Searching for an optimal control in the presence of saddles on the quantum-mechanical observable landscape

Gregory Riviello,[1] Re-Bing Wu,[2] Qiuyang Sun,[1] and Herschel Rabitz[1]

[1]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*

[2]*Department of Automation, Tsinghua University, and Center for Quantum Information Science and Technology, TNlist, Beijing 100084, China*

The broad success of theoretical and experimental quantum optimal control is intimately connected to the topology of the underlying control landscape. For several common quantum control goals, including the maximization of an observable expectation value, the landscape has been shown to lack local optima if three assumptions are satisfied: (i) the quantum system is controllable, (ii) the Jacobian of the map from the control field to the evolution operator is full rank, and (iii) the control field is not constrained. In the case of the observable objective, this favorable analysis shows that the associated landscape also contains saddles, i.e., critical points that are not local suboptimal extrema. In this paper, we investigate whether the presence of these saddles affects the trajectories of gradient-based searches for an optimal control. We show through simulations that both the detailed topology of the control landscape and the parameters of the system Hamiltonian influence whether the searches are attracted to a saddle. For some circumstances with a special initial state and target observable, optimizations may approach a saddle very closely, reducing the efficiency of the gradient algorithm. Encounters with such attractive saddles are found to be quite rare. Neither the presence of a large number of saddles on the control landscape nor a large number of system states increases the likelihood that a search will closely approach a saddle. Even for applications that encounter a saddle, well-designed gradient searches with carefully chosen algorithmic parameters will readily locate optimal controls.

I. INTRODUCTION

The last two decades have seen a significant expansion of the boundaries of quantum optimal control experiments (OCEs) due to technological advances in experimental resources, especially femtosecond lasers and pulse-shaping capabilities [1–12]. OCEs have been successfully performed for a wide range of goals, including the control of molecular vibrational [13–20] and electronic states [21–29], the generation and coherent manipulation of x rays [30–34], the control of decoherence processes [35,36], the selective cleavage and formation of chemical bonds [37–43], the manipulation of energy flow in macromolecular complexes [44–47], and the control of photoisomerization reactions [48–52]. Optimal control theory (OCT) [7,9,53–56] has provided insights into the coherent control of a variety of quantum phenomena, such as electron transfer [57,58], molecular photoisomerization [59–62] and photodissociation [63–67], the manipulation of trapped Bose-Einstein condensates [68–70], strong-field ionization [71], quantum-information processing [72–94], and spin squeezing in atomic ensembles [95,96].

The primary goal of OCEs and OCT simulations is to find a control $\varepsilon(t)$ that yields the global maximum or minimum value of a cost functional $J = J[\varepsilon(t)]$. This cost functional represents control objectives such as the distance between the unitary evolution operator and a target unitary transformation, the probability of a transition between two states, or the expectation value of an observable [7]. Several recent studies [97–99] strongly indicate that the success of numerous OCEs and OCT simulations is related to the favorable topology of the quantum control landscape defined by the functional dependence of J on $\varepsilon(t)$ [7,9,100]. In particular, it has been shown that the control landscape lacks local optima (referred

to as traps) if three conditions are satisfied: (i) the quantum system is *controllable*; i.e., any unitary evolution operator can be produced by some admissible control field beyond some finite time; (ii) the Jacobian matrix mapping the control field $\varepsilon(t)$ to the final-time evolution operator $U(T,0)$ is of full rank everywhere on the landscape; (iii) there are no constraints on the control field [98,101–108]. The absence of local suboptimal extrema is of central importance to optimization; numerical studies have described the appearance of local traps on the control landscape due to the violation of assumption (i) [109] and shown that the violation of assumption (ii) [110] can, in special cases, prevent a gradient search from identifying globally optimal controls. Recent work [111] has shown that assumptions (i) and (ii) are *almost always* satisfied. Thus, the satisfaction of assumption (iii) (which depends in practice on access to adequate system-specific control resources) is generally the key criterion that determines whether OCEs or OCT searches will optimize successfully, especially with a local gradient-based algorithm. In this work, we assume that assumptions (i), (ii), and (iii) are satisfied and that the control landscape lacks local optima; this behavior was confirmed by the success of *all* simulations.

Even when the three assumptions are satisfied, however, the control landscapes for the unitary and observable objectives both contain suboptimal critical points. These critical points are saddles rather than local extrema, and cannot in principle trap a gradient-based search. However, gradient-based methods typically converge more slowly when they come near any such critical point. A prior numerical study of the unitary control objective indicated that saddles have little effect on gradient-based searches [112]. In this work, therefore, we focus on observable control, for which the landscape may have

a much larger number of saddles. Recent OCEs performed on a two-spin system located saddles on the observable control landscape at the predicted objective values and of the right character [113], providing empirical support for the theoretical analysis.

The trajectory of a gradient search is influenced by both the landscape topology (which is fully defined by the initial state and target observable) and the local, nontopological geometry of the landscape (which depends on those two operators as well as the form of the Hamiltonian and the nature of the initial control field). We perform a large number of numerical OCT searches on a variety of control problems in order to identify physical parameters or characteristics that determine whether an optimal search will approach a saddle closely during an optimization. Using a specially designed metric [114], we quantify the attractiveness of saddles and measure their influence on the efficiency of seeking optimal controls. The present work considers gradient-based simulations, which can be very sensitive to saddles. In the laboratory, it is more common to employ stochastic algorithms, but the presence of a high density of attractive saddles could nonetheless be a challenge to optimization. The findings in the present work are therefore relevant for effective performance in OCEs.

The remainder of the paper is organized as follows: Section II discusses the theoretical basis for the classification of critical points, as well as the observable objective and the topology of the corresponding control landscape. Section III describes the numerical methods employed in this work and the metric used to evaluate the effects of saddles during a gradient-based search. In Sec. IV we examine the factors that cause landscape saddles to influence searches for optimal controls. Our concluding remarks are given in Sec. V.

II. BACKGROUND AND LANDSCAPE ANALYSIS

A. Background

The control illustrations in this paper involve closed N-level quantum systems with Hamiltonians of the form

$$H(t) = H_0 - \mu \varepsilon(t), \qquad (1)$$

within the electric dipole approximation. H_0 is the field-free diagonal operator, the control field $\varepsilon(t)$ is a real-valued function of time defined on the interval $[0,T]$, and μ is the dipole operator that couples the system to the field. In the Schrödinger picture, the state of the system at a time t is described by the density matrix $\rho(t) = U(t)\rho_0 U^\dagger(t)$, where $\rho_0 \equiv \rho(0)$ is the initial density matrix and $U(t) \equiv U(t,0)$ is the propagator or evolution operator. The propagator satisfies the Schrödinger equation:

$$i\hbar \frac{d}{dt} U(t) = H(t)U(t), \quad U(0) = \mathbb{I}, \qquad (2)$$

where $\mathbb{I}$ is the N-dimensional identity operator. In the present work, we only consider evolution-operator controllable systems [9,54], i.e., systems for which any unitary operator W is the solution of the Schrödinger equation (2) at sufficiently long time T with some control field $\varepsilon(t)$. In the absence of controllability, it has been shown that the control landscape may contain traps [109].

The topology of a quantum control landscape is determined by characterizing its *critical points*, where

$$\frac{\delta J}{\delta \varepsilon(t)} = 0, \quad \forall t \in [0,T]. \qquad (3)$$

Critical points can be classified as global extrema, local extrema, or saddles, according to the properties of second- and higher-order functional derivatives of J with respect to the control field [7,100]. For example, the Hessian matrix,

$$\mathsf{H}(t,t') = \frac{\delta^2 J}{\delta \varepsilon(t)\delta \varepsilon(t')},$$

describes the local curvature near a critical point. At a saddle, the Hessian has both positive and negative character. The existence of landscape saddles has practical significance for OCT optimizations, since their presence may influence searches with a gradient algorithm [110,115] or even hinder the convergence efficiency of global stochastic algorithms [116]. The topic assessed in this paper is the role of saddles in seeking optimal controls, as reflected in the performance of a gradient-based algorithm which was chosen due to its sensitivity to landscape saddle features.

The landscape analysis for the objective J can be performed using either the *dynamic formulation*, in which the control landscape $J = J[\varepsilon(t)]$ is defined on the L^2 space of control fields, or the *kinematic formulation*, in which the control landscape $J = J(U_T)$ is defined on the unitary group U(N). In order to clarify the relationship between these two formulations, we partition the relationship between J and the control field $\varepsilon(t)$ by representing J as a function of the final-time evolution operator $U_T \equiv U(T)$, and U_T in turn as a functional of the control field, i.e., $J = J(U_T)$ and $U_T = U_T[\varepsilon(t)]$. Using the chain rule, Eq. (3) can be rewritten as

$$\frac{\delta J}{\delta \varepsilon(t)} = \left\langle \nabla J(U_T), \frac{\delta U_T}{\delta \varepsilon(t)} \right\rangle = 0, \quad \forall t \in [0,T], \qquad (4)$$

where $\langle \cdot, \cdot \rangle$ is the Hilbert-Schmidt inner product, $\nabla J(U_T)$ is the gradient of J with respect to U_T, and the Jacobian matrix $\delta U_T / \delta \varepsilon(t)$ is the first-order functional derivative of U_T with respect to the control field. Adopting satisfaction of assumption (ii), to which we referred in Sec. I, leads to the conclusion that Eq. (3) is equivalent to the kinematic result,

$$\nabla J(U_T) = 0. \qquad (5)$$

Therefore, the dynamic and kinematic perspectives yield the same landscape critical point specifications.

B. Formulation and landscape topology of the control objective

The OCT simulations in this work consider the goal of maximizing the expectation value of a Hermitian quantum observable θ at time T:

$$J = \langle \theta(T) \rangle = \mathrm{Tr}(U_T \rho_0 U_T^\dagger \theta). \qquad (6)$$

In order to fully describe the landscape topology of this objective from the kinematic perspective $J(U_T)$, the multiplicities of the eigenvalues of ρ_0 and θ must also be specified. Consider that ρ_0 has r distinct eigenvalues $p_1 > p_2 > \ldots > p_r$ with corresponding multiplicities $a_1, a_2, \ldots, a_r$ and that θ has q distinct eigenvalues $o_1 > o_2 > \ldots > o_q$ with corresponding

TABLE I. The contingency table C^i, which describes an alignment between the distinct eigenvalues of ρ_0 and θ corresponding to the critical submanifold M^i. The column and row sums of C^i are $a_1, \ldots, a_r$ and $b_1, \ldots, b_q$, respectively.

	a_1	$\cdots$	a_r
b_1	c_{11}^i	$\cdots$	c_{1r}^i
$\vdots$	$\vdots$	$\ddots$	$\vdots$
b_q	c_{q1}^i	$\cdots$	c_{qr}^i

multiplicities $b_1, b_2, \ldots, b_q$, where $q, r \leqslant N$. It has been demonstrated that ρ_0 and θ can always be treated as diagonal in the eigenbasis of H_0 and with their eigenvalues sorted in descending order, i.e.,

$$\rho_0 = \mathrm{diag}\{p_1, \ldots, p_1; \ldots; p_r, \ldots p_r\},$$
$$\theta = \mathrm{diag}\{o_1, \ldots, o_1; \ldots; o_q, \ldots o_q\}, \tag{7}$$

with no loss of generality in the landscape analysis [105]. In the kinematic formulation, it has also been shown that the sufficient and necessary condition for U_T to be a critical point of the landscape is that the final-time density matrix $\rho(T) = U_T \rho_0 U_T^\dagger$ commutes with the target observable θ [98,105,106], i.e.,

$$[\rho(T), \theta] = 0. \tag{8}$$

With ρ_0 and θ in the form of Eq. (7), the condition in Eq. (8) is satisfied if and only if the unitary matrix U_T lies in the double coset

$$U_T = P \Pi Q^\dagger, \quad P \in \mathcal{U}(\mathbf{b}), \quad Q \in \mathcal{U}(\mathbf{a}) \tag{9}$$

of some N-dimensional permutation matrix Π, where $\mathcal{U}(\mathbf{a}) = \mathcal{U}(a_1) \times \ldots \times \mathcal{U}(a_r)$ is the product of unitary groups of dimension $a_1, \ldots, a_r$ and $\mathcal{U}(\mathbf{b}) = \mathcal{U}(b_1) \times \ldots \times \mathcal{U}(b_q)$ is the product of unitary groups of dimension $b_1, \ldots, b_q$ [105]. In general, however, Π is not unique and the evolution operators U_T that satisfy Eq. (9) are not permutation matrices.

