# Search complexity and resource scaling for the optimal control of quantum observables

Gregory Riviello, Raj Chakrabarti, Katherine Moore, and Herschel Rabitz

January 24, 2013

**Abstract**

Theoretical and experimental quantum optimal control of observable expectation values has enjoyed wide success on systems ranging from atoms to large biomolecules while employing limited computational resources. Prior analysis of this objective established that local extrema in the search landscape of controllable systems have null measure, guaranteeing the convergence of local search algorithms. However, the local structure of the landscape (and thus the computational effort required to achieve convergence) may vary depending on the structure of the control Hamiltonian, observable operator, and quantum density matrix. This work investigates the implications of these features for the dynamics of the maximization of the expectation values of Hermitian observables for several classes of quantum systems. When the density matrix and observable are rank-deficient, the computational effort is largely insensitive to Hamiltonian parameters and scales approximately logarithmically with the Hilbert space dimension of the system. Due to the difficulty of obtaining accurate Hamiltonian estimates, this favorable scaling is essential for the coherent control of systems with large Hilbert spaces. Its origin, including the influence of local saddle critical points and the factors which decouple computational effort from the control Hamiltonian, is examined. These results suggest conditions for feasible optimal control of complex molecular dynamics.

# 1 Introduction

Three classes of problems - gate or propagator control, state or density matrix control, and observable control - have been studied in the quantum control community to date. The primary approaches to quantum optimal control are learning feedback (also called optimal control experiments, or OCE), which iteratively updates the control field based on an experimentally measured value of the performance index, and open loop control (also called optimal control theory, or OCT), wherein the control that maximizes the performance index is identified numerically or analytically.

For problems such as state [1, 2, 3] or propagator [4, 5] control, optimal control theory (OCT) is essential since estimating the entire system propagator or state experimentally at each step, which is required for learning control, is very expensive. Still, application of optimal control theory requires accurate estimates of the time-independent Hamiltonian. Most proposals for quantum computing (which require propagator control) are based on coupled qubit systems, for which the Hamiltonian may be known quite accurately **(ref)**. However, as recently demonstrated for state tomography, estimation error increases rapidly with Hilbert space dimension and often has poor performance for larger systems. Although the Hamiltonian need only be estimated onc, the quantum Hamiltonian estimation problem is more nonlinear in the parameters than state or propagator estimation, and is hence even more prone to estimation error and poor finite sample performance, especially for large Hilbert space dimensions. To date, much of the literature on Hamiltonian estimation has focused on estimation of a small number of unknown Hamiltonian parameters [6]. For larger systems where the Hamiltonian estimation error can be considerable, as is the case with most molecules, OCT may be more useful as a means of obtaining an initial guess for the optimal control $\bar{\varepsilon}(\cdot)$, which may then need to be refined using learning control methods. Alternatively, the problem may be solved entirely through the use of learning control methods. This has been the most common approach used thus far, with numerous laboratory successes to date across a wide variety of molecular systems [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

Due to the necessity of applying learning feedback to identify optimal controls for complex molecular dynamics, the scaling of search effort for learning control optimization plays a central role in determin-

ing the feasibility of such strategies for complex molecules. Experimental evidence indicates that this search effort typically falls between $10^2$ and $10^3$ algorithmic iterations regardless of molecular size (and complexity) for systems studied to date [7, 9, 10, 14, 17, 18, 19, 20, 21, 22, 23, 24]. Recent analytic results on the topology of quantum observable control landscapes show that the measure of false traps in these landscapes is null in the space of possible control fields, facilitating the convergence of learning feedback algorithms. However, this favorable topology is in itself not sufficient to explain the rapid and uniform convergence observed across molecules of qualitatively different complexity in OCE studies. In this work, we study the optimization dynamics for observable learning feedback control for realistic molecular systems and assess the influence of landscape local structure and critical points on these search dynamics. We relate these landscape features to physical properties of the molecular systems - including control Hamiltonian, internal Hamiltonian, density matrix spectrum and type of observable of interest - and assess the conditions under which the search complexity is sufficiently low for rapid convergence of learning feedback strategies.

