Landscape for optimal control of quantum-mechanical unitary transformations

Herschel Rabitz,[1] Michael Hsieh,[1] and Carey Rosenthal[2]
[1]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Department of Chemistry, Drexel University, Philadelphia, Pennsylvania 19104, USA*
(Received 14 October 2003; published 30 November 2005)

The optimal creation of a targeted unitary transformation W is considered under the influence of an external control field. The controlled dynamics produces the unitary transformation U and the goal is to seek a control field that minimizes the cost $J=\|W-U\|$. The optimal control landscape is the cost J as a functional of the control field. For a controllable quantum system with N states and without restrictions placed on the controls, the optimal control landscape is shown to have extrema with $N+1$ possible distinct values, where the desired transformation at $U=W$ is a minimum and the maximum value is at $U=-W$. The other distinct $N-1$ extrema values of J are saddle points. The results of this analysis have significance for the practical construction of unitary transformations.

DOI: 10.1103/PhysRevA.72.052337 PACS number(s): 03.67.Lx, 03.67.−a, 03.65.Ta

I. INTRODUCTION

The control of quantum systems is increasingly being explored for a variety of fundamental and practical reasons [1]. One area of application is to quantum information sciences [2], where a general objective is to create quantum "machines" that have unusual capabilities. At the heart of a quantum computer are devices constructed to act as particular unitary transformations. A natural desire is to achieve the best quality unitary transformations possible [3] in relation to the specified target transformations under the influence of an external control field $C(t)$, likely of an electromagnetic nature. The present work seeks to explore the control of devices operating to produce a specified transformation under ideal conditions as a limit of what may be expected. The particular focus is on the extrema that may be encountered when searching for optimal controls to generate a unitary transformation.

II. THE OPTIMAL CONTROL LANDSCAPE

Consider a quantum system of N states undergoing unitary evolution in the presence of a control field $C(t)$. The goal is to steer the dynamics described by the unitary matrix $U(t)$ over the interval $0 \leqslant t \leqslant T$ such that $U \equiv U(T)$ is as close as possible to the desired target unitary transformation W. The system is assumed to be controllable [4], implying that at least one field $C(t)$ exists for constructing $U=W$. Regardless of whether numerical design is first performed or direct laboratory control discovery [5] is employed, both routes involve optimization of a target cost function J which may be defined as

$$J = \sum_{i,j} |W_{ij} - U_{ij}|^2 = 2N - 2 \text{ Re Tr}[W^\dagger U]. \tag{1}$$

In the laboratory, U would not be directly measured, but through other observations, in principle, it could be constructed for use in J. For purposes of analysis, here we adopt Eq. (1) and the goal is to minimize $J[C(t)]$, which is a functional of the control field $C(t)$, through $U[C(t)]$. A concern in seeking to minimize $J[C(t)]$ is the possible presence of extrema with values $J>0$. If the search for an optimal control leads to trapping in a nonoptimal value for J, then the resultant quantum information device likely would not perform as desired. Another possibility is the presence of a large number of saddle point extrema which could make the search inefficient for locating a control satisfying $J=0$. Thus, a basic question is the possible existence of solutions to

$$\frac{\delta J}{\delta C(t)} = 0 \tag{2}$$

which may correspond to local extrema with $J>0$. The control $C(t)$ will be allowed to vary freely, implying that all landscape extrema will satisfy Eq. (2). The analysis does not require knowledge of the Hamiltonian, except for specification that the system be controllable at some time T. The assumptions of full controllability and unfettered access to any control permits a complete assessment of the general control landscape features. The actual accessible landscape domain explored in any particular case will be restricted by the details of the system Hamiltonian and especially any constraints placed on the controls (e.g., constrained bandwidth, a limited time T to reach the target, etc.).