By substituting Eq. (9) into Eq. (6), the objective functional J at a critical point can be rewritten as

$$J_{\mathrm{crit}} = \mathrm{Tr}(P \Pi Q^\dagger \rho_0 Q \Pi^\dagger P^\dagger \theta) = \mathrm{Tr}(\Pi \rho_0 \Pi^\dagger \theta). \tag{10}$$

Thus, critical points on the observable objective landscape only exist at a finite number of discrete values of J; these values only depend on the eigenvalues of ρ_0 and θ, not on the control field or Hamiltonian [98,105]. More specifically, each critical J value corresponds to the sum of the product of the permuted eigenvalues of ρ_0 with the eigenvalues of θ. Further characterization of the critical points of J was accomplished via the *contingency table* method described in Ref. [105]. The contingency table C is a $q \times r$ matrix whose nonnegative integer-valued elements $\{c_{jk}\}$, the so-called *overlap numbers*, are the number of positions on the diagonals of θ and $\Pi \rho_0 \Pi^\dagger$ where the distinct eigenvalues o_j and p_k, respectively, both appear. The column and row sums of C are $a_1, \ldots, a_r$ and $b_1, \ldots, b_q$, respectively.

A specific contingency table C^i is shown in Table I. The critical points of the landscape $J(U_T)$ that correspond to C^i collectively comprise a *critical submanifold* of the control

landscape, which we denote as M^i. All critical points in M^i share the same objective value,

$$J_i = \sum_{j,k=1}^{q,r} c_{jk}^i o_j p_k, \tag{11}$$

although two critical submanifolds may have identical objective values. We will denote the objective values corresponding to the global maximum and minimum of the landscape as $J_{\max}$ and $J_{\min}$, respectively. If both ρ and θ are full rank, then each permutation Π generates a distinct contingency table and thus there are $N!$ critical submanifolds on the landscape. In this case, the critical submanifolds are disjoint N-tori, and analysis of the Hessian spectrum shows that two of them are the global maximum and global minimum of J while the remainder are saddles [98,105]. Graphically, we can visualize the ith critical submanifold as an infinitely thin "pancake" of some shape in the function space of controls at its corresponding saddle value $J_i[\varepsilon(t)]$, where the gradient $\delta J_i / \delta \varepsilon(t) = 0$ and the Hessian $\mathsf{H}(t, t')$ has an indefinite nonzero spectrum and an infinite null space. If any eigenvalues of ρ or θ are degenerate, then the same contingency table can be produced from multiple permutations Π, and the critical submanifold corresponding to that contingency table results from the merging of several N-tori. In this degenerate case, the landscape has fewer than $N! - 2$ saddles. The fewest landscape critical submanifolds arise when $\rho_0 = |i\rangle\langle i|$ and $\theta = |f\rangle\langle f|$, i.e., when ρ_0 and θ are projectors onto the pure states $|i\rangle$ and $|f\rangle$, respectively. This special case of the observable objective is called the *state-transition objective* and corresponds to maximizing the probability of a transition from $|i\rangle$ to $|f\rangle$. The landscape for the state-transition control contains no saddles, so such problems are not considered in this paper; see Ref. [117] for a numerical study of state-transition landscapes.

For a particular ρ_0 and θ, each permutation Π leads to the construction of a (not necessarily unique) contingency table C, as described above. By repeating this process, all of the contingency tables for the landscape $J(U_T)$ can be determined, and the corresponding objective values indicate whether each table corresponds to the global maximum, the global minimum, or a saddle. As an example, consider a 4-level control problem with $\theta = \mathrm{diag}\{0.5, 0.2, 0.2, 0.1\}$ and $\rho = \mathrm{diag}\{0, 0, 0, 1\}$. θ has three distinct eigenvalues, $o_1 = 0.5$, $o_2 = 0.2$, and $o_3 = 0.1$, with multiplicities $b_1 = 1$, $b_2 = 2$, and $b_3 = 1$, respectively. ρ_0 has two distinct eigenvalues, $p_1 = 1$ and $p_2 = 0$, with multiplicities $a_1 = 1$ and $a_2 = 3$, respectively. Therefore, the contingency table corresponding to each critical submanifold is a 3×2 matrix with row sums $\{1, 2, 1\}$ and column sums $\{1, 3\}$. Under the permutation

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$\Pi \rho_0 \Pi^\dagger = \mathrm{diag}\{1, 0, 0, 0\}$. The overlap numbers for the contingency table corresponding to this permutation are determined by comparing the diagonal of the permuted density matrix with the diagonal of the observable θ. $c_{11} = 1$ because the distinct eigenvalues $o_1 = 1$ and $p_1 = 0.5$ simultaneously appear at the first position (and no other positions) on the diagonals of

θ and $\Pi \rho_0 \Pi^\dagger$, respectively. Similarly, the remaining overlap numbers are determined to be $c_{21} = c_{31} = c_{12} = 0$, $c_{22} = 2$, and $c_{32} = 1$. If this process is repeated for all four-dimensional permutation matrices, then three distinct contingency tables are identified:

$$C^1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 1 \end{pmatrix}, \quad C^2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad C^3 = \begin{pmatrix} 0 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

Using Eq. (11), the objective values for each critical submanifold are calculated to be $J_1 = 0.5$, $J_2 = 0.2$, and $J_3 = 0.1$. Therefore, the contingency tables C^1 and C^3 correspond to the global maximum and minimum of the landscape, respectively, while C^2 corresponds to a saddle submanifold. The enumeration of these critical submanifolds fully describes the landscape topology for the observable control problem.

III. METHODOLOGY

A. Optimal control procedure

In this work, a gradient-based method will be employed to investigate local landscape saddle features because this procedure is "myopic"; i.e., each step taken during the search is dictated by the local geometry of the control landscape at the current control field and thus is particularly sensitive to the presence of saddles. Each search is parametrized by the dimensionless index $s \geq 0$, which denotes the changes made to the field in the course of the optimization through the notation $\varepsilon(s,t)$. The search trajectory is generated by solving the initial value problem

$$\frac{\partial \varepsilon(s,t)}{\partial s} = \gamma \frac{\delta J[\varepsilon(s,t)]}{\delta \varepsilon(s,t)}, \quad \varepsilon(0,t) \equiv \varepsilon_0(t), \qquad (12)$$

where the initial field is $\varepsilon_0(t)$, and the step size γ is a positive constant. The functional derivative $\delta J / \delta \varepsilon(s,t)$ that appears in Eq. (12) is calculated using the chain rule [as in Eq. (4)] along with the previously derived [98] relation

$$\frac{\delta U_T}{\delta \varepsilon(t)} = \frac{i}{\hbar} U_T U^\dagger(t) \mu U(t).$$

The result is [98,102,107,115]

$$\frac{\delta J}{\delta \varepsilon(t)} = \frac{2}{\hbar} \text{ImTr}[U_T^\dagger \theta U_T \rho_0 U^\dagger(t) \mu U(t)]. \qquad (13)$$

We solve Eq. (12) numerically using the MATLAB routine ODE45, a fourth-order Runge-Kutta integrator with a variable step size (i.e., it determines γ at each iteration) [118]. ODE45 requires that an absolute error tolerance τ be specified, and we use the conservative value $\tau = 10^{-8}$ unless otherwise stated. The *search effort*, defined as the number of iterations required for convergence, is an important measure of algorithmic efficiency. For the goal of maximizing the objective functional, the simulation is considered to have successfully converged when the search arrives at a control field $\varepsilon(s_f,t)$ that corresponds to an objective value $J \geq [J_{\max} - 0.001 \times (J_{\max} - J_{\min})]$.

In this paper, $\varepsilon(t)$ was discretized over L evenly spaced intervals,

$$\varepsilon(t) = \{\varepsilon_l | t \in (t_{l-1}, t_l]\}_{l=1}^L, \qquad (14)$$

where $t_l = l \Delta t$ and $\Delta t = T / L$. The overall evolution operator $U(t_l) \equiv U(t_l, 0)$ is a product of incremental evolution operators,

$$U(t_l, t_{l-1}) = \exp \left[-\frac{i}{\hbar} (H_0 - \mu \varepsilon_l) \Delta t \right],$$

$$U(t_l) = U(t_l, t_{l-1}) \cdots U(t_2, t_1) U(t_1, 0), \qquad (15)$$

where the final-time evolution $U_T = U(t_L)$. The control variables are the L real, independently addressable field values $\{\varepsilon_l\}$, which can generate arbitrary pulse shapes as long as L is sufficiently large. The lth value of the initial field has the parametrized form

$$\varepsilon_l(0) \equiv \quad \varepsilon_0(t_l) = A(t_l) \sum_{m=1}^M a_m \cos(\omega_m t_l), \qquad (16)$$

where $A(t_l) = A_0 \exp[-(t_l - T/2)^2 / (2\eta^2)]$ is the Gaussian envelope function. The width of the envelope is specified by $\eta = T/10$, and ensures that $\varepsilon_0(t) \approx 0$ at $t = 0$ and $t = T$. The $M = 20$ frequencies $\{\omega_m\}$ are randomly selected from a uniform distribution on $[\omega_{\min}, \omega_{\max}]$, where $\omega_{\min}$ and $\omega_{\max}$ are the smallest and largest transition frequencies in H_0, respectively. The amplitudes $\{a_m\}$ are randomly selected from a uniform distribution on $[0,1]$. The normalization constant A_0 is chosen so that the fluence, $F = \int_0^T \varepsilon^2(t) dt$, of the initial field $\varepsilon_0(t)$ has the value F_0.