The paper is organized as follows. **Section 2 develops the theoretical and computational foundation for quantum observable control, including a discussion of the landscape structure and algorithm employed in this work. Section 3 describes a series of metrics designed to characterize the local and global properties of the control landscape. Section 4 introduces a diverse group of generic quantum systems, as well as control Hamiltonians derived from common objectives in molecular spectroscopy, in order to illustrate the range of landscape structures experienced during quantum observable optimizations. Section 5 studies the computational search effort and scaling of control resources for both sets of systems, while Section 6 connects these trends in computational effort to landscape structure via the metrics in Section 3.**

# 2 Theory and Methodology

## 2.1 Dynamical Formulation of the Control Objective

Consider an $N$-dimensional isolated quantum system whose dynamics are governed by the time-dependent Schrödinger equation,

$$i\hbar\frac{\partial U(t,0)}{\partial t} = H(\kappa,t)U(t,0), \qquad U(0,0) \equiv \mathbb{I}, \tag{1}$$

where $H(\kappa,t)$ is the time-dependent Hamiltonian whose control variables are denoted as $\kappa$. We will set $\hbar = 1$. The unitary propagator at some final time $T$ is

$$U(T,0) = \mathbf{T}\text{exp}\left(-i\int_0^T H(\kappa,t)dt\right), \tag{2}$$

where $\mathbf{T}$ is the time-ordering operator and $U(T,0)$ is implicitly understood to be a function of $\kappa$.

We are interested in the class of control objective functionals

$$J[U(\cdot,0),\ \varepsilon(\cdot)] = F(U(T,0)). \tag{3}$$

The endpoint control objective $F$ may be defined as the expectation value of a specified quantum observable $\Theta$, which may be written as follows:

$$F(U) = \text{Tr}(U\rho_0 U^\dagger\Theta). \tag{4}$$

where $\rho_0$ denotes the initial density matrix of the system. The quantum system is assumed to be controllable such that (i) any desired $U$ can be generated by some choice of $\kappa$ at time $T$. In addition, to facilitate convergence analysis, we assume (ii) any local variation of $U$ can be generated via some variation of $\kappa$. Condition (i) is equivalent to the requirement that the Lie algebra generated from $H_0$ and $\mu$ forms a complete set of operators [25], which can be easily verified for any given control system by computing repeated commutators of $H_0$ and $\mu$. Condition (ii) is equivalent to requiring that the components $\mu_{ij}(t)$ of the time evolved density matrix (in the Heisenberg picture) are linearly independent functions of time. Denoting by $\nu(A) = \mathbf{a}$, where the components of $\mathbf{a}$ are the $N^2$ independent parameters of a Hermitian

matrix $A$, this condition is equivalent to demanding that the matrix

$$\mathrm{S}_{ij,pq}^{\varepsilon,T} \quad = \quad \int_0^T \nu\left(\mu(t)\right)\left[\nu\left(\mu(t)\right)\right]' \,\mathrm{dt}$$

where $'$ denotes the matrix transpose, is full rank (nonsingular) **(ref)**. $S^{\varepsilon,T}$, a Gramian matrix (i.e., the tensor product of vectors in a real vector space), is an $N \times N$-dimensional positive-semidefinite symmetric matrix. This condition, which facilitates convergence of local search algorithms, is verified numerically in Section **??**.

In this work, the objective $F$ of Eq. (4) is optimized by shaping an external electric field that interacts with an isolated quantum system, i.e. $\kappa \rightarrow \varepsilon(t)$. We consider a controllable quantum system of $N$ levels $|1\rangle, \ldots, |N\rangle$ whose dynamics are given by the time-dependent Hamiltonian $H(t) = H_0 - \mu \cdot \varepsilon(t)$, where $H_0$ describes the free dynamics of the system, $\mu$ is the dipole operator, and $\varepsilon(t)$ is the control field. To solve for optimal controls $\varepsilon(t)$ that maximize or minimize Eq. (3), it is useful to define a Lagrangian functional $\bar{J}$ that directly imposes the dynamical constraint in Eq. (1):

$$\bar{J} = \mathrm{Tr}(U(T,0)\rho_0 U^\dagger(T,0)\Theta) \; + \int_0^T \mathrm{Tr}\left[\phi^\dagger(t,0)\left(-i(H_0 - \mu\cdot\varepsilon(t))U(t,0) - \frac{dU(t,0)}{dt}\right)\right] \, dt \qquad (5)$$

where $\phi(t)$ is a Lagrange multiplier matrix function.

