deterministic schemes.

From an implementation perspective, one may distinguish two main scenarios, depending on whether the decoupler is specified by a continuous or discrete control group $\mathcal{G}$. In the former case, the decoupling time scale Δt is *defined* by the independence requirement between $U_c(t)$ and $U_c(t + s)$, condition (iii) entering as a design constraint. Note that bounded-strength controls might suffice as long as Δt is finite. If $\mathcal{G}$ is discrete, the required random walk of $U_c(t)$ may be enforced through a sequence of equally spaced bang-bang pulses randomly drawn from $\hat{\mathcal{G}}$. In this case, the independence requirement is automatically satisfied by identifying Δt with the separation between consecutive kicks. Either way, it is important to stress that random decoupling (unlike deterministic decoupling) places *no* restriction on the temporal behavior of $H_0(t)$, only on its maximum eigenvalue.

Random decoherence suppression.—The above formalism can be extended to the suppression of noise effects arising from the coupling between the target system S and an uncontrollable quantum environment E. Let the total drift Hamiltonian be expressed in the form $H_0(t) = \mathbb{1}_S \otimes H_E + \sum_a J_a(t) \otimes B_a$, where H_E accounts for the (typically unknown) evolution of E and the internal evolution of S is included among the interaction operators, with $\mathrm{tr}[J_a(t)] = 0$ for all t. The action of the decoupler is understood as $U_c(t) \otimes \mathbb{1}_E$. Physically, it is meaningful to define a pure-state error probability that depends only on the *reduced* state of S in the logical frame. That is, $\tilde{\rho}_S(T)$ in Eq. (1) is now calculated as $\tilde{\rho}_S(T) = \mathrm{tr}_E[\tilde{U}(T)\tilde{\rho}_{SE}(0)\tilde{U}^\dagger(T)]$, $\tilde{\rho}_{SE}(0) = \rho_{SE}(0)$ being the joint initial state and, as before, the logical propagator $\tilde{U}(t)$ describing the combined evolution in a frame that explicitly removes the control field. By purifying the environment, we can assume that $\rho_{SE}(0) = P_S \otimes P_E$, both P_S and P_E being one-dimensional projectors. The derivation of a bound for $\epsilon_T(P_S)$ may be formally carried out following

the same steps as in the uncoupled case. It suffices to observe that Eq. (1) is equivalent to

$$\epsilon_T(P_S) = \mathbb{E}\{\mathrm{tr}_{S,E}[P_S^{\perp} \otimes \mathbb{1}_E \tilde{U}'(T) P_S \otimes P_E \tilde{U}'(T)^{\dagger}]\},$$

with the propagator

$$\tilde{U}'(t) = U_E^{\dagger}(t) U_c^{\dagger}(t) U(t) = \mathcal{T} \exp\left\{-i \int_0^t du \tilde{H}'(u)\right\}$$

describing the evolution in a frame where *both* the applied control and the environment dynamics $U_E(t) = \exp(-iH_E t)$ are explicitly removed, and $\tilde{H}'(t) = \sum_a U_c^{\dagger}(t) J_a(t) U_c(t) \otimes B_a$. We thus have the following:

Theorem 2: *Let $\mathcal{G}$ act irreducibly on $\mathcal{H}_S$ and satisfy the same uniformity and independence assumptions as in Theorem 1. If $||\sum_a J_a(t) \otimes B_a||_2$ is uniformly bounded in time by $\lambda > 0$, then*

$$\epsilon_T = O(T \Delta t \lambda^2) \quad \text{for } T \Delta t \lambda^2 \ll 1. \tag{5}$$

Formally, λ is a measure of the overall *noise strength* as defined in the context of quantum error-correction theory [14]. As pointed out in this reference, caution is required in treating infinite-dimensional environments. Physically, $1/\lambda = \tau_c$ is of the order of the shortest correlation time scale present in the interaction to be removed. While the latter provides the relevant time scale to consider in the absence of additional information about the environment's initial state, power spectrum, and internal dynamics, such properties may critically impact the decoupling performance in actual applications [15]. Thus, *lower* error bounds tend to be fairly example specific.

According to the above theorems, ϵ_T can in principle be made arbitrarily small by appropriate control design, implying the possibility to *arbitrarily suppress on average* the unwanted evolution in the logical frame. This is especially surprising for decoherence suppression considering that, in the physical frame, the applied random field appears to be in general a source of decoherence. It is worth noting that the possibility to exploit randomization was considered earlier for specific decoupling problems. Preservation of coherence of a lossy radiation mode via the random modulation of a system parameter was established in [16]. More recently, a randomized refocusing algorithm was proposed in [17] in the context of efficient simulation of quantum computation starting from few-body Hamiltonians on n qubits. While revisiting such specific situations in the light of the present analysis is interesting in itself, our main goal in what follows is to continue developing a model-independent formulation of random decoupling in general control-theoretic terms.