III. ENUMERATION OF CRITICAL VALUES

In order to explore the cost function landscape, it is convenient to write $W=\exp(iB)$ and $U=\exp(iA)$ with $B^\dagger=B$ and $A^\dagger=A$, without any loss of generality. The $N\times N$ matrix B is *a priori* assumed known along with its eigenstates $\{|\beta_i\rangle\}$ and eigenvalues $\{b_i\}$ where $B|\beta_i\rangle=b_i|\beta_i\rangle$. We may express the trace operation in Eq. (1) in terms of the basis $\{|\beta_i\rangle\}$ combined with $U=\exp(iA)$ to rewrite Eq. (2) as

$$\frac{\delta J}{\delta C(t)} = -\sum_i \sum_{pq} \left(\exp(-ib_i) \frac{\partial}{\partial A_{pq}} \langle \beta_i | \exp(iA) | \beta_i \rangle + \exp(ib_i) \frac{\partial}{\partial A_{pq}} \langle \beta_i | \exp(-iA) | \beta_i \rangle \right) \frac{\delta A_{pq}}{\delta C(t)} = 0. \quad (3)$$

Since the quantum system is assumed to be controllable, then each of the matrix elements A_{pq} must be uniquely addressable by the control field $C(t)$ for all p and q values in keeping with A being Hermitian. These points imply that the set of N^2 functions $\{\delta A_{pq}/\delta C(t)\}$ should be linearly independent [4], leading to the criteria for satisfying Eq. (3) as being

$$\Delta_{qp} = \sum_i \left(\exp(-ib_i) \frac{\partial}{\partial A_{pq}} \langle \beta_i | \exp(iA) | \beta_i \rangle + \exp(ib_i) \frac{\partial}{\partial A_{pq}} \langle \beta_i | \exp(-iA) | \beta_i \rangle \right) = 0 \quad (4)$$

for all values of p and q. The transfer from Eq. (2) to Eq. (4) defining the optimal solutions is central to the analysis. An understanding of the mapping $C(t) \to U$ in Eq. (2) is generally highly complex and system-specific. The analysis of Eq. (4) assures satisfaction of Eq. (2) and (4) has a generic character as it only depends on the matrices A and B. Thus, the analysis of Eq. (4) can lead to some general conclusions about the nature of optimal control of unitary transformations.

The derivatives in Eq. (4) may be evaluated with the aid of the identity

$$\frac{\partial}{\partial A_{pq}} \exp(iA) = i \int_0^1 ds \exp[iA(1-s)] \frac{\partial A}{\partial A_{pq}} \exp(iAs).$$

The matrix A is represented in any convenient basis $\{|\ell\rangle\}$, and the eigenvectors $\{|\alpha_i\rangle\}$ and eigenvalues $\{a_i\}$ of A satisfy $A|\alpha_i\rangle = a_i|\alpha_i\rangle$. Utilizing that $\{|\ell\rangle\}$ and $\{|\alpha_i\rangle\}$ are complete sets of states in the space of dimension N yields

$$\begin{aligned}
\frac{\partial}{\partial A_{pq}} \langle \beta_i | \exp(iA) | \beta_i \rangle
&= i \int_0^1 ds \sum_{\ell,\ell'} \sum_{j,k} \langle \beta_i | \alpha_j \rangle \exp[ia_j(1-s)] \langle \alpha_j | \ell \rangle \times \langle \ell | \frac{\partial A}{\partial A_{pq}} | \ell' \rangle \langle \ell' | \alpha_k \rangle \exp(ia_k s) \langle \alpha_k | \beta_i \rangle \\
&= i \int_0^1 ds \sum_{\ell,\ell'} \sum_{j,k} \langle \beta_i | \alpha_j \rangle \exp[ia_j(1-s)] \langle \alpha_j | \ell \rangle \times \delta_{\ell p} \delta_{\ell' q} \langle \ell' | \alpha_k \rangle \exp(ia_k s) \langle \alpha_k | \beta_i \rangle \\
&= \sum_{j,k} \langle \beta_i | \alpha_j \rangle \langle \alpha_j | p \rangle \langle q | \alpha_k \rangle \langle \alpha_k | \beta_i \rangle \left(\frac{\exp(ia_j) - \exp(ia_k)}{a_j - a_k} \right).
\end{aligned} \quad (5)$$