A necessary condition for solving problem of maximizing or minimizing (3) subject to the dynamical constraint is satisfaction of the first-order conditions (Euler-Lagrange equations) for the Lagrangian $\bar{J}$. The first Euler-Lagrange equation is simply the Schrödinger equation (1). To derive the remaining first-order conditions, it is convenient to integrate by parts and rewrite $\bar{J}$ in terms of the associated *PMP-Hamiltonian function* $\mathbf{H}$ [26, 27]:

$$\bar{J} = F(U(T,0)) - i\mathrm{Tr}(\phi^\dagger(T,0)U(T,0)) + i\mathrm{Tr}(\phi^\dagger(0,0)U(0,0)) + \int_0^T \mathbf{H}(U(t,0),\phi(t,0),\varepsilon(t)) + \mathrm{Tr}\left(\frac{d\phi^\dagger(t,0)}{dt}U(t,0)\right) \, dt, \qquad (6)$$

Here, the PMP-Hamiltonian is:

$$\begin{aligned}\mathbf{H}[U(t,0),\phi(t,0),\varepsilon(t)] &= -\langle\phi(t,0), iH_0U(t,0)\rangle + \varepsilon(t)\langle\phi(t,0), i\mu U(t,0)\rangle, \\ &= -\langle U^\dagger(T,0)\phi(T,0), iU^\dagger(t,0)H_0U(t,0)\rangle + \varepsilon(t)\langle U^\dagger(T,0)\phi(T,0), iU^\dagger(t,0)\mu U(t,0)\rangle,\end{aligned} \tag{7}$$

where $\langle A, B\rangle$ denotes the Hilbert-Schmidt inner product $\mathrm{Tr}(A^\dagger B)$,

$$\phi(T,0) = U(T,0)U^\dagger(t,0)\phi(t,0), \tag{8}$$

and

$$\frac{d\phi(t,0)}{dt} = -i(H_0 - \varepsilon(t)\cdot\mu)\phi(t,0), \tag{9}$$

which is the second Euler-Lagrange equation of (5). $\phi(T,0)$ must satisfy the following boundary condition

$$\phi(T,0) = \nabla_{U(T,0)}F(U(T,0)). \tag{10}$$

For $F(U)$ given by (4), we have [28]

$$\nabla_{U(T,0)}F(U(T,0)) = U(T,0)\left[U(T,0)\rho U^\dagger(T,0), \Theta\right]. \tag{11}$$

The third E-L equation (critical condition) is $\frac{\partial\mathbf{H}}{\partial\varepsilon(t)} = 0$. For control system satisfying Eq. (1),

$$\begin{aligned}\frac{\partial\mathbf{H}}{\partial\varepsilon(t)} &= -i\mathrm{Tr}\left(U^\dagger(T,0)\phi(T,0)U^\dagger(t,0)\mu U(t,0)\right) \\ &= -i\mathrm{Tr}\left(\left[U(T,0)\rho U^\dagger(T,0), \Theta\right]U^\dagger(t,0)\mu U(t,0)\right) = 0.\end{aligned} \tag{12}$$

Let us introduce the control-propagator map $V_T:\ \varepsilon(\cdot)\mapsto U(T)$ and the composition of maps $\tilde{J}\equiv F\circ V_T$. Then, the functional derivative $\frac{\delta\tilde{J}}{\delta\varepsilon(\cdot)}$ evaluated at time $t$, which we denote $\frac{\delta\tilde{J}}{\delta\varepsilon(t)} = \frac{\partial\mathbf{H}}{\partial\varepsilon(t)}$. For simplicity of exposition, we adopt the symbol $J$ for $\tilde{J}$ and refer simply to the derivative $\frac{\delta J}{\delta\varepsilon(t)}$. Control fields that satisfy Eqs. (1), (9), and (12) constitute the critical points of the control landscape $J(\varepsilon)$.