Comparison with cyclic decoupling.—In order to assess the performance and usefulness of random decoupling schemes, a comparative error bound for deterministic decoupling is needed. We focus on the standard situation where the drift Hamiltonian H_0 is time independent, and decoupling is accomplished by cyclic averaging over a finite group of order $|\mathcal{G}| > 1$. Apart from the redundant ensemble expectation, Eqs. (1) and (2) still define a valid worst-case pure state error probability. The deterministic counterpart to Theorem one is then the following:

Theorem 3: *Suppose that (i) $\mathcal{G}$ acts irreducibly on $\mathcal{H}_S$. (ii) $U_c(t)$ is assigned according to a cyclic path over $\mathcal{G}$, with $U_c(t) = \hat{g}_j$ for $t \in [j\Delta t, j + 1\Delta t)$, $j = 0, \ldots, |\mathcal{G}| - 1$, $\Delta t > 0$, and $T_c = |\mathcal{G}|\Delta t$. (iii) $||H_0||_2$ is bounded by $k > 0$, with $kT_c < 1$. Then*

$$\epsilon_T = O((TT_c k^2)^2) \quad \text{for } TT_c k^2 \ll 1 - kT_c. \tag{6}$$

The proof follows from a direct evaluation of the logical propagator $\tilde{U}(T)$ using average Hamiltonian theory [1],

$$\tilde{U}(T) = e^{-i\overline{H}T}, \qquad \overline{H} = \sum_{\ell=0}^{\infty} \overline{H}^{(\ell)},$$

where $\overline{H}$ is computed from the Magnus expansion under the averaging and convergence conditions, $\overline{H}^{(0)} = 0$ and $kT_c < 1$, respectively [12]. We now provide an intuitive justification to the error bounds we found.

Write $R = T\Delta t k^2 = (k\Delta t)^2 (T/\Delta t)$. For the random method, each control step can accumulate an error amplitude of up to $k\Delta t$. Randomizing the decoupler has the net effect that the amplitudes add up probabilistically. Therefore, over an evolution time T, the total error probability is bounded by the number $T/\Delta t$ of such intervals, times the error probability $(k\Delta t)^2$ of each step. Notice that the bound of Theorem 1 is indeed $\epsilon_T^R = O(R)$.

For the cyclic method using $|\mathcal{G}|$ steps of duration Δt in each cycle, the dominant errors are due to $\overline{H}^{(1)}$. That is, they arise from noncommuting contributions associated with pairs of intervals in a cycle. Thus, for each cycle the error amplitude is bounded by $|\mathcal{G}|^2 (k\Delta t)^2$, and a total time T contains $T/(|\mathcal{G}|\Delta t)$ such cycles. If, as assumed, each cycle is identical and the interaction is constant, the total error amplitude is bounded by the sum, yielding $|\mathcal{G}|R$. By squaring and using that $|\mathcal{G}|\Delta t = T_c$, the bound of Theorem 3 emerges, $\epsilon_T^D = O(|\mathcal{G}|^2 R^2)$.

The above analysis shows that the worst-case errors of the two procedures compare as follows:

$$\epsilon_T^R = O(R) \quad \text{vs } \epsilon_T^D = O((|\mathcal{G}|^2 R) R),$$

the quantity $|\mathcal{G}|^2 R$ becoming a relevant figure of merit for performance. Thus, cyclic decoupling tends to perform better if any time dependence or fluctuations in the interactions to be removed have time scale *longer* than $T_c = |\mathcal{G}|\Delta t$ and, in addition, $|\mathcal{G}|^2 R \ll 1$. Superior performance of random decoupling is expected instead in situations where the effective correlation or fluctuations have time scales large compared to Δt but *short* compared to $|\mathcal{G}|\Delta t$, or, alternatively, $|\mathcal{G}|^2 R \gg 1$.

Generalizations and applications.—The above results lend themselves to a number of generalizations. The extension to reducible group actions (hence selective decoupling) is conceptually straightforward. Procedures for universal decoupled control may be designed similarly to [7], by randomly modulating the applied control

Hamiltonians to compensate for the decoupler action if necessary. This paves the way to schemes for randomly controlled noise-suppressed universal quantum computation. In addition, one may envisage a variety of hybrid control schemes where deterministic and random operations are simultaneously exploited. At least two options are worth considering. First, one may randomize the decouplers. If multiple decouplers are available to effect a desired averaging, which one to apply may be picked at random at every cycle. Or, with a single decoupler, one may randomize the cycles, by randomly choosing which path to follow to traverse $\mathcal{G}$. While a clever concatenation of deterministic and random protocols could merge advantageous features from both methods, quantitative error estimates as well as studies of the *typical* performance in specific situations will be reported elsewhere.