Substituting the result of Eq. (5) into Eq. (4) and without loss of generality applying the unitary transformation $\Sigma_{pq} \langle \alpha_r | q \rangle \Delta_{qp} \langle p | \alpha_s \rangle$ to Eq. (4) produces the result

$$\sum_{p,q} \langle \alpha_r | q \rangle \Delta_{qp} \langle p | \alpha_s \rangle = \sigma_{rs} F(a_r, a_s) + \sigma_{sr}^* F(a_r, a_s)^* = 0, \quad (6)$$

where

$$F(a_r, a_s) = \begin{cases} i \exp(ia_r) & \text{if } a_r = a_s, \\ \dfrac{\exp(ia_r) - \exp(ia_s)}{a_r - a_s} & \text{if } a_r \neq a_s, \end{cases}$$

and

$$\sigma_{rs} = \sum_i \langle \alpha_r | \beta_i \rangle \langle \beta_i | \alpha_s \rangle \exp(-ib_i) = \langle \alpha_r | \exp(-iB) | \alpha_s \rangle \quad (7)$$

for all r and s. In Eqs. (6) and (7), first considering the particular case $a_r = a_s$ produces the result $\sigma_{rr} \exp(ia_r) - \sigma_{rr}^* \exp(-ia_r) = 0$. Defining $\sigma_{rr} = |\sigma_{rr}| \exp(i\delta_r)$, with δ_r being a real phase, reveals that $a_r + \delta_r = n_r \pi$, with n_r being an integer. Considering Eq. (6) again for $a_r \neq a_s$ yields

$$\langle \alpha_r | [\exp(-iB), \exp(iA)] + [\exp(iB), \exp(-iA)] | \alpha_s \rangle = 0. \quad (8)$$

Equation (8) is also satisfied for $a_r = a_s$, given that $\{|\alpha_i\rangle\}$ are eigenstates of A. Thus, Eq. (8) is valid for all r and s. It follows [6] that $[A, B] = 0$, implying that A and B have simultaneous eigenstates.

Armed with the results above, we may return to the cost function in Eq. (1), which can be written as

$$\begin{aligned}
J &= 2N - 2\,\mathrm{Re} \sum_r \left(\sum_i \exp(-ib_i) \langle \beta_i | \alpha_r \rangle \langle \alpha_r | \beta_i \rangle \right) \exp(ia_r) \\
&= 2N - 2\,\mathrm{Re} \sum_r |\sigma_{rr}| \exp[i(a_r + \delta_r)] \\
&= 2N - 2 \sum_r |\sigma_{rr}| (-1)^{n_r},
\end{aligned} \quad (9)$$

where $a_r + \delta_r = n_r \pi$ was used. It also follows that $\delta_r = -b_r$ from Eq. (7) by setting $r = s$ and noting that $|\alpha_r\rangle$ is an eigenstate of B; this shows that the eigenvalues of A and B at the control extrema, $a_r - b_r = n_r \pi$, differ from each other by integral multiples of π. Similarly utilizing Eq. (7), it is evident that $|\sigma_{rr}| = 1$. We may finally conclude that the cost function in Eq. (1) only has the following possible $N+1$ extrema values:

$$J = 0, 4, \cdots, 4N. \quad (10)$$

A shift in the overall phase φ of the target $W \to W \exp(i\varphi)$ just correspondingly produces a shift in the phase of the solutions $U \to U \exp(i\varphi)$ or $A \to A + i\varphi$, but the cost function values in Eq. (10) remains the same and the topology of the landscape discussed below does not change. Each of the extrema corresponding to a particular A matrix and J value would have an associated control field $C(t)$ in an actual application.