Assuming $S^{\varepsilon,T}$ is nonsingular [28] (i.e., that the matrix elements of $U^\dagger(t,0)\mu U(t,0)$ are linearly independent functions of time, a necessary condition for control optimality following from the third Euler-Lagrange equation is $[U(T,0)\rho U^\dagger(T,0),\Theta] = 0$, or equivalently that the time-evolved density matrix and initial observable operator commute. Since unitary evolution can only reorder the eigenvalues of $\Theta$,

we have $\tilde{\Theta}(T) = diag\{\sigma_{\Pi(1)}, \cdots, \sigma_{\Pi(N)}\}$, where $\Pi(i)$ denotes the $i$-th permuted eigenvalue of $\Theta(0)$. Denoting the eigenvalues of $\rho(0)$ by $\lambda_i$, the objective function values at the critical points are then $F = \mathrm{Tr}(\rho\Theta(T)) = \sum_{i=1}^{N} \lambda_i \sigma_{\Pi(i)}$. For full-rank, nondegenerate $\rho$ and $\Theta$, there are thus $N!$ critical values of the objective function.

## 2.2 Landscape Topology

For regular extremals, since the optimality status of critical points is equivalent to that of the critical points of the function $F(U)$ on $U(N)$, it is simplest to enumerate the number of positive, negative and null directions by examination of the Hessian quadratic form with $\delta U$ parametrized as $UA$, where $A$ is an arbitrary skew-Hermitian matrix. The vector $U(T)A$ in the tangent space $\mathcal{T}_{U(T)}U(N)$ induced by a control variation $\delta\varepsilon$ is $\int_0^T \frac{\delta U(T)}{\delta\varepsilon(t)} \cdot \delta\varepsilon(t)\, dt$, with associated Lie algebra direction $A = U^\dagger(T) \int_0^T \frac{\delta U(T)}{\delta\varepsilon(t)} \cdot \delta\varepsilon(t)\, dt$ in $u(N)$. It is straightforward to verify that $\frac{\delta U(T)}{\delta\varepsilon(t)} = U(T)\mu(t)$ **(ref)**. Thus, we may rewrite expression (20) for the second order variation as:

$$\mathcal{Q}(\delta U, \delta U) = \mathrm{Tr}\left\{A^2([\rho, U^\dagger\Theta U]_+ - 2A\rho A, U^\dagger\Theta U\right\} \tag{13}$$

Alternatively, using the exponential map to write $U + \delta U = Ue^{As}$,

$$\begin{aligned}
\mathcal{Q}(\delta U, \delta U) &= \left.\frac{\partial^2 F}{\partial s \partial s'}\right|_{s,s'=0} \\
&= \left.\frac{\partial}{\partial s'}\mathrm{Tr}\left\{A(\rho e^{-As'}U^\dagger\Theta U e^{As'} - e^{-As'}U^\dagger\Theta U e^{As'}\rho)\right\}\right|_{s'=0} \\
&= \mathrm{Tr}\left\{UA\left(\rho U^\dagger\Theta U A - U^\dagger\Theta U A\rho - \rho A U^\dagger\Theta U + U^\dagger\Theta U\rho A\right)\right\} \\
&= \mathrm{Tr}\left\{A^2\rho U^\dagger\Theta U - U^\dagger\Theta U A\rho A - A\rho A U^\dagger\Theta U + A^2 U^\dagger\Theta U\rho\right\} \\
&= \mathrm{Tr}\left\{A^2[\rho, U^\dagger\Theta U]_+ - 2A\rho A U^\dagger\Theta U\right\}.
\end{aligned}$$

This expression indicates that the kernel $\mathcal{H}(t, t')$ has finite rank. At the critical points, the HQF can be used to evaluate the rank of the Hessian and the optimality status. **We now expand this expression as a polynomial in $A_{ij}$ to obtain the optimality status of the critical points of the landscape...** At a critical point, $[\rho, \Theta(T)]_+ = 2\rho\Theta(T)$. The optimality status of any critical point corresponding to

the same critical value is the same. So, it is sufficient to consider critical points for which $\rho = \tilde{\rho}$ and $\Theta(T) = \tilde{\Theta}(T)$, where $\tilde{\cdot}$ denotes the operator in its diagonal basis. Then,