After the field values $\varepsilon_l(s)$ are set in Eq. (16) (i.e., for $s > 0$), they are allowed to change according to the discrete version of Eq. (12):

$$\frac{\partial \varepsilon_l(s)}{\partial s} = \gamma \frac{\delta J}{\delta \varepsilon_l(s)} \simeq \gamma \Delta t \frac{\partial J}{\partial \varepsilon(t_l)}. \qquad (17)$$

B. Critical distance metric

The effect of saddles on a gradient search depends in part on how closely the search trajectory approaches the saddle submanifold. We quantify this distance using the unitless *critical distance metric* $D^i(U_T)$, which is a measure of the distance between a control U_T and a particular critical submanifold M^i on the kinematic observable landscape [114]. Suppose that ρ_0 and θ are represented as diagonal matrices as in Eq. (7), with their eigenvalues sorted in descending order. U_T can be divided into $q \times r$ rectangular blocks U_{jk}, each of dimension $b_j \times a_k$:

$$U_T = \begin{pmatrix} U_{11} & \cdots & U_{1r} \\ \vdots & \ddots & \vdots \\ U_{q1} & \cdots & U_{qr} \end{pmatrix}. \qquad (18)$$

U_{jk}, which is generally not unitary, contains the elements of U_T that correspond to the alignment of o_j and p_k. Let the singular value decomposition of U_{jk} be

$$U_{jk} = X_{jk} S_{jk} Y_{jk}^\dagger, \qquad (19)$$

where the columns of the unitary matrices X_{jk} and Y_{jk} are the left and right singular vectors of U_{jk}, respectively, and S_{jk} is a diagonal matrix containing the singular values σ_{jkl} of U_{jk}

sorted in descending order. According to Theorem 1 of [114], U_T belongs to the critical submanifold M^i of the observable landscape in Eq. (6) if and only if the first c^i_{jk} singular values σ_{jkl} of each block U_{jk} are equal to 1 and the remaining singular values of each block are equal to 0. Thus, the critical distance metric is defined by comparing each singular value of U_{jk} to either 1 or 0, as appropriate [114]:

$$
\begin{aligned}
D^i(U_T) &= \sum_{j,k=1}^{q,r} \left[\sum_{l \leqslant c^i_{jk}} (1 - \sigma_{jkl})^2 + \sum_{l > c^i_{jk}} \sigma^2_{jkl} \right] \\
&= 2 \sum_{j,k=1}^{q,r} \sum_{l \leqslant c^i_{jk}} (1 - \sigma_{jkl}).
\end{aligned}
\tag{20}
$$

$D^i(U_T) = 0$ if and only if U_T belongs to the critical submanifold M^i corresponding to the contingency table C^i. For a particular control problem, the range of D^i depends on the specific degeneracies of ρ_0 and θ, and the maximum possible distance from each critical submanifold is not necessarily the same. For all control problems, however, the distance between any two critical submanifolds labeled i and i', defined as

$$
D^{i \to i'} = \sum_{j,k=1}^{q,r} \left| c^i_{jk} - c^{i'}_{jk} \right|,
\tag{21}
$$

cannot exceed $2N$.

IV. EFFECT OF SADDLES ON GRADIENT OPTIMIZATIONS

Previous simulations of observable control problems with gradient-based algorithms have regularly reached the landscape maximum value [110,119]. However, searches may converge more slowly while coming close to saddles, increasing the search effort. In this section, we investigate the practical effects of saddles on OCT simulations in a variety of control problems. Many optimization parameters, details of the landscape topology, and features of the Hamiltonian affect whether searches approach saddles closely, so this paper cannot comprehensively address all the relevant aspects of any particular control problem. However, we discuss several key parameters that significantly influence saddle attraction. Each parameter is studied, as independently of the others as possible, in order to evaluate its individual role.

All simulations in this section are performed on one of two quantum systems, either rigid-rotor-like,

$$
H_0 = \sum_{j=0}^{N-1} j(j+1)|j\rangle\langle j|,
\tag{22}
$$

or an anharmonic oscillator,

$$
H_0 = \sum_{j=0}^{N-1} \left[\kappa \left(j + \frac{1}{2} \right) - \frac{\kappa^2}{\lambda} \left(j + \frac{1}{2} \right)^2 \right] |j\rangle,
\tag{23}
$$

where $\kappa = 2$ and $\lambda = 320$. In both cases, the dipole matrix is

$$
\mu = \sum_{j \neq k}^{N-1} \frac{d^{|j-k|}}{d} |j\rangle\langle k|,
\tag{24}
$$

where the parameter $d \geqslant 0$. For the purposes of this paper, the field-free Hamiltonians H_0 in Eqs. (22) and (23) were chosen to illustrate two extreme cases of increasing and decreasing energy level spacing, respectively. Correspondingly, the freedom in choosing d in the dipole allows for sampling different degrees of coupling structure in μ.

A. Degeneracy of the initial state and the target observable

When the three assumptions described in Sec. I are satisfied, the topology of the observable landscape is fully determined by the number and multiplicities of the distinct eigenvalues of the initial density matrix ρ_0 and the target observable θ. Each permutation Π of the eigenvalues of ρ_0 with respect to those of θ corresponds to a particular critical submanifold, as shown in Eq. (10). The multiplicities of the two sets of eigenvalues determine how many distinct permutations coincide with the global maximum, the global minimum, and each saddle [105]. If more than one permutation corresponds to a critical submanifold M^i, then evolution operators U_T that coincide with M^i can take a wider range of forms. Here, we perform optimizations on five related control problems in order to determine whether this additional freedom in the form of the critical U_T influences the proximity of gradient searches to the saddles. Each search is performed on the rigid-rotor-like system from Eqs. (22) and (24), with $N = 8$ and $d = 0.2$. The final time is $T = 20$ and the control period [0,20] is divided into $L = 512$ intervals. The initial state is $\rho_0 = |0\rangle\langle 0|$; ρ_0 has two distinct eigenvalues $p_1 = 1$ and $p_0 = 0$, of multiplicities $a_1 = 1$ and $a_2 = 7$, respectively. Five target observables $\{\theta^m\}$ are considered:

$$
\begin{aligned}
\theta^1 &= \frac{4}{9}|6\rangle\langle 6| + \frac{5}{9}|7\rangle\langle 7|, \\
\theta^2 &= \sum_{j=5}^{6} \frac{4}{13}|j\rangle\langle j| + \frac{5}{13}|7\rangle\langle 7|, \\
\theta^3 &= \sum_{j=4}^{6} \frac{4}{17}|j\rangle\langle j| + \frac{5}{17}|7\rangle\langle 7|, \\
\theta^4 &= \sum_{j=3}^{6} \frac{4}{21}|j\rangle\langle j| + \frac{5}{21}|7\rangle\langle 7|, \\
\theta^5 &= \sum_{j=2}^{6} \frac{4}{25}|j\rangle\langle j| + \frac{5}{25}|7\rangle\langle 7|.
\end{aligned}
\tag{25}
$$

Each of the target observables θ^m have three distinct eigenvalues o^m_1, o^m_2, and o^m_3, with $o^m_3 = 0$ in each case. The eigenvalues o^m_1 and o^m_2 are different for each control problem but chosen to ensure that every observable has unit trace: for example, $o^2_2 = 4/13$ and $o^2_1 = 5/13$. In general, the multiplicities of the eigenvalues of θ^m are $b^m_1 = 1$, $b^m_2 = m$, and $b^m_3 = 7 - m$, respectively. The latter case of b^m_3 implies that the observable θ^m has $7 - m$ zero eigenvalues, which are associated with the system states not explicitly shown in Eq. (25).

As an example of using the method described in Sec. II B, the contingency tables for the control problem corresponding

TABLE II. Optimization results for several target observables θ^m. The multiplicity of the second-largest eigenvalue o_2^m of each observable is m. One hundred runs were performed for each observable.

Observable	Multiplicity of o_2^m	$\overline{D_{\min}^{\text{sadd}}}$	MSE
θ^1	1	3.35×10^{-1}	141
θ^2	2	2.34×10^{-2}	289
θ^3	3	8.12×10^{-3}	477
θ^4	4	7.36×10^{-4}	2745
θ^5	5	1.92×10^{-4}	9562

to the observable θ^5 were determined to be

$$C^{\max} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \\ 0 & 2 \end{pmatrix},$$

$$C^{\text{sadd}} = \begin{pmatrix} 0 & 1 \\ 1 & 4 \\ 0 & 2 \end{pmatrix}, \qquad (26)$$

$$C^{\min} = \begin{pmatrix} 0 & 1 \\ 0 & 5 \\ 1 & 1 \end{pmatrix},$$

with objective values $J_{\max} = 0.2$, $J_{\text{sadd}} = 0.16$, and $J_{\min} = 0$. Permutations that align the eigenvalue $p_1 = 1$ of ρ_0 with the eigenvalues $o_1^5 = 0.2$, $o_2^5 = 0.16$, and $o_3^5 = 0$ of θ^5 correspond to the global maximum, saddle, and global minimum of the landscape, respectively. The contingency tables for the other four problems were constructed in the same way. Each control landscape contains one saddle, since every target observable has the same number of distinct nonzero eigenvalues and each control problem has the same initial state. In general, for each problem, permutations that align the eigenvalue p_1 with the eigenvalues o_1^m, o_2^m, and o_3^m of θ^m correspond to the global maximum, saddle, and global minimum critical submanifolds, respectively.

For each of the five control problems, we performed one hundred optimization runs using the control procedure described in Sec. III A. Each run began at a different initial field $\varepsilon_0(t)$ as defined in Eq. (16), with fluence $F_0 = 10$. Every search converged successfully, and the results of these optimizations are reported in Table II. In addition, the distance to each of the critical submanifolds was calculated at every step of each optimization, using the distance metric described in Sec. III B. At an iteration of the gradient search denoted by the index s, the control field $\varepsilon(s,t)$ corresponds to an evolution operator $U_T(s)$ and in turn to a particular value $D^i[U_T(s)]$ of the critical distance metric for each critical submanifold M^i. The smallest value of D^i over the interval $0 \leqslant s \leqslant s_f$ (i.e., the shortest distance to the saddle manifold M^i at any point during a given search) was denoted as $D_{\min}^i$. We use the mean value of $\overline{D_{\min}^{\text{sadd}}}$ for a set of one hundred optimizations to measure how closely gradient-based searches approach a saddle, on average, for a given control problem. For each set of optimizations, the mean search effort (MSE), i.e., the mean number of iterations, is also reported.
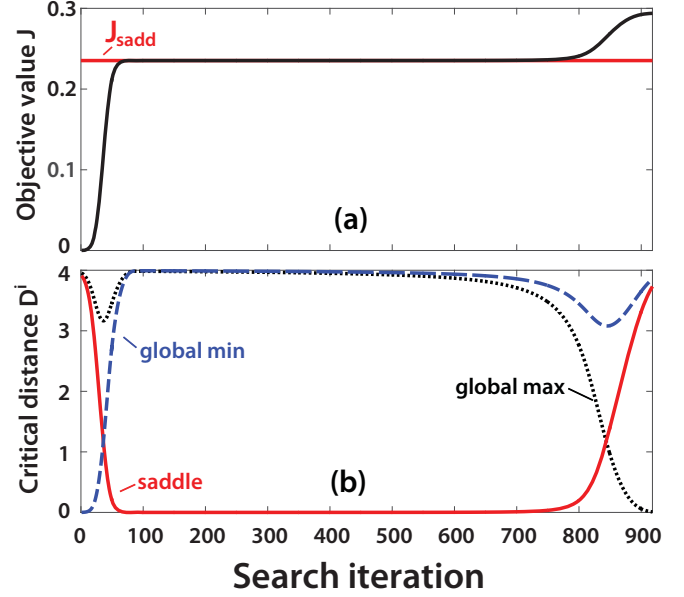

FIG. 1. (a) The objective value at each iteration for a particular optimization run involving the target observable θ^3 (black line) and the objective value of the saddle submanifold, J_{sadd} (red line), for this control problem. (b) The value of the critical distance metric at each iteration of the same search, for $D^{\max}$ (dotted black line), D^{sadd} (solid red line), and $D^{\min}$ (dashed blue line).