IV. HESSIAN ANALYSIS OF CRITICAL POINTS

The locally optimal solutions in Eq. (10) range from that of perfect control $J = 0$ in creating the target $U = W$ out to the outcome at $J = 4N$ corresponding to $U = -W$. The latter two outcomes for J are, respectively, a minimum and a maximum of the cost function, both corresponding to perfect solutions (up to phase). The topology of the remaining N–1 distinct extrema in Eq. (10) can be deduced by considering the definite nature of

$$\mathcal{H} = \int_0^T \int_0^T \omega(t) \frac{\delta^2 J}{\delta C(t) \delta C(t')} \omega(t') dt\, dt', \qquad (11)$$

where $\omega(t)$ is an arbitrary function. Utilizing Eq. (1) and defining the operator [7]

$$g = \sum_{pq} v_{pq}(\partial / \partial A_{pq})$$

with

$$v_{pq} = \int_0^T \omega(t)[\delta A_{pq} / \delta C(t)] dt$$

leads to

$$\mathcal{H} = - \mathrm{Tr}[\exp(-iB) g^2 \exp(iA) + \exp(iB) g^2 \exp(-iA)], \qquad (12)$$

where it is understood that $\mathcal{H}$ is to be evaluated at one of the extrema corresponding to Eq. (10). The prior analysis shows that the various extrema are expressed in terms of a relationship between the eigenvalues of A and B, given by $a_i = b_i + n_i \pi$ with n_i being integers. Thus, each extremum matrix A^ℓ, $\ell = 1, 2, 3, \ldots$ has an associated set of integers $n^\ell = (n_1^\ell, n_2^\ell, \ldots, n_N^\ell)$ such that $A^\ell = B + D^\ell$ with $\langle \beta_i | D^\ell | \beta_j \rangle = \delta_{ij} \pi n_i^\ell$ being diagonal in the representation of the eigenstates of B. Noting that $g^\dagger = -g$, expressing the trace operation in Eq. (12) in terms of the eigenstates $\{|\beta_j\rangle\}$ and evaluating the expression for $\mathcal{H}^\ell$ at the ℓth extremum leads to

$$\mathcal{H}^\ell = 2 \sum_j \langle \beta_j | g^\dagger g | \beta_j \rangle (-1)^{n_j^\ell}. \qquad (13)$$

Analysis of Eq. (13), considering $\omega(t)$ to be an arbitrary function residing in g, shows that each extremum of the cost function J falls into one of three classes:

(i) minima $(\mathcal{H}^\ell > 0)$:$\{n_j^\ell\}$ all even integers.

(ii) maxima $(\mathcal{H}^\ell < 0)$:$\{n_j^\ell\}$ all odd integers.

(iii) saddle points ($\mathcal{H}^\ell$ indefinite);

$\{n_j^\ell\}$ mixed even and odd integers. (14)

Classes (i) and (ii), respectively, correspond to $U = W$ and $U = -W$ with class (i) being a perfect control. All the remaining extrema fall into class (iii) as saddle points. Thus, in seeking to minimize J, there are no false extrema to get trapped in although many saddle points may be encountered on the way. The key results of this paper in Eqs. (10) and (14) have several fundamental implications for theoretical design or laboratory discovery of optimal controls $C(t)$ to create useful unitary transformations. These points will be discussed in the remainder of the paper.

As a simple example of the analysis above, consider the target unitary transformation $W = \mathrm{CNOT}$, (CNOT represents the controlled-NOT operation), where

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \qquad (15)$$

In its diagonal basis, $W = Q e^{iE_B} Q^\dagger$, where

$$E_B = \begin{pmatrix} e^{ib_1} & 0 & 0 & 0 \\ 0 & e^{ib_2} & 0 & 0 \\ 0 & 0 & e^{ib_3} & 0 \\ 0 & 0 & 0 & e^{ib_4} \end{pmatrix}, \qquad (16)$$

with $e^{ib_1} = -1$ and $e^{ib_2} = e^{ib_3} = e^{ib_4} = 1$. For J to take on a critical value, it is necessary and sufficient that $U = Q e^{iE_A} Q^\dagger$, where

$$E_A = \begin{pmatrix} e^{ia_1} & 0 & 0 & 0 \\ 0 & e^{ia_2} & 0 & 0 \\ 0 & 0 & e^{ia_3} & 0 \\ 0 & 0 & 0 & e^{ia_4} \end{pmatrix}, \qquad (17)$$