$$\begin{aligned}
\mathcal{Q}(\delta U, \delta U) &= \mathrm{Tr}\left\{A^2[\rho, U^\dagger \Theta U]_+ - 2A\rho A U^\dagger \Theta U\right\} \\
&= \mathrm{Tr}\left\{2A^2 diag\left(\lambda_1\sigma_{\Pi(1)}, \cdots, \lambda_N\sigma_{\Pi(N)}\right) + 2A diag\left(\lambda_1, \cdots, \lambda_N\right) A diag\left(\sigma_{\Pi(1)}, \cdots, \sigma_{\Pi(N)}\right)\right\} \\
&= 2\sum_i \lambda_i \sigma_{\Pi(i)} \sum_j |A_{ij}|^2 - 2\sum_i \sum_j |A_{ij}|^2 \lambda_i \sigma_{\Pi(j)} \\
&= 2\sum_i \sum_j (\lambda_i \sigma_{\Pi(i)} - \lambda_i \sigma_{\Pi(j)})|A_{ij}|^2 \\
&= 2\sum_i \sum_{j>i} (\lambda_i \sigma_{\Pi(i)} + \lambda_j \sigma_{\Pi(j)} - \lambda_i \sigma_{\Pi(j)} - \lambda_i \sigma_{\Pi(i)})|A_{ij}|^2 \\
&= 2\sum_i \sum_{j>i} (\lambda_i - \lambda_j)(\sigma_{\Pi(i)} - \sigma_{\Pi(j)})|A_{ij}|^2.
\end{aligned}$$

The number of positive, negative and zero eigenvalues for each critical manifold (i.e., each distinct permutation $\Pi$) can be read off the signs of the coefficients in this quadratic polynomial.

### 2.2.1 Hessian Rank Away From Critical Points

Away from the critical points, $\mathcal{H}(t, t')$ must be diagonalized directly to evaluate the Hessian rank; note that $\langle \nabla F(U(T)), \delta^2 U(T) \rangle$ does not appear in the second variation for regular control fields, but the term does contribute elements to the Hessian; i.e., the second variation is a quadratic form only at critical points. (Writing a matrix representation of the Hessian matrix (Hermitian two-form) of $F$ on $U(N)$ is inconvenient because it resides in the second-order tensor product of the cotangent space $T^{2*\otimes}_{U(T)}U(N)$.)

## 2.3 Algorithms

The majority of algorithms for quantum optimal control have been developed for cost functionals of the form $F(U(T)) + \int_0^T L(\varepsilon(t))\, dt$, which contain an auxiliary cost (e.g., the total field fluence, where $L(\varepsilon(t)) = -\frac{1}{2}\varepsilon^2(t)$). In these cases, so-called iterative algorithms [29] are typically used to obtain optimal controls, wherein an initial guess for $\varepsilon(t)$ is used to integrate the Schrödinger equation, with initial condition $U_0$, and the costate equation Eq. (9), with final condition $U^\dagger(T,0)\nabla_{U(T,0)}F(U(T,0))$;

these steps are iterated self-consistently [30], i.e.,

$$\frac{dU_k(t,0)}{dt} = -\frac{i}{\hbar}(H_0 - \varepsilon_k(t)\cdot\mu)U_k(t,0),\ U(0,0) = U_0$$
$$\frac{d\phi_k(t,0)}{dt} = \frac{i}{\hbar}(H_0 - \varepsilon_k(t)\cdot\mu)\phi_k(t,0),\ \phi_k(T,0) = \nabla_{U_k(T,0)}F(U_k(T,0))$$
$$\varepsilon_{k+1}(t) = -\frac{i}{\hbar}\mathrm{Tr}\left(U_k^\dagger(T,0)\nabla_{U_k(T,0)}F(U_k(T,0))U_k(^\dagger t)\mu U_k(t,0)\right),$$

until convergence. Since Eq. (12) only holds precisely at an optimal control $\bar{\varepsilon}(t)$, convergence is not guaranteed and additional adjustable parameters are sometimes added to these equations to ensure monotonic convergence, i.e., that $J_{k+1} - J_k \geq 0,\ \forall k$ [29, 30]. The above approach is used extensively in the literature for quantum state control [3], where the convergence tolerance is large enough that competition between the endpoint $F(U(T))$ and Lagrange $\int_0^T L(\varepsilon(t))\ dt$ terms does not compromise the fidelity of the endpoint objective at the global optimum.