These simulations indicate that gradient-based optimizations for which the multiplicity of the observable eigenvalue o_2^m is larger involve a greater mean search effort and approach the saddle more closely. This trend is consistent with an expression for the dimension of critical submanifolds on the observable landscape obtained in Ref. [105]. Since the critical submanifold dimension does not take the system dynamics into account, its value is not predictive of the attractiveness of the saddle as measured in this paper. However, the submanifold dimension qualitatively matches the trend in Table II; the dimension of the saddle increases with the multiplicity of o_2^m. Figure 1, which illustrates the value of the objective J and the distance to each critical submanifold at each iteration of a particular optimization run corresponding to θ^3, shows that this increase in search effort results from a large number of iterations spent near the saddle submanifold.

Each of the control problems included in Table II has a control landscape with one saddle. The saddle corresponds to permutations Π that align the eigenvalues p_1 (of ρ_0) and o_2^m (of θ^m). The correlation between the multiplicity of o_2^m and the observed proximity to the saddle suggests that a broader range of critical unitary transformations U_T, made possible by this greater multiplicity, makes the saddle more attractive to a gradient search for these problems. In general, the simulations performed in this paper only encounter attractive saddles when ρ_0 and θ have respective eigenvalues that are highly degenerate and thereby have a strong influence on the nature of the saddle permutations Π. Importantly, we show later in Sec. IV C that optimizations of several control problem cases with a large number of saddles have a *significantly greater* mean distance of approach to any saddle than for the case with one saddle here. Thus, attractive saddles are expected to be rare in realistic control problems.

Other searches approached the saddle more closely than the example in Fig. 1; one optimization corresponding to the observable θ^5 required over 7×10^4 iterations to converge and reached a minimum distance of $D_{min}^{sadd} = 8.95 \times 10^{-7}$ from the saddle. Despite the numerical challenges presented by such runs (for which the magnitude of the gradient becomes very small) none of the searches failed. These results corroborate prior numerical studies, which concluded that the observable objective is amenable to gradient-based optimization when the landscape lacks local extrema. Figure 1(b) exhibits a notable feature of this optimization; after the search has passed its point of closest approach to the saddle submanifold and the objective value J has passed J_{sadd}, the distance to the global minimum briefly decreases just as the gradient ascent resumes. In general, the nonmonotonic behavior of the metric values D^{max}, D^{sadd}, and D^{min} may reflect the varied and possibly complex shape of the critical submanifolds themselves over the space of controls. The phenomenon in Fig. 1(b), however, was observed in each of the optimizations summarized in Table II and is therefore unlikely to depend on the particular gradient path taken to the top of the landscape. Instead, this behavior may be interpreted as a reflection of the relationship between the critical submanifolds for this set of control problems. Using Eq. (21), the distance between *any* two of the three critical submanifolds is calculated to be 4. As a result of this symmetric relationship, each critical submanifold is at the maximum distance from the other two critical submanifolds, and thus any ascent or descent from one of them will be accompanied by an immediate decrease in the distance to all other critical submanifolds.

B. Influence of Hamiltonian parameters

The landscape topology, i.e., the characterization of the critical points of J, depends only on ρ_0 and θ. Section IV A shows that the nature of the topology has a significant effect on whether a gradient-based search is attracted to a saddle submanifold. However, the local geometry (i.e., the nontopological features) of the control landscape is also important and depends on the Hamiltonian. For the Hamiltonian defined in Eq. (1), the dipole matrix μ, the field-free Hamiltonian H_0, and the particular initial field $\varepsilon_0(t)$ each may influence whether an ascent of the landscape closely approaches saddles. In this section, we independently consider the effect of each of these factors.

The parametrization of the dipole matrix elements in Eq. (24) allows μ to take a variety of forms. For $d = 1$, all transitions $|j\rangle \rightarrow |k\rangle, j \neq k$ are equally allowed. For $0 \leqslant d \leqslant 1$, all system transitions $|j\rangle \rightarrow |k\rangle, j \neq k$ are still allowed, but the value of the dipole moment coupling $\langle j|\mu|k\rangle$ decreases exponentially with the difference $|j - k|$. In the limit $d \rightarrow 0^+$, only transitions $|j\rangle \rightarrow |j \pm 1\rangle$ between adjacent system states are allowed. To determine the effect of the dipole coupling parameter d on the attraction of a gradient search to the saddle, we performed a set of optimizations utilizing the same $\rho_0 = |0\rangle\langle 0|$ and θ^m that were used in Sec. IV A. In particular, consider the target observable $\theta^5 = \sum_{j=2}^{6} 0.16|j\rangle\langle j| + 0.2|7\rangle\langle 7|$, which we will denote as case (I) in the remainder of this work. The dipole matrix is given in Eq. (24), and simulations were performed using both the rigid-
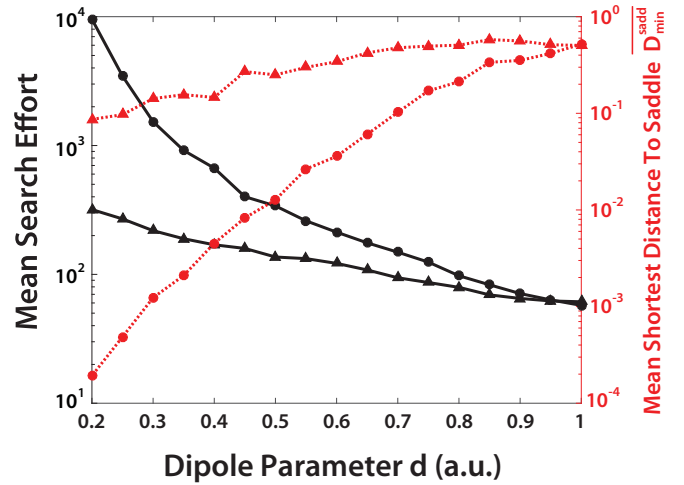

FIG. 2. The mean search effort (solid black lines) and the mean shortest distance to the saddle submanifold $\overline{D_{min}^{sadd}}$ (dotted red lines), as a function of the dipole parameter d, for optimizations of case (I) that use the rigid-rotor-like (circles) and anharmonic oscillator (triangles) systems. Both systems have $N = 8$ levels, and one hundred optimization runs were performed for each value of d.

rotor-like and anharmonic oscillator field-free Hamiltonians [Eqs. (22) and (23), respectively]. In all simulations, the system had $N = 8$ levels, the control interval of $T = 20$ was divided into $L = 512$ steps, and the initial field fluence was $F_0 = 10$. The contingency tables corresponding to case (I) are given in Eq. (26). Accordingly, the landscape contains three critical submanifolds: the global maximum, the global minimum, and a saddle. The high search effort associated with case (I) (i.e., the observable θ^5) for the rotor-like Hamiltonian in Table II indicates that the saddle is very attractive to gradient-based searches when the dipole parameter $d = 0.2$ is used. In this section, we selected values of the dipole parameter over the range $0.2 \leqslant d \leqslant 1$ and we performed one hundred optimizations for each value of d using both forms of H_0. The distance to each critical submanifold was measured using the metric D^i, and the mean of the smallest distance to the saddle, $\overline{D_{min}^{sadd}}$, was determined for each set of one hundred runs.

Figure 2 shows that for both choices of H_0, larger values of d correspond to smaller search effort and to searches that are less attracted to the saddle. The dipole moment for transitions between adjacent states is the same for any value of d, i.e, $\langle j \pm 1|\mu|j\rangle = 1 \; \forall d$. However, the dipole moment for a transition between nonadjacent states decreases as d decreases. Therefore, these results indicate that the landscape saddle for this control problem is more attractive to gradient searches when nonadjacent states are coupled less strongly to one another. While the control problem corresponding to the observable θ^5 contains a very attractive saddle for $d = 0.2$, as observed in Sec. IV A, searches performed for larger values of d do not approach the saddle closely. This result demonstrates that the values of the dipole elements play a significant role in determining the effect of saddles on gradient optimizations of case (I).

The expression for the objective J at critical points in Eq. (10) helps to clarify the role of the dipole matrix elements in this control problem. For the simulations in this section, the

global maximum of the landscape corresponds to evolution operators U_T that align $p_1 = 1$ and $o_1^5 = 0.2$, which are the largest eigenvalues of ρ_0 and θ^5, respectively. For case (I), the density matrix eigenvalues satisfy $\rho_j = 0 \forall j \geqslant 2$, so aligning p_1 and o_1^5 strictly assures an optimal solution at the top of the landscape that lies on the global maximum submanifold of optimal solutions. As Eq. (18) shows, the alignment of any pair of eigenvalues from the initial state and the target observable, respectively, corresponds to one of the blocks U_{jk} into which we divide U_T. Since both p_1 and o_1^5 are of unit multiplicity, the block U_{11} is a single element of U_T. The relevant element is $\langle 7|U_T|0 \rangle$, since p_1 is associated with the ground state $|0\rangle$ of the system and o_1^5 is associated with the state $|7\rangle$. Therefore, the optimal objective value is achieved via an alignment of ρ_0 and θ that corresponds to the system transition $|0\rangle \rightarrow |7\rangle$.

For $d \approx 1$, the dipole moment $\langle 7|\mu|0 \rangle = d^6$ for this transition may be large enough to achieve the alignment of ρ_0 and θ required for a globally optimal control directly, via a $|0\rangle \rightarrow |7\rangle$ transition. In contrast, for small values of d, the dipole moment for the $|0\rangle \rightarrow |7\rangle$ transition is much smaller than the dipole moment for transitions between adjacent states. As a result, the optimal evolution operator U_T is more likely to correspond to a series of adjacent-state transitions (e.g., $|0\rangle \rightarrow |1\rangle \rightarrow \ldots \rightarrow |7\rangle$) constructively interfering along with additional companion pathways in order to reach the top of the landscape. Analogously, the saddle submanifold corresponds to unitary transformations that align the eigenvalue $p_1 = 1$ with the eigenvalue $o_2^5 = 0.16$. The expression for θ^5 in Eq. (25) shows that o_2^5 encompasses the states $|2\rangle, |3\rangle, \ldots, |6\rangle$, so the transitions $|0\rangle \rightarrow |j\rangle, 2 \leqslant j \leqslant 6$ are associated with the saddle. The dipole moment for each of these transitions is larger than the dipole moment for the transition $|0\rangle \rightarrow |7\rangle$ associated with the global maximum. For case (I), this disparity grows as d decreases, making it more likely that the gradient search will come even closer to the saddle.

Figure 2 also indicates that the relationship between the dipole parameter d and the attractiveness of the saddle is much more dramatic for the rigid-rotor-like system than for the anharmonic oscillator system. While the two field-free Hamiltonians lead to similar values of the mean search effort and of $\overline{D_{\min}^{\text{sadd}}}$ when $d = 1$, smaller values of the dipole parameter lead to a disparity between the rotor-like and oscillator optimizations. When $d = 0.2$, the rigid-rotor-like simulations required a mean search effort of 9562 iterations and led to a mean shortest saddle distance of $\overline{D_{\min}^{\text{sadd}}} = 1.92 \times 10^{-4}$. On the other hand, the anharmonic oscillator simulations required a mean search effort of 318 iterations and a mean shortest saddle distance of $\overline{D_{\min}^{\text{sadd}}} = 8.66 \times 10^{-2}$. Therefore, the form of the field-free Hamiltonian also influences the attractiveness of the saddle in case (I). However, the correlation between the dipole parameter d and the attractiveness of the saddle is qualitatively similar for both the rotor-like and oscillator systems. The same trends were observed with optimizations using other observables θ^m (not shown here) for the anharmonic oscillator system; as in the rigid-rotor simulations summarized in Table II, a greater multiplicity of the second-largest observable eigenvalue corresponded to searches that approached the saddle more closely.