where $e^{ia_k} = \pm 1$ for $k = 1, 2, 3, 4$. In this case for Eq. (15) Q is not unique due to the degenerate eigenvalues for E_B. What is relevant here is the spectrum of A in Eq. (17) to establish the landscape topology. From Eq. (13), we may verify that for $e^{ia_1} = -1$, $e^{ia_2} = e^{ia_3} = e^{ia_4} = 1$, we attain a global maximum $J = 16$ with all Hessian eigenvalues negative, and for $e^{ia_1} = 1$, $e^{ia_2} = e^{ia_3} = e^{ia_4} = -1$ a global minimum $J = 0$ with all Hessian eigenvalues positive. In all other cases we obtain saddle points with landscape heights $J = 4$, 12 and a mixture of positive and negative Hessian eigenvalues.

The result in Eq. (10) was also verified for a number of arbitrary target transformations W of dimension $N = 4$ and 8. As controllability is assumed, we may sidestep the actual solution of Schrödinger's equation and numerically solve Eq. (4) treating the elements of A as variables to be searched over with a gradient algorithm. The particular local extrema found depended on the initial trial matrix A, but in all cases only the stable minima were found corresponding to $J = 0$. Similarly, seeking maximization of J only lead to the solutions at $J = 4N$. As the saddle points are unstable, the simple gradient algorithm did not directly discover them, or get trapped at them. Although no saddle points were found by direct searching, it was verified by numerical tests that the extrema satisfying the criterion in case (iii) of Eq. (13) were saddle points. The presence of local saddle point extrema $J > 0$ implies that global search algorithms [5,8] (e.g., genetic-type algorithms) could be advantageous for seeking at least one control satisfying $J = 0$, although local algorithms could also work as indicated in the simulations above. Fortunately, the suboptimal extrema are not local minima, which eliminates the prospect of being trapped. The connection between the critical topology of the landscape and the efficacy of various types of global and local search algorithms for finding effective controls for creating desired unitary transformations is explored in a separate work [15].

V. CONSEQUENCES FOR LABORATORY GENERATION OF UNITARY TRANSFORMATIONS

The computational design of optimal controls to generate unitary transformations can be very expensive [9] for large values of N, especially when using global search algorithms. However, in the laboratory, the high duty cycle of realizations employing shaped laser pulses as controls [10] may permit a thorough search for a control $C(t)$ that gives the absolute minimum value in the control landscape. The scaling of the required search effort with the dimension of the unitary matrix N is a concern. Many proposed practical quantum information systems may involve large numbers of qubits $n \sim 10^3$ with the total Hilbert space dimension being $\mathcal{N}=2^n$. One hope is that a quantum information machine of Hilbert space dimension $\mathcal{N}$ may be constructed by linking together a number N_s of modular subunits with each having a small number n_i of qubits [11], where $\Sigma_i n_i = n$. In this scenario, after learning how to control each of the subunits, they would be coupled together to operate in a cooperative fashion thereby producing a total functioning quantum information machine. Optimal searching for desired unitary transformations on the Hilbert space for the ith subunit would encounter $2^{n_i}+1$ landscape extrema values. The actual operational scenario in the laboratory will likely depend on the physical realization and many other factors. Regardless of these issues, the analysis here reveals another important feature with regard to the control search landscape to be traversed in making the operational decisions.

In the laboratory, there will always be nonidealities to consider, including decoherence, noise in the controls, observation errors, etc. In the desirable regime of weak noise and decoherence, the landscape structure remains the same, but viewed at lower resolution [12]. Besides these latter points, the inevitable constraints on practically constructing control fields [10] can alter the extrema values in Eq. (10) and introduce other local extrema besides those identified under ideal conditions. This comment may be understood by considering a situation where the field constraints only allow for a restricted exploration of any local region of A (e.g., around a particular exact solution $J=0$); the local character of the landscape could appear distorted by the field restrictions. If possible, the control field should be left with sufficient flexibility to freely roam over the landscape of J, hopefully including the domain of at least one solution $J=0$.