Let $s$ parametrize the algorithmic time evolution of the search trajectory of any algorithm that seeks to maximize $F(U(T,0))$. The simplest first-order algorithm is the gradient flow of the objective function; the gradient flow trajectory is the solution $\varepsilon(s,t)$ to the initial value problem

$$\frac{\partial\varepsilon(s,t)}{\partial s} = \alpha(s)\ \frac{\delta J}{\delta\varepsilon(s,t)}, \tag{14}$$

for a specified initial guess for the control $\varepsilon(0,t)$, where $\alpha(s)$ is an adaptive step size. In the numerical simulations in this work, Eq. (14) will be solved using a variable step size fourth order Runge-Kutta integrator built into MATLAB [31]. Our primary concern in this paper is the convergence rate of such algorithms. In Section **??**, we will demonstrate that local suboptima for the endpoint cost defined in Eq. (4) have null measure, guaranteeing convergence of such gradient algorithms. In this paper we focus on the scaling of search effort required for control optimization using gradient algorithms. Recently, such algorithms have been successfully implemented in the laboratory for maximization of second harmonic generation (SHG) **(ref)**. In the majority of laboratory adaptive feedback strategies, genetic or evolutionary algorithms are used [7, 9, 10, 14, 17, 18, 19, 20, 21, 22, 23, 24]. Although the convergence of gradient algorithms is guaranteed for unconstrained controls, in the laboratory, the constraints of finite available bandwidth and a discrete set of control parameters may cause the underlying

trap-free character of quantum control landscapes. Nonetheless, the trajectories of quantum control gradient flows play a dominant role in determining the convergence rates of evolutionary search algorithms (some of which explicitly use first- or second-order information **(ref)**), and these trajectories may be analytically characterized.

# 3 Landscape Structure Metrics

**The distance traveled over the control landscape is intuitively related to the complexity of a search trajectory. The search is parameterized in terms of the algorithmic index $s$, which progresses in $m$ steps of size $\Delta s$ from $s = 0$ initially to $s = s_f$ at the final optimized field. The Euclidean or straight-line distance between $\epsilon(0,t)$ and $\epsilon(s_f,t)$ is**

$$||\Delta_E \epsilon(t)|| = ||\epsilon(s_f,t) - \epsilon(0,t)|| = \left( \int_0^T \left[\epsilon(s_f,t) - \epsilon(0,t)\right]^2 dt \right)^{1/2} \approx \left( \Delta t \sum_{j=0}^{n-1} \left[\epsilon(s_f, j\Delta t) - \epsilon(0, j\Delta t)\right]^2 \right)^{1/2} \tag{15}$$

**where the control field is divided into $n$ steps of size $\Delta t$. The path length, or the actual distance of the search trajectory from $\epsilon(0,t)$ to $\epsilon(s_f,t)$, is obtained by integrating over the entire trajectory:**

$$||\Delta_P \epsilon(t)|| = \int_0^{s_f} \left( \int_0^T \left( \frac{\partial \epsilon(s,t)}{\partial s} \right)^2 dt \right)^{1/2} ds \approx \sum_{k=0}^{m-1} \left( \Delta t \sum_{j=0}^{n-1} \left[\epsilon((k+1)\Delta s, j\Delta t) - \epsilon(k\Delta s, j\Delta t)\right]^2 \right)^{1/2} \tag{16}$$

**The ratio $R_\epsilon = \frac{||\Delta_P \epsilon(t)||}{||\Delta_E \epsilon(t)||}$ represents the directness of the search path and is one measure of search complexity.**

**Additional local structure metrics arise from the terms of a Taylor expansion of the objective functional $J$ with respect to the field:**

$$J[\epsilon(t) + \delta\epsilon(t)] = J[\epsilon(t)] + \int_0^T \nabla_\epsilon J[\epsilon(t)]\ \delta\epsilon(t)\ dt + \frac{1}{2} \int_0^T \int_0^T \mathcal{H}(t,t')\ \delta\epsilon(t)\ \delta\epsilon(t')\ dt\ dt' + \ldots \tag{17}$$

**The first-order term yields the slope metric $\mathcal{G}_k$, defined at a point $s_k$ on the search trajectory as**

$$\mathcal{G}_k = ||\nabla J|_k|| = \left( \int_0^T \left[ \frac{\delta J}{\delta \epsilon(s_k, t)} \right]^2 \right)^{1/2} \approx \left( \Delta t \sum_{j=0}^{n-1} \left[ \frac{\partial J}{\partial \epsilon(k\Delta s, j\Delta t)} \right]^2 \right)^{1/2} \tag{18}$$