We anticipate that randomization might offer substantial benefits whenever a large number of control time slots is involved. An extreme example is maximal decoupling in n arbitrarily coupled qubits, $d = 2^n$. Deterministic group-based schemes require averaging over the Pauli error basis $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n}$, with $|\mathcal{G}| = d^2 = 4^n$ [3]. For fixed control parameters $T, \Delta t$ such that $R \ll 1$, the condition $|\mathcal{G}|^2 R \ll 1$ becomes exponentially harder to meet as n increases. Equivalently, for a fixed tolerable error ϵ_T, an interval Δt that shrinks exponentially with n is needed to compensate $|\mathcal{G}|^2$ in this case. A randomized implementation of Pauli decoupling is indeed at the heart of the simulation algorithm mentioned above [17]. In addition, the recently proposed Pauli-random-error-correction method for coherent errors [18] may also be understood as an ingenious application of the present control framework, random Pauli rotations being repeatedly applied to average *static* imperfections, and permutations of the original logic gates ensuring the intended decoupled control. While cyclic schemes with quadratic complexity [19] are known for bilinearly coupled qubits as assumed in [18], randomized schemes may still be attractive for large n and/or time-varying couplings. In the same spirit, the cancellation of rapidly fluctuating *dynamical* imperfections reported in [20] may be suggestively reinterpreted as a random self-decoupling effect. Lastly, efficiency improvements are to be expected from decoupling according to the symmetric group S_n acting on n qubits, which otherwise involves factorial overheads, and is relevant to the synthesis of collective noise [21].

Conclusion.—We introduced an approach to dynamical decoupling that relies on random control design. Beside being interesting *per se* as a largely unexplored setting for coherent and error control, random dynamical decoupling carries the potential for faster convergence and relaxed timing constraints compared to deterministic counterparts in relevant situations. While additional work is needed to expand the present analysis, we believe that our results add to the significance of decoupling methods as a control-theoretic tool and allow a step forward toward making them a practical error control strategy in quantum information science.

L. V. acknowledges support from LANL during the early stages of this work. E. K. was supported by the U.S. NSA. We thank Howard Barnum for discussions. Contributions to this work by NIST, an agency of the U.S. government, are not subject to copyright laws.

*Corresponding author.
Electronic address: lorenza.viola@dartmouth.edu

[1] U. Haeberlen, *High Resolution NMR in Solids: Selective Averaging* (Academic Press, New York, 1976).
[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[3] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).
[4] M. S. Byrd and D. A. Lidar, Phys. Rev. Lett. **89**, 047901 (2002); E. M. Fortunato *et al.*, New J. Phys. **4**, 5.1 (2002).
[5] N. Boulant *et al.*, Quant. Info. Proc. **1**, 135 (2002); K. Khodjasteh and D. A. Lidar, Phys. Rev. A **68**, 022322 (2003).
[6] See, e.g., L. Viola, Phys. Rev. A **66**, 012307 (2002).
[7] L. Viola, S. Lloyd, and E. Knill, Phys. Rev. Lett. **83**, 4888 (1999).
[8] L. Viola and E. Knill, Phys. Rev. Lett. **90**, 037901 (2003).
[9] K. Khodjasteh and D. A. Lidar, quant-ph/0408128.
[10] Knowledge of the path may be ignored if only operators commuting with $\mathcal{G}$ are used, as is usually the case in the context of self-averaging of intermolecular interactions.
[11] In the special case of $\mathcal{G}$ finite with order $|\mathcal{G}|$, integration relative to $\nu_{\mathcal{G}}$ simply means $\int_{\mathcal{G}} d\nu_{\mathcal{G}}(\cdot) = 1/|\mathcal{G}| \sum_{g \in \mathcal{G}} (\cdot)$.
[12] See EPAPS Document No. E-PRLTAO-94-073509 for appendix containing the proofs of Theorems 1 and 3. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.
[13] P. Zanardi, Phys. Rev. A **60**, R729 (1999).
[14] E. Knill, R. Laflamme, and L. Viola, Phys. Rev. Lett. **84**, 2525 (2000).
[15] L. Faoro and L. Viola, Phys. Rev. Lett. **92**, 117905 (2004); L. Falci, A. D'Arrigo, A. Mastellone, and E. Paladino, Phys. Rev. A**70**, 040101 (2004).
[16] S. Mancini, D. Vitali, P. Tombesi, and R. Bonifacio, Europhys. Lett. **60**, 498 (2002).
[17] M. J. Bremner, J. Dodd, M. A. Nielsen, and D. Bacon, Phys. Rev. A **69**, 012313 (2004).
[18] O. Kern, G. Alber, and D. L. Shepelyansky, quant-ph/0407262.
[19] M. Rötteler and P. Wocjan, quant-ph/0409135, and references therein.
[20] P. Facchi, S. Montangero, R. Fazio, and S. Pascazio, quant-ph/0407098.
[21] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **85**, 3520 (2000); L.-A. Wu and D. A. Lidar, *ibid.* **88**, 207902 (2002).