The critical topology of the function described in Eq. (1), formally a mapping $J: \mathrm{U}(N) \rightarrow \mathbb{R}$, could have been alternately analyzed from the perspective of the geometry of the unitary group $\mathrm{U}(N)$, and such approaches have been adopted previously [14]. Although these methods have some mathematical advantages, the underlying physical structure of the problem becomes obscured. In the present work, the optimization analysis is carried out in a way that has a more facile physical interpretation. Using the degrees of freedom in matrices A and B produces a physically transparent picture of the control process being one of optimization over the *action variables* represented in A. The latter Hermitian matrix can be interpreted as the collective, time-ordered dynamics generated by the Hamiltonian. The influence of nonidealities including control constraints on the critical structure of the landscape, are most naturally stated in terms of the Hamiltonian, to which the A matrix has a clear qualitative connection.

VI. CONCLUSION

The key to the analysis in this paper rests on identifying the underlying kinematics [13] of the general optimization problem, and thereby separating the Hamiltonian-specific aspects of any particular realization. This kinematic focus was facilitated by using $U=\exp(iA)$ and assuming full controllability to permit analysis of Eq. (4) rather than Eq. (2). Equation (4) was amenable to a thorough analysis of its extrema and hence the optimal control landscape of J. The kinematic reduction yields the rather surprising conclusion that all optimal control problems for U of the same dimension N and target W have equivalent landscapes with respect to A, no matter what form the Hamiltonian takes, provided that the system is controllable [4]. Differences in search behavior over the landscape can in practice arise due to the particular nature of the Hamiltonian and the control. Nevertheless, the optimal control landscape topology identified in this work is universal.

ACKNOWLEDGMENTS

The authors acknowledge valuable discussions with Robert Kosut and Ian Walmsley, and support from the NSF and DARPA. M.H. acknowledges the support of NDSEG.

[1] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, Science **288**, 824 (2000).

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York 2000).

[3] C. M. Tesch, L. Kurtz, R. de Vivie-Riedle, Chem. Phys. Lett. **342**, 633 (2001); J. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003).

[4] V. Ramakrishna, M. V. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, Phys. Rev. A **51**, 960 (1995). The assumption of full controllability is actually stronger than necessary in the analysis of the paper, as $\{\delta A_{pq}/\delta C(t)\}$ is only required to be linearly independent. Full controllability assures this prospect as well as the independence of all higher-order functional derivatives.

[5] R. S. Judson and H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).

[6] G. H. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins University Press, Maryland, 1983).

[7] The derivative with respect to A_{pq} is understood to be taken separately with the actual variables $\mathrm{Re}\, A_{pq}$ and $\mathrm{Im}\, A_{pq}$.

[8] D. Zeidler, S. Frey, K. L. Kompa, and M. Motzkus, Phys. Rev.

A **64**, 023420 (2001); B. J. Pearson, J. L. White, T. C. Weinacht, and P. H. Bucksbaum, *ibid.* **63**, 063412 (2001).
[9] J. Palao, and R. Kosloff, Phys. Rev. Lett. **89**, 188301 (2002).
[10] A. M. Weiner, Rev. Sci. Instrum. **71**, 1929 (2000).
[11] M. Oskin and F. Chong, and I. Chuang, IEEE Comp. **35**, 79 (2002).
[12] H. Rabitz, M. Hsieh, and C. Rosenthal, Science **303**, 1998 (2004).
[13] The word *kinematics* is used here in the sense that Eq. (4) does not depend on the evolution with respect to any particular Hamiltonian.
[14] T. Frankel in, *Differential and Combinatorial Topology*, edited by S. S. Cairns (Princeton University Press, Princeton 1965); L. Geissinger, Lin. and Mult. Alg. **5**, 321(1978); M. Shayman (unpublished).
[15] K. Moore, M. Hsieh, and H. Rabitz (unpublished).