**$\mathcal{G}_k$ is equal to the magnitude of the gradient at the point** $s_k$, and is uniformly bounded by a constant that scales linearly with the Hilbert space dimension. This bound indicates that the quantum control landscape has "smooth slopes" **(ref)**. To derive this bound, we apply the triangle and Cauchy-Schwarz inequalities in succession to expression (12):

$$\begin{aligned}
\left| \frac{\delta J}{\delta \varepsilon(t)} \right| &= \left| \mathrm{Tr} \left\{ [U(T,0)\rho U^\dagger(T,0), \Theta] \mu(t) \right\} \right| \\
&= \left| \mathrm{Tr} \left( U(T,0)\rho U^\dagger(T,0)\Theta\mu(t) \right) - \mathrm{Tr} \left( \Theta U(T,0)\rho U^\dagger(T,0)\mu(t) \right) \right| \\
&\leq \left| \mathrm{Tr}(U(T,0)\rho U^\dagger(T,0)\Theta)\mu(t) \right| + \left| \mathrm{Tr}(\Theta U^\dagger(T,0)\rho U(T,0))\mu(t) \right| \\
&= ||U(T,0)\rho U^\dagger(T,0)\Theta\mu(t)|| + ||\Theta U^\dagger(T,0)\rho U(T,0))\mu(t)|| \\
&\leq ||U(T,0)\rho U^\dagger(T,0)\Theta||||\mu(t)|| + ||\Theta U^\dagger(T,0)\rho U(T,0))||||\mu(t)|| \\
&\leq ||\rho||\,||\Theta||\,||\mu(t)|| + ||\Theta||\,||\rho||\,||\mu(t)|| \\
&= 2\,||\Theta||\,||\mu||. \quad \textbf{N absent from expression?}
\end{aligned}$$

**Large values of $\mathcal{G}_k$ should indicate that the algorithm is ascending the landscape more rapidly in the region of $s_k$ and therefore studying $\mathcal{G}_k$ at the initial point(s) of the search trajectory may aid in the prediction of overall search effort.**

**The second-order term in Equation 17 contains the second-derivative or Hessian matrix.** Thus, it will be useful to have an explicit expression for the Hessian kernel $\mathcal{H}(t, t')$. For concision, we will adopt the notation $U(T) \equiv U(T, 0)$. The second-order variation of $J$ may be written **(ref)**

$$\delta^2 J = \mathcal{Q}_F(\delta U(T), \delta U(T)) + \langle \nabla F(U(T)), \delta^2 U(T) \rangle. \tag{19}$$

where at critical points $\mathcal{Q}_F(\delta U(T), \delta U(T))$ denotes the Hessian quadratic form (HQF). It can be shown

**(ref)** that for $F(U(T))$ given by (4), this reduces to

$$\int_0^T \int_0^T \delta\varepsilon(t)\cdot \mathrm{Tr}\left\{\rho\Theta(T)\mu(t')\mu(t) + \Theta(T)\rho\mu(t)\mu(t') - \mu(t)\rho\mu(t')\Theta(T) - \mu(t')\rho\mu(t)\Theta(T) + [\rho,\Theta(T)][\mu(t),\mu(t')]\right\}\cdot\delta\varepsilon(t')\,dt\,dt \tag{20}$$

So, the Hessian symmetric kernel $\mathcal{H}(t,t')$ of the linear integral operator associated with $\delta^2 J$ is

$$\mathcal{H}(t,t') = \mathrm{Tr}\left\{\rho\Theta(T)\mu(t')\mu(t) + \Theta(T)\rho\mu(t)\mu(t') - \mu(t)\rho\mu(t')\Theta(T) - \mu(t')\rho\mu(t)\Theta(T) + [\rho,\Theta(T)][\mu(t),\mu(t')]\right\}, \tag{21}$$

which we will approximate near critical points by

$$\mathcal{H}(t,t') \approx \mathrm{Tr}\left\{\rho\Theta(T)[\mu(t),\mu(t')]_+ - \Theta(T)(\mu(t)\rho\mu(t') + \mu(t)\rho\mu(t'))\right\} \tag{22}$$

where we have used the critical point condition $[\rho,\Theta(T)] = 0$ and where $[,]_+$ denotes the symmetric commutator (anticommutator).