TABLE III. Optimization results for case (I), for various values of the initial field fluence F_0. The rigid-rotor-like system with $N = 8$ states was used, and the dipole parameter $d = 0.2$. One hundred optimization runs were performed for each F_0 value.

F_0	$\overline{D_{\min}^{\text{sadd}}}$	MSE
10^1	1.92×10^{-4}	9562
5×10^1	2.66×10^{-2}	572
10^2	1.13×10^{-1}	372
5×10^2	4.57×10^{-1}	179
10^3	4.89×10^{-1}	150

Additionally, the strength of the field-system interaction is proportional to both the transition dipole moment and the amplitude of the control field. Therefore, a transition for which the dipole moment is very small can still occur with significant probability if the amplitude of the field is sufficiently large. To determine whether the amplitude of the control field $\varepsilon(t)$ influences the form taken by U_T during the course of an optimization and thus affects whether gradient optimizations are attracted to the saddle, we performed additional sets of simulations with larger values of the initial field fluence F_0 over the range $50 \leqslant F_0 \leqslant 1000$. These optimizations used the rigid-rotor-like system in Eqs. (22) and (24), with $d = 0.2$. All other parameters were the same as for the prior simulations in this section. One hundred optimizations were performed for each value of F_0.

As reported in Table III, the use of a larger initial field fluence significantly reduced both the search effort and the mean proximity to the saddle during a search. Comparing this result to Fig. 2, we find that gradient optimizations of case (I) are significantly attracted to the saddle submanifold only when both the dipole parameter d and the initial field fluence F_0 are sufficiently small. These results emphasize the importance of carefully choosing optimization parameters, such as the initial fluence, in order for an OCT search of this problem to be as efficient as possible. Even if the critical topology of the landscape and the form of the dipole matrix yield saddles that are likely to attract a gradient search, an optimization of case (I) is unlikely to approach these saddles if the amplitude of the field is sufficiently large. The collective results from Table III reflect that the rapidly evolving search at high initial fluence quickly passes the saddle, in contrast to the behavior in Fig. 1.

C. The number of saddles

The optimizations in Secs. IV A and IV B were performed on 8-level control problems for which the landscape has only one saddle submanifold, due to the degeneracy of the initial state ρ_0 (i.e., all but one of the eigenvalues were equal to zero) and the nature of the target observable. Thus, we showed that saddles may attract gradient searches when ρ_0 is a projector onto a pure state and θ has the particular structure described in Eq. (25). In contrast, the landscape can contain a greater number of saddles when ρ_0 and θ have different structures, reaching the extreme when both operators are full rank and nondegenerate. In this section, we investigate the effect of a large number of saddles by performing optimizations of J on two additional control problem cases:

(II) The initial density matrix is $\rho_0 = |0\rangle\langle 0|$, as in case (I), but a different full rank target observable θ is used for each simulation. Each observable is defined as

$$\theta = \mathcal{N}_o \sum_{j=1}^{N} \tilde{o}_j |j\rangle\langle j|, \qquad (27)$$

where the normalization $\mathcal{N}_o = 1/\sum_j \tilde{o}_j$, and each of the values $\tilde{o}_j$ are randomly selected from a uniform distribution on the interval [0,1]. The critical topology for this problem was determined using the methods in Sec. II B. There are eight contingency tables, each of which is an 8×2 matrix with column sums $a_1 = 1, a_2 = 7$ and row sums $b_k = 1 \quad \forall k$:

$$C^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \dots, \quad C^8 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}. \qquad (28)$$

C^1 corresponds to the global maximum, C^8 corresponds to the global minimum, and the remaining contingency tables correspond to the six saddle submanifolds of the landscape. For this case (II), the expression in Eq. (11) for the critical objective values simplifies to $J_i = o_i$.

(III) The initial density matrix and the target observable are both full rank and lack degeneracy, and each optimization uses a different choice of both ρ_0 and θ. The observable is defined as in Eq. (27), and ρ_0 is analogously defined as

$$\rho_0 = \mathcal{N}_p \sum_{k=1}^{N} \tilde{p}_k |k\rangle\langle k|, \qquad (29)$$

where the normalization $\mathcal{N}_p = 1/\sum_k \tilde{p}_k$ and each value $\tilde{p}_k$ is selected randomly from the interval [0,1]. Each of the $8! = 40\,320$ contingency tables is one of the N-dimensional permutation matrices Π and its associated critical objective value is determined using Eq. (10). Two of the critical submanifolds correspond to the global maximum and minimum, and the remainder are saddles.

Cases (II) and (III) differ from each other only in the form of ρ_0. For both cases, the rigid-rotor-like system from Eqs. (22) and (24) is used, with $N = 8$ levels and the dipole parameter is $d = 0.2$. The final time is $T = 20$ and the control period is divided into $L = 512$ intervals. A random initial control field $\varepsilon_0(t)$ with a fluence $F_0 = 10$ was used for each search, and one hundred runs were performed for cases (II) and (III). The distance $D^i(U_T)$ to each critical submanifold was measured during every run. The control landscape for cases (II) and (III) have many saddles, and $D_{\min}^{\mathrm{sadd}}$ is defined as the shortest distance to *any* saddle during an optimization. The results of these optimizations are reported in Table IV, with the results for case (I) (from Sec. IV A) included for comparison.

For case (I), the single saddle on the control landscape was extremely attractive to a gradient search. However, when the same parameters were used for cases (II) and (III), the optimizations did not approach any saddle very closely, especially for case (III). For case (II), the mean shortest

TABLE IV. Optimization results for cases (I) through (III). For each case, the rigid-rotor-like system with $N = 8$ levels and dipole parameter $d = 0.2$ was used. One hundred optimization runs were performed for each control case.

Case	ρ_0	θ	No. of saddles	$\overline{D_{\min}^{\mathrm{sadd}}}$	MSE
(I)	Pure state	θ^5	1	1.92×10^{-4}	9562
(II)	Pure state	Full rank	6	9.93×10^{-2}	302
(III)	Full rank	Full rank	40318	5.56×10^{-1}	332

distance to any saddle was $\overline{D_{\min}^{\mathrm{sadd}}} = 9.93 \times 10^{-2}$, and no search passed closer than $D_{\min}^{\mathrm{sadd}} = 4.45 \times 10^{-3}$ to any saddle. For case (III), the mean shortest distance to any saddle was $\overline{D_{\min}^{\mathrm{sadd}}} = 0.556$, and no search passed closer than $D_{\min}^{\mathrm{sadd}} = 0.194$ to any saddle. Thus, the landscape saddles for cases (II) and (III) are much less likely to attract gradient searches than the saddle in case (I). Figure 3 demonstrates this point by illustrating the value of the objective J and the distance to twenty randomly selected critical submanifolds at each iteration of one optimization of case (III). The large number of saddles on the corresponding control landscape made it necessary to use a random sample for graphical purposes, but this sample is qualitatively representative of the entire set of 40 320 critical submanifolds. The single saddle for the optimization in Fig. 1 is far more attractive than any of the saddles represented in Fig. 3. While the small "kinks" in Fig. 3(a) indicate points at which the optimization was attracted to a saddle, the distance to each of the $8! - 2 = 40\,318$ saddles was monitored for each optimization and none of them exhibit more than these minor effects.
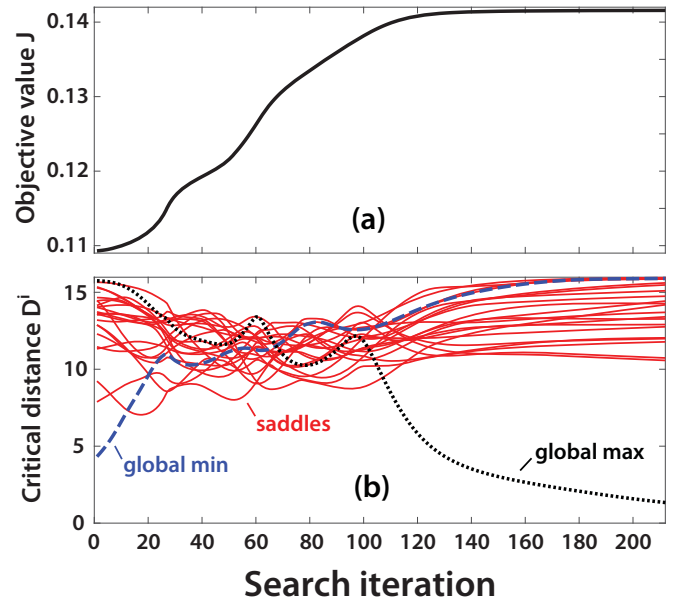

FIG. 3. (a) The objective value at each iteration for an optimization of case (III). (b) The value of the critical distance metric at each iteration of the same search, for D^{max} (dotted black line), D^{min} (dashed blue line), and D^{sadd} for twenty randomly selected saddles (solid red lines).

For cases (II) and (III), the eigenvalues of ρ_0 and θ that must be aligned in order for a control to lie on the saddle all have unit multiplicity. Therefore, the simulations in this section support our previous conclusion that the multiplicity of the eigenvalues is correlated with attractive landscape saddles for these cases. Furthermore, this result suggests that when the landscape has a large number of saddle submanifolds, a gradient search is less likely to to be significantly attracted to any one of them. This conclusion is significant for the control of nominally complex systems where ρ_0 and θ may have high rank.

D. The number of system states

For the optimizations in Secs. IV A and IV B, the control landscape has only one saddle submanifold due to the degeneracy in the initial state and the observable. For the optimizations in Sec. IV C, one or both of θ and ρ_0 are full rank and the resulting control landscape has many saddles. In addition, all of the previous simulations in this paper were performed on eight-level systems. Since problems of physical interest often involve systems with many states, in this section we investigate whether the effect of saddles on a gradient search depends on the number of levels N.

All simulations in this section were performed on the rigid-rotor-like system from Eqs. (22) and (24). Gradient searches were performed for control cases (I) through (III), and each case was generalized to N-level systems. For case (I), the initial state and observable are still defined as in Sec. IV B, with $\rho_0 = |0\rangle\langle 0|$ and $\theta^5 = \sum_{j=2}^6 0.16|j\rangle\langle j| + 0.2|7\rangle\langle 7|$. Therefore, the only change to the eigenvalue spectra of ρ_0 and θ for different values of N is the multiplicity of the smallest (zero) eigenvalue of each operator. For all values of N, the landscape contains three critical submanifolds, which again correspond to the global maximum, the global minimum, and a single saddle. For case (II), there are N critical submanifolds, of which $N-2$ are saddles. The contingency tables are the set of $N \times 2$ matrices with column sums $a_1 = 1, a_2 = N - 1$ and row sums $b_k = 1, 1 \leqslant k \leqslant N$. For case (III), there are $N!$ critical submanifolds, of which $N! - 2$ are saddles. Each contingency table is one of the N-dimensional permutation matrices Π.

First, we performed optimizations of case (I) for a number of states ranging over $12 \leqslant N \leqslant 40$. The dipole parameter was $d = 0.5$, the final time was $T = 20$, and $L = 512$ time intervals were used. One hundred optimizations using initial fields with fluence $F_0 = 10$ were performed for each value of N, and the distance D^i to the saddle submanifold was monitored during each optimization. The results are illustrated in Fig. 4, with the prior results for $N = 8$ and $d = 0.5$ from Sec. IV B included as well. They indicate that both the mean search effort and the mean value of $D_{\min}^{\text{sadd}}$ for case (I) remain relatively constant regardless of the number of states. As discussed in Sec. IV A, the saddle submanifold for this case corresponds to controls that align the eigenvalues p_1 and o_2 of ρ_0 and θ, respectively. Since the multiplicity of neither of these eigenvalues increases with N, it is intuitive that the attractiveness of the saddle is relatively invariant to the number of system states. In addition, all system states $|j\rangle, j > 7$ are associated with the smallest (i.e., zero) eigenvalue of both the initial state and the target observable. Evidently, the higher-lying states play a very limited role in the optimal search.
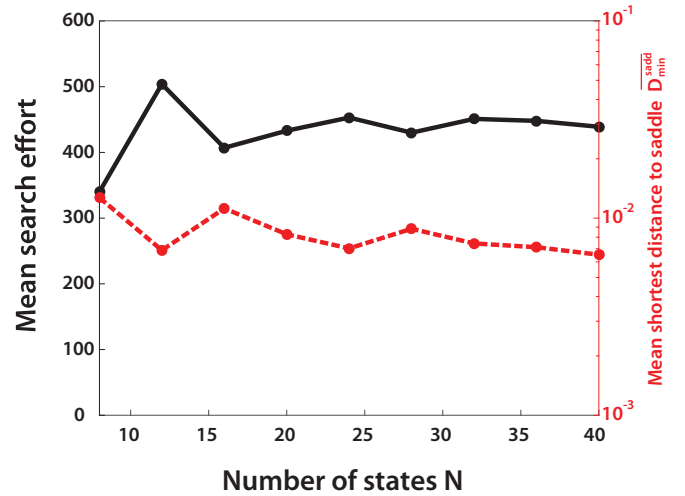

FIG. 4. The mean search effort (solid black line) and the mean shortest distance to the saddle submanifold, $\overline{D_{\min}^{\text{sadd}}}$ (dashed red line), as a function of the number of states N, for gradient optimizations of case (I). The dipole parameter $d = 0.5$, and one hundred optimization runs were performed for each value of N.

Optimizations were also performed for cases (II) and (III), for $3 \leqslant N \leqslant 16$ (the simulations in Sec. IV C used $N = 8$). In these simulations, the dipole parameter $d = 0.2$ and the control time $T = 20$ was divided into $L = 512$ intervals for $N < 10$ and into $L = 2048$ intervals for $N \geqslant 10$. The initial field fluence was $F_0 = 10$, and one hundred optimizations were performed for each value of N. The distance D^i to the saddle submanifold was measured at each step of the search for every value of N for case (II) and for $3 \leqslant N \leqslant 10$ for case (III). These values were used to determine $D_{\min}^{\text{sadd}}$, which is defined as in Sec. IV C, i.e., the shortest distance to any saddle during a given optimization. For cases (II) and (III), we also define the quantity $D_{\min}^{\text{mean(sadd)}}$ as the mean of the shortest distance to each saddle for a given optimization. Neither measure was calculated for $N > 10$ in case (III) due to the factorial scaling of the number of critical submanifolds (e.g., for $N = 11$, the landscape contains $\sim 3.99 \times 10^7$ saddles).

Figure 5 illustrates that the search effort increases with the number of system states N for both cases, reflecting the complexity of these control problems in many-level systems. For case (I), a critical unitary evolution U_T is optimal (i.e., corresponds to the global maximum of the landscape) if and only if it aligns the largest eigenvalue of ρ_0 with the largest eigenvalue of θ. For case (III), however, a critical U_T must simultaneously align each of the N eigenvalues of the initial state ρ_0 with a particular eigenvalue of the target observable θ in order to be optimal. Despite this scaling with N, the mean search effort for cases (II) and (III) at $N = 16$ (3255 and 2049 iterations, respectively) was still significantly less than for optimizations of case (I) for $N = 8$ and $d = 0.2$ (9562 iterations). Once again, we observe that additional system complexity does not significantly impact the effort of optimization for these cases, even with a greater number of saddles present. Among these various cases, there also exist other subtle trends that must result from the details of the dynamics involved.
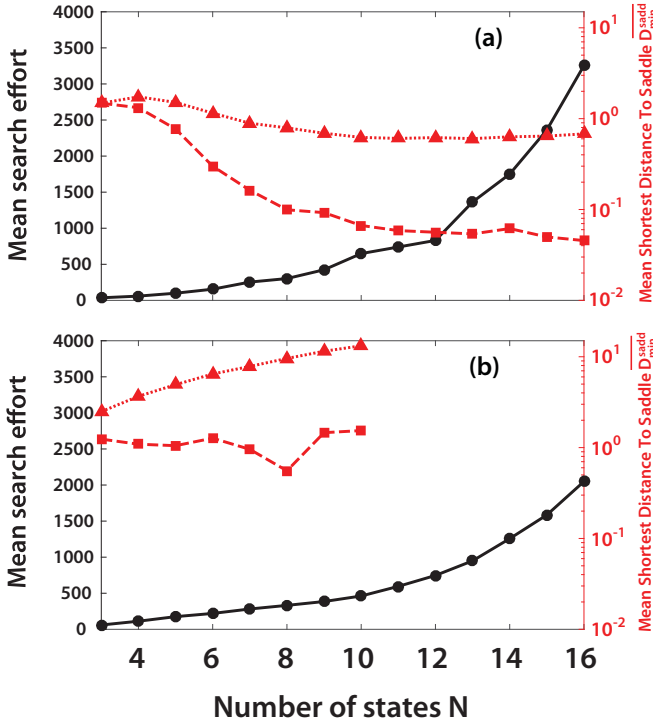
FIG. 5. The mean search effort (black circles), the mean shortest distance to any saddle submanifold, $\overline{D_{\min}^{\text{sadd}}}$ (red squares), and the mean shortest distance averaged over all saddle submanifolds, $\overline{D_{\min}^{\text{mean(sadd)}}}$ (red triangles), as a function of the number of states N, for gradient optimizations of (a) case (II) and (b) case (III). The dipole parameter $d = 0.2$, and one hundred optimization runs were performed for each value of N.
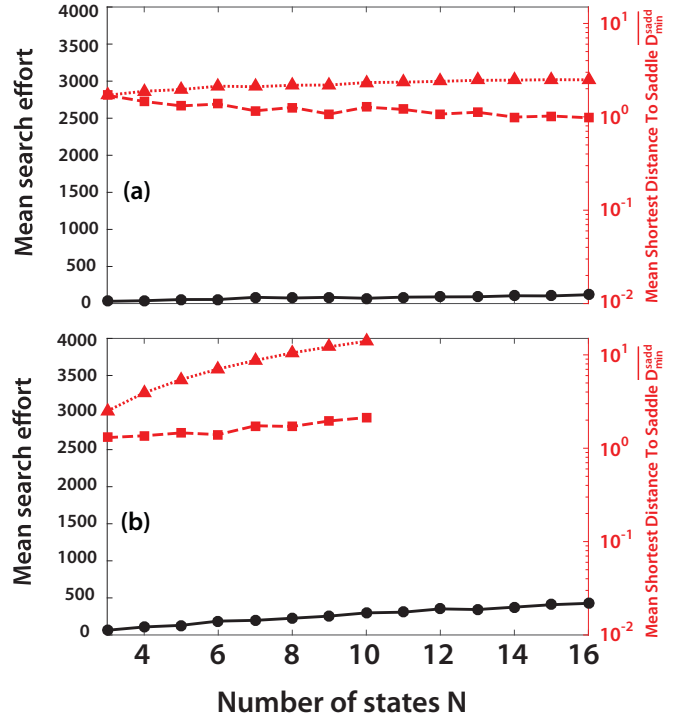
FIG. 6. The mean search effort (black circles), the mean shortest distance to any saddle submanifold, $\overline{D_{\min}^{\text{sadd}}}$ (red squares), and the mean shortest distance averaged over all saddle submanifolds, $\overline{D_{\min}^{\text{mean(sadd)}}}$ (red triangles), as a function of the number of states N, for gradient optimizations of (a) case (II) and (b) case (III). The dipole parameter $d = 1$, and one hundred optimization runs were performed for each value of N.

Additionally, Fig. 5 shows that the mean value $\overline{D_{\min}^{\text{sadd}}}$ remains comfortably large for all values of N for cases (II) and (III), indicating that optimizations do not closely approach any saddles. For case (II), the mean shortest distance to any saddle was $\overline{D_{\min}^{\text{sadd}}} = 1.48$ for $N = 3$; while this value initially decreases as N grows, it remains relatively constant at $\overline{D_{\min}^{\text{sadd}}} \approx 0.08$ for $N \geqslant 8$. Most saddles are not approached even this closely, as the mean shortest distance averaged over all saddles was $\overline{D_{\min}^{\text{mean(sadd)}}} \approx 0.7$ for $N \geqslant 8$. For case (III), $\overline{D_{\min}^{\text{sadd}}} \approx 1$ for all values of N, while $\overline{D_{\min}^{\text{mean(sadd)}}} = 2.5$ for $N = 3$ and increases to $\overline{D_{\min}^{\text{mean(sadd)}}} = 13.2$ for $N = 10$. Although the number of saddles on the landscape corresponding to case (III) increases factorially with N, the trend in $\overline{D_{\min}^{\text{mean(sadd)}}}$ indicates that the average attractiveness of each saddle decreases with N. This behavior is consistent with a mathematical analysis of the kinematic volume fraction near critical submanifolds that was performed on a related observable problem [120]. These competing trends may explain why, although the number of saddles increases by a factor of 3.5×10^{12} between $N = 3$ and $N = 16$, the mean search effort only increases by a factor of 34 over the same range (60 for $N = 3$ and 2049 for $N = 16$). This dramatic disparity shows that case (III) is surprisingly amenable to gradient optimization, despite its very large number of landscape saddles.

Like the saddle attraction measured in Sec. IV B, the observed scaling of search effort with N for cases (II) and (III) also depends on Hamiltonian parameters. We repeated the optimizations represented in Fig. 5 in the same manner as described above, but with the dipole parameter $d = 1$ rather than $d = 0.2$. The results of this set of simulations are illustrated in Fig. 6. All measures of the distance to saddles are at least equal to the values observed for $d = 0.2$, and the search effort scaling is significantly less; at $N = 16$, the mean effort for cases (II) and (III) is 119 and 427, respectively. The strong coupling may have accelerated the searches for these cases by preventing them from lingering near any of the saddles.

In conclusion, the optimizations of case (I) in this section show that increasing the number of system states does not affect the search effort or the attractiveness of the saddle submanifold. While the mean search effort increases with N for cases (II) and (III), we demonstrate that this scaling is not due the saddles becoming more attractive, as the point of closest approach to a saddle does not change significantly with N. On average, the degree of attraction to any individual saddle appears invariant to N for case (II) and decreases with N for case (III).

E. Control constraints

In the ODE45 algorithm, the accuracy demanded of the solutions to Eq. (12) is determined by the error tolerance τ. In

TABLE V. Optimization results for case (I) and various values of τ, the absolute error tolerance in ODE45. The system has $N = 8$ states and one hundred optimization runs were performed for each τ value.

τ	No. failed	$\overline{D_{\text{fail}}^{\text{sadd}}}$
10^{-1}	100	1.66
10^{-2}	100	1.35×10^{-2}
10^{-3}	97	2.68×10^{-4}
10^{-4}	89	9.20×10^{-5}
10^{-5}	50	1.41×10^{-5}
10^{-6}	14	2.10×10^{-6}
10^{-7}	0	

a previous numerical study that used a gradient algorithm to investigate the role of control constraints in OCT optimization, the effect of changing this tolerance was studied [119] and it was determined that a choice of $\tau = 10^{-8}$ yields sufficiently accurate solutions to find optimal controls for the state-transition objective. It was also shown that large values of τ, and the resulting inaccurate solutions to Eq. (12), are a severe constraint that may result in search failure (i.e., a decrease in the value of the objective functional after an iteration). However, the simulations in Ref. [119] involved the state-transition landscape, which lacks saddles. When a gradient search is close to a critical point of the landscape, the norm of the gradient is small, and more accuracy may be required of the solutions to Eq. (12) in order to ensure successful optimization. Therefore, a smaller value of τ may need to be used when the control landscape has saddles, particularly if the saddles attract gradient searches. The simulations in this section explore whether less accurate solutions to Eq. (12) can cause searches to fail in close proximity to a saddle submanifold.

We performed optimizations of case (I) as defined in Sec. IV B, with $N = 8$ levels and $d = 0.2$. Case (I) was chosen because its control landscape contains the most attractive saddle identified in this work. The control period was $T = 20$, the time discretization was $L = 512$, and each initial field had fluence $F_0 = 10$. One hundred runs were performed for each value of the absolute error tolerance τ over the range $10^{-7} \leqslant \tau \leqslant 10^{-1}$. For each failed optimization, the distance to the saddle at the final iteration, $D_{\text{fail}}^{\text{sadd}}$, was recorded. The results of these optimizations are reported in Table V, and they confirm that a large value of τ will cause searches that use ODE45 to fail. All searches failed for $\tau \geqslant 10^{-2}$, and at least one search failed for $\tau \geqslant 10^{-6}$. The mean distance from the saddle at which searches fail increases with τ, suggesting that the appropriate error tolerance for a particular optimization of this problem is determined by the attractiveness of the saddle(s) on the corresponding control landscape. When the search approaches a saddle more closely, a smaller value of τ is required in order to avoid search failure. The results of the simulations in Ref. [119] support this conclusion; for a control problem that lacks saddles, all optimizations were successful for $\tau \leqslant 2 \times 10^{-3}$. The searches on a landscape that has an attractive saddle require significantly more accurate solutions to Eq. (12), as $\tau \leqslant 10^{-7}$ is required to ensure that all searches succeed. Thus, the choice of $\tau = 10^{-8}$ for this paper is adequate for the field to reach its optimal form.

V. CONCLUSION

The critical topology of the quantum control landscape has been analyzed theoretically [83,97,98,100–102,105,121–124] in consideration of the mounting successes of diverse optimal control experiments and simulations. These theoretical works have shown that the landscape lacks local optima when three assumptions are met: controllability, a full-rank Jacobian matrix $\delta U_T / \delta \varepsilon(t)$ everywhere on the landscape, and an unconstrained control field $\varepsilon(t)$. A recent numerical work [119] suggests that only the latter condition is of prime importance to avoid significant resource constraints. Satisfaction of these assumptions ensures that all intermediate critical points (i.e., those that do not correspond to the global maximum or minimum) are saddles. For the observable objective considered here, the initial state ρ_0 and the target observable θ determine the topology of the control landscape, which may have as many as $N! - 2$ saddle submanifolds for an N-level system. This paper has investigated the effect of these saddles on gradient searches.

At one extreme, we identified a control problem, case (I), with a landscape that contains a highly attractive saddle submanifold (i.e., a saddle that almost all gradient searches will approach closely at some point during the optimization). For this problem, the majority of algorithmic iterations takes place very close to the saddle; the distance to the saddle was calculated using the critical distance metric [114]. We also identified features of the landscape topology that lead to this phenomenon. For this case, a saddle submanifold is more attractive to a gradient search when the eigenvalues that correspond to the saddle alignment have greater multiplicity, and the proximity of the search to the saddle is also influenced by parameters of the system Hamiltonian; optimizations experience a significantly greater attraction to the saddle when the field-free Hamiltonian has a rigid-rotor-like energy structure rather than that of an anharmonic oscillator. Even with a rigid-rotor-like H_0 and the particular degeneracy of ρ_0 and θ described above, the saddle in case (I) is only observed to attract gradient searches when the dipole moment for transitions between nonadjacent states and the initial fluence of the control field are both sufficiently small. This choice of values may result in a tendency for the gradient search to initially drive the unitary evolution operator towards alignments of the eigenvalues of ρ_0 and θ that correspond to a saddle submanifold before optimizing, rather than driving the evolution toward the global maximum directly.

We also studied cases for which the target observable, or both the initial state and the observable, are full rank. The landscapes for these problems have multiple saddles (a very large number, in the latter case). Importantly, we demonstrated that the greater number of saddles for these cases does not imply a correspondingly greater probability that a search will closely approach *any* saddle. In the case of the control problem for which the landscape contains the largest possible number of saddles, the average attractiveness of the saddles *decreases* significantly as N *increases*. This result has significance for many normally complex laboratory circumstances where a high density of occupied states may be involved and the observable also involves many states. In this regard as well, the gradient-based algorithm used in this

work is likely the most conservative method; typical use of stochastic search algorithms in the laboratory likely will be less sensitive to saddles, as they may be able to "step over" them. Thousands of numerical OCT searches were performed in this paper, and they were only significantly attracted to a saddle for a very particular choice of the initial state, target observable, and several Hamiltonian parameters. The great majority of the optimizations that were performed do not approach any saddle closely. Even for a control problem that corresponds to a landscape with an attractive saddle, it was shown that a careful choice of algorithmic parameters ensures successful optimization. This conclusion is based on the extensive numerical studies performed for several control problem cases in this work. While there is still the need for a rigorous mathematical understanding of its origin that

builds on the foundations in Ref. [120], the results for these cases indicated that control landscapes free of local optima are highly favorable for optimization, even when saddles are present.

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

[2] R. J. Levis and H. A. Rabitz, J. Phys. Chem. A **106**, 6427 (2002).

[3] D. Goswami, Phys. Rep. **374**, 385 (2003).

[4] T. Brixner and G. Gerber, ChemPhysChem **4**, 418 (2003).

[5] M. Dantus and V. V. Lozovoy, Chem. Rev. **104**, 1813 (2004).

[6] P. Nuernberger, G. Vogt, T. Brixner, and G. Gerber, Phys. Chem. Chem. Phys. **9**, 2470 (2007).

[7] C. Brif, R. Chakrabarti, and H. Rabitz, New J. Phys. **12**, 075008 (2010).

[8] M. Wollenhaupt and T. Baumert, Faraday Discuss. **153**, 9 (2011).

[9] C. Brif, R. Chakrabarti, and H. Rabitz, in *Advances in Chemical Physics*, Vol. 148, edited by S. A. Rice and A. R. Dinner (Wiley, New York, 2012), pp. 1–76.

[10] I. R. Sola, J. Gonzalez-Vazquez, R. de Nalda, and L. Banares, Phys. Chem. Chem. Phys. **17**, 13183 (2015).

[11] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, Eur. Phys. J. D **69**, 279 (2015).

[12] C. P. Koch, J. Phys.: Condens. Matter **28**, 213001 (2016).

[13] T. Hornung, R. Meier, and M. Motzkus, Chem. Phys. Lett. **326**, 445 (2000).

[14] T. C. Weinacht, R. Bartels, S. Backus, P. H. Bucksbaum, B. Pearson, J. M. Geremia, H. Rabitz, H. C. Kapteyn, and M. M. Murnane, Chem. Phys. Lett. **344**, 333 (2001).

[15] R. A. Bartels, T. C. Weinacht, S. R. Leone, H. C. Kapteyn, and M. M. Murnane, Phys. Rev. Lett. **88**, 033001 (2002).

[16] J. Konradi, A. K. Singh, and A. Materny, J. Photochem. Photobiol. A **180**, 289 (2006).

[17] D. B. Strasfeld, S.-H. Shim, and M. T. Zanni, Phys. Rev. Lett. **99**, 038102 (2007).

[18] J. Konradi, A. Scaria, V. Namboodiri, and A. Materny, J. Raman Spectrosc. **38**, 1006 (2007).

[19] A. Scaria, J. Konradi, V. Namboodiri, and A. Materny, J. Raman Spectrosc. **39**, 739 (2008).

[20] A. C. W. van Rhijn, A. Jafarpour, M. Jurna, H. L. Offerhaus, and J. L. Herek, Faraday Discuss. **153**, 227 (2011).

[21] C. J. Bardeen, V. V. Yakovlev, K. R. Wilson, S. D. Carpenter, P. M. Weber, and W. S. Warren, Chem. Phys. Lett. **280**, 151 (1997).

[22] T. Brixner, N. H. Damrauer, P. Niklaus, and G. Gerber, Nature (London) **414**, 57 (2001).

[23] V. I. Prokhorenko, A. M. Nagy, and R. J. D. Miller, J. Chem. Phys. **122**, 184502 (2005).

[24] O. Nahmias, O. Bismuth, O. Shoshana, and S. Ruhman, J. Phys. Chem. A **109**, 8246 (2005).

[25] L. Bonacina, J. Extermann, A. Rondi, V. Boutou, and J.-P. Wolf, Phys. Rev. A **76**, 023408 (2007).

[26] D. G. Kuroda, C. P. Singh, Z. Peng, and V. D. Kleiman, Science **326**, 263 (2009).

[27] P. van der Walle, M. T. W. Milder, L. Kuipers, and J. L. Herek, Proc. Natl. Acad. Sci. USA **106**, 7714 (2009).

[28] M. Roth, L. Guyon, J. Roslund, V. Boutou, F. Courvoisier, J.-P. Wolf, and H. Rabitz, Phys. Rev. Lett. **102**, 253001 (2009).

[29] J. Roslund, M. Roth, L. Guyon, V. Boutou, F. Courvoisier, J.-P. Wolf, and H. Rabitz, J. Chem. Phys. **134**, 034511 (2011).

[30] R. Bartels, S. Backus, E. Zeek, L. Misoguti, G. Vdovin, I. P. Christov, M. M. Murnane, and H. C. Kapteyn, Nature (London) **406**, 164 (2000).

[31] R. A. Bartels, M. M. Murnane, H. C. Kapteyn, I. Christov, and H. Rabitz, Phys. Rev. A **70**, 043404 (2004).

[32] D. H. Reitze, S. Kazamias, F. Weihe, G. Mullot, D. Douillet, F. Augé, O. Albert, V. Ramanathan, J. P. Chambaret, D. Hulin, and P. Balcou, Opt. Lett. **29**, 86 (2004).

[33] T. Pfeifer, C. Spielmann, and G. Gerber, Rep. Prog. Phys. **69**, 443 (2006).

[34] C. Winterfeldt, C. Spielmann, and G. Gerber, Rev. Mod. Phys. **80**, 117 (2008).

[35] M. P. A. Branderhorst, P. Londero, P. Wasylczyk, C. Brif, R. L. Kosut, H. Rabitz, and I. A. Walmsley, Science **320**, 638 (2008).

[36] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature (London) **458**, 996 (2009).

[37] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber, Science **282**, 919 (1998).

[38] M. Bergt, T. Brixner, B. Kiefer, M. Strehle, and G. Gerber, J. Phys. Chem. A **103**, 10381 (1999).

[39] R. J. Levis, G. M. Menkir, and H. Rabitz, Science **292**, 709 (2001).

[40] Š. Vajda, A. Bartelt, E.-C. Kaposta, T. Leisner, C. Lupulescu, S. Minemoto, P. Rosendo-Francisco, and L. Wöste, Chem. Phys. **267**, 231 (2001).

[41] C. Daniel, J. Full, L. González, C. Lupulescu, J. Manz, A. Merli, Š. Vajda, and L. Wöste, Science **299**, 536 (2003).

[42] P. Nuernberger, D. Wolpert, H. Weiss, and G. Gerber, Proc. Natl. Acad. Sci. USA **107**, 10366 (2010).

[43] J. Plenge, A. Wirsing, I. Wagner-Drebenstedt, I. Halfpap, B. Kieling, B. Wassermann, and E. Rühl, Phys. Chem. Chem. Phys. **13**, 8705 (2011).

[44] J. L. Herek, W. Wohlleben, R. J. Cogdell, D. Zeidler, and M. Motzkus, Nature (London) **417**, 533 (2002).

[45] W. Wohlleben, T. Buckup, J. L. Herek, R. J. Cogdell, and M. Motzkus, Biophys. J. **85**, 442 (2003).

[46] T. Buckup, T. Lebold, A. Weigel, W. Wohlleben, and M. Motzkus, J. Photochem. Photobiol. A **180**, 314 (2006).

[47] J. Savolainen, R. Fanciulli, N. Dijkhuizen, A. L. Moore, J. Hauer, T. Buckup, M. Motzkus, and J. L. Herek, Proc. Natl. Acad. Sci. U.S.A. **105**, 7641 (2008).

[48] G. Vogt, G. Krampert, P. Niklaus, P. Nuernberger, and G. Gerber, Phys. Rev. Lett. **94**, 068305 (2005).

[49] G. Vogt, P. Nuernberger, T. Brixner, and G. Gerber, Chem. Phys. Lett. **433**, 211 (2006).

[50] B. Dietzek, B. Brüggemann, T. Pascher, and A. Yartsev, Phys. Rev. Lett. **97**, 258301 (2006).

[51] V. I. Prokhorenko, A. M. Nagy, S. A. Waschuk, L. S. Brown, R. R. Birge, and R. J. D. Miller, Science **313**, 1257 (2006).

[52] M. Greenfield, S. D. McGrane, and D. S. Moore, J. Phys. Chem. A **113**, 2333 (2009).

[53] H. Rabitz and W. S. Zhu, Acc. Chem. Res. **33**, 572 (2000).

[54] D. D'Alessandro, *Introduction to Quantum Control and Dynamics* (Chapman & Hall/CRC, Boca Raton, FL, 2007).

[55] J. Werschnik and E. K. U. Gross, J. Phys. B: At. Mol. Opt. Phys. **40**, R175 (2007).

[56] G. G. Balint-Kurti, S. Zou, and A. Brown, in *Advances in Chemical Physics*, Vol. 138, edited by S. A. Rice (Wiley, New York, 2008), pp. 43–94.

[57] D. Kammerlander, A. Castro, and M. A. L. Marques, Phys. Rev. A **83**, 043413 (2011).

[58] A. Castro, J. Werschnik, and E. K. U. Gross, Phys. Rev. Lett. **109**, 153603 (2012).

[59] Y. Ohtsuki, K. Ohara, M. Abe, K. Nakagami, and Y. Fujimura, Chem. Phys. Lett. **369**, 525 (2003).

[60] M. Artamonov, T.-S. Ho, and H. Rabitz, Chem. Phys. **305**, 213 (2004).

[61] M. Artamonov, T.-S. Ho, and H. Rabitz, J. Chem. Phys. **124**, 064306 (2006).

[62] Y. Kurosaki, M. Artamonov, T.-S. Ho, and H. Rabitz, J. Chem. Phys. **131**, 044306 (2009).

[63] R. Kosloff, S. A. Rice, P. Gaspard, S. Tersigni, and D. J. Tannor, Chem. Phys. **139**, 201 (1989).

[64] S. Shi and H. Rabitz, Comput. Phys. Commun. **63**, 71 (1991).

[65] P. Gross, D. Neuhauser, and H. Rabitz, J. Chem. Phys. **96**, 2834 (1992).

[66] K. Nakagami, Y. Ohtsuki, and Y. Fujimura, J. Chem. Phys. **117**, 6429 (2002).

[67] K. Krieger, A. Castro, and E. K. U. Gross, Chem. Phys. **391**, 50 (2011).

[68] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, Phys. Rev. A **75**, 023602 (2007).

[69] U. Hohenester, J. Grond, and J. Schmiedmayer, Fortschr. Phys. **57**, 1121 (2009).

[70] J. Grond, G. von Winckel, J. Schmiedmayer, and U. Hohenester, Phys. Rev. A **80**, 053625 (2009).

[71] E. Räsänen and L. B. Madsen, Phys. Rev. A **86**, 033426 (2012).

[72] J. P. Palao and R. Kosloff, Phys. Rev. Lett. **89**, 188301 (2002).

[73] C. M. Tesch and R. de Vivie-Riedle, Phys. Rev. Lett. **89**, 157901 (2002).

[74] J. P. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003).

[75] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[76] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, Phys. Rev. A **72**, 042331 (2005).

[77] U. Hohenester, Phys. Rev. B **74**, 161307 (2006).

[78] S. Montangero, T. Calarco, and R. Fazio, Phys. Rev. Lett. **99**, 170501 (2007).

[79] M. Grace, C. Brif, H. Rabitz, I. A. Walmsley, R. L. Kosut, and D. A. Lidar, J. Phys. B: At. Mol. Opt. Phys. **40**, S103 (2007).

[80] R. de Vivie-Riedle and U. Troppmann, Chem. Rev. **107**, 5082 (2007).

[81] M. Wenin and W. Pötz, Phys. Rev. A **78**, 012358 (2008).

[82] M. Wenin and W. Pötz, Phys. Rev. B **78**, 165118 (2008).

[83] J. Dominy and H. Rabitz, J. Phys. A **41**, 205305 (2008).

[84] V. Nebendahl, H. Häffner, and C. F. Roos, Phys. Rev. A **79**, 012312 (2009).

[85] S. Schirmer, J. Mod. Opt. **56**, 831 (2009).

[86] S. Safaei, S. Montangero, F. Taddei, and R. Fazio, Phys. Rev. B **79**, 064524 (2009).

[87] M. Wenin, R. Roloff, and W. Pötz, J. Appl. Phys. **105**, 084504 (2009).

[88] R. Roloff, M. Wenin, and W. Pötz, J. Comput. Electron. **8**, 29 (2009).

[89] R. Roloff, M. Wenin, and W. Pötz, J. Comput. Theor. Nanosci. **6**, 1837 (2009).

[90] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm, Phys. Rev. Lett. **102**, 090401 (2009).

[91] P. Rebentrost and F. K. Wilhelm, Phys. Rev. B **79**, 060507 (2009).

[92] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Phys. Rev. Lett. **103**, 110501 (2009).

[93] M. D. Grace, J. Dominy, R. L. Kosut, C. Brif, and H. Rabitz, New J. Phys. **12**, 015001 (2010).

[94] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, J. Phys. B: At. Mol. Opt. Phys. **44**, 154013 (2011).

[95] C. M. Trail, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. **105**, 193602 (2010).

[96] L. M. Norris, C. M. Trail, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. **109**, 173603 (2012).

[97] H. Rabitz, M. Hsieh, and C. Rosenthal, Science **303**, 1998 (2004).

[98] T.-S. Ho and H. Rabitz, J. Photochem. Photobiol. A **180**, 226 (2006).

[99] K. Moore, M. Hsieh, and H. Rabitz, J. Chem. Phys. **128**, 154117 (2008).

[100] R. Chakrabarti and H. Rabitz, Int. Rev. Phys. Chem. **26**, 671 (2007).

[101] H. Rabitz, M. Hsieh, and C. Rosenthal, Phys. Rev. A **72**, 052337 (2005).

[102] H. Rabitz, T.-S. Ho, M. Hsieh, R. Kosut, and M. Demiralp, Phys. Rev. A **74**, 012721 (2006).

[103] H. Rabitz, M. Hsieh, and C. Rosenthal, J. Chem. Phys. **124**, 204107 (2006).

[104] M. Hsieh and H. Rabitz, Phys. Rev. A **77**, 042306 (2008).

[105] R. B. Wu, H. Rabitz, and M. Hsieh, J. Phys. A **41**, 015006 (2008).

[106] M. Hsieh, R. B. Wu, and H. Rabitz, J. Chem. Phys. **130**, 104109 (2009).

[107] T.-S. Ho, J. Dominy, and H. Rabitz, Phys. Rev. A **79**, 013422 (2009).

[108] M. Hsieh, R. B. Wu, H. Rabitz, and D. Lidar, Phys. Rev. A **81**, 062352 (2010).

[109] R.-B. Wu, M. A. Hsieh, and H. Rabitz, Phys. Rev. A **83**, 062306 (2011).

[110] G. Riviello, C. Brif, R. Long, R.-B. Wu, K. M. Tibbetts, T.-S. Ho, and H. Rabitz, Phys. Rev. A **90**, 013404 (2014).

[111] B. Russell, H. Rabitz, and R. Wu, arXiv:1608.06198.

[112] K. W. Moore, R. Chakrabarti, G. Riviello, and H. Rabitz, Phys. Rev. A **83**, 012326 (2011).

[113] Q. Sun, I. Pelczer, G. Riviello, R.-B. Wu, and H. Rabitz, Phys. Rev. A **91**, 043412 (2015).

[114] Q. Sun, G. Riviello, R.-B. Wu, and H. Rabitz, J. Phys. A: Math. Theor. **48**, 465305 (2015).

[115] K. W. Moore and H. Rabitz, J. Chem. Phys. **137**, 134113 (2012).

[116] J. G. Digalakis and K. G. Margaritis, Int. J. Comput. Math. **77**, 481 (2001).

[117] K. W. Moore and H. Rabitz, Phys. Rev. A **84**, 012109 (2011).

[118] MathWorks, *MATLAB* (MathWorks, Natick, MA, 1994).

[119] G. Riviello, K. M. Tibbetts, C. Brif, R. Long, R.-B. Wu, T.-S. Ho, and H. Rabitz, Phys. Rev. A **91**, 043401 (2015).

[120] J. Dominy and H. Rabitz, J. Phys. A **44**, 255302 (2011).

[121] C. Altafini, Syst. Contr. Lett. **58**, 213 (2009).

[122] R. B. Wu, A. Pechen, H. Rabitz, M. Hsieh, and B. Tsou, J. Math. Phys. **49**, 022108 (2008).

[123] M. Hsieh, T.-S. Ho, and H. Rabitz, Chem. Phys. **352**, 77 (2008).

[124] Z. Shen, M. Hsieh, and H. Rabitz, J. Chem. Phys. **124**, 204106 (2006).