Several important second-order metrics are derived from the Hessian. The Hessian trace or mean curvature $\mathrm{Tr}(\mathcal{H}) = \sum_{i=0}^{n}\mathcal{H}(t_i,t_i)$ is the first of these. The norm of the Hessian matrix is another local metric for assessing optimization search effort using first-order algorithms. The bound on its elements is

$$\begin{aligned}
||\mathcal{H}(t,t')|| &= |\mathrm{Tr}\left\{\rho\Theta(T)\mu(t')\mu(t) + \Theta(T)\rho\mu(t)\mu(t') - \mu(t)\rho\mu(t')\Theta(T) - \mu(t')\rho\mu(t)\Theta(T) + [\rho,\Theta(T)][\mu(t),\mu(t')]\right\}| \\
&= ||\rho\Theta(T)\mu(t')\mu(t) + \Theta(T)\rho\mu(t)\mu(t') - \mu(t)\rho\mu(t')\Theta(T) - \mu(t')\rho\mu(t)\Theta(T) + [\rho,\Theta(T)][\mu(t),\mu(t')]|| \\
&\leq 2\,||\rho\Theta(T)\mu(t)\mu(t')|| + 2\,||\mu(t)\rho\mu(t')\Theta(T)|| + 4\,||\rho\Theta(T)\mu(t)\mu(t')|| \\
&\leq 8\,||\rho||\,||\Theta||\,||\mu||^2 \\
&\leq 8\,||\Theta||\,||\mu||^2
\end{aligned}$$

and the norm of the Hessian matrix is therefore

$$||\mathcal{H}|| = \mathbf{??} \tag{23}$$

At critical points,

$$||\mathcal{H}(t,t')|| \leq 4\,||\rho||\,||\Theta||\,||\mu||^2 \leq 4\,||\Theta||\,||\mu||^2.$$

However, the control optimization gradient flows have global dynamics that cannot be assessed using such norms. In particular, these flows may be attracted toward critical points of the landscape. In order to assess these global properties of the flow, **a metric representing the distance to the nearest saddle from the current point on the search trajectory was developed. This metric makes use of the critical properties of the transition matrix $\mathcal{T} = U(T,0)^* \circ U(T,0)$, which is the Hadamard (entrywise) product of the unitary propagator with its complex conjugate. We assume that only $n_\rho$ and $n_O$ of the eigenvalues of $\rho$ and $O$, respectively, are nonzero and nondegenerate while the remainder of both spectra are zero, and define $n_< = \min[n_\rho, n_O]$. $\{\rho_p\}_<$ and $\{O_q\}_<$ are the subsets of the $n_<$ largest eigenvalues of each operator. If $\rho$ and $O$ are both full-rank matrices, $\mathcal{T}$ is a permutation matrix at landscape critical points. When either is rank-deficient, only the rows (columns) of $\mathcal{T}$ corresponding to the elements of $\{\rho_j\}_<$ ($\{O_k\}_<$) resemble those of a permutation. The saddle distance metric at the point $s_k$ on the search trajectory is**

$$\mathcal{S}_k = \mathcal{N}\left(n_< - \sum_{\{\rho_p\}_<} \langle p|\mathcal{T}_k\mathcal{T}_k^\dagger|p\rangle\right) = \mathcal{N}\left(n_< - \sum_{\{O_q\}_<} \langle q|\mathcal{T}_k^\dagger\mathcal{T}_k|q\rangle\right) \tag{24}$$

**where the sums are over all indices $p$ ($q$) such that $\rho_p$ ($O_q$) is an element of $\{\rho_p\}_<$ ($\{O_q\}_<$). The normalization $\mathcal{N} = \frac{N}{n_<(N-1)}$ guarantees that $\mathcal{S}_k$ is always $\leq 1$, since the matrix elements of $\mathcal{T}\mathcal{T}^\dagger$ and $\mathcal{T}^\dagger\mathcal{T}$ can be rewritten as**

$$\langle p|\mathcal{T}_k\mathcal{T}_k^\dagger|p\rangle = \sum_{p'=1}^{N} \langle p|\mathcal{T}_k|p'\rangle^2 \qquad \langle q|\mathcal{T}_k^\dagger\mathcal{T}_k|q\rangle = \sum_{q'=1}^{N} \langle q'|\mathcal{T}_k|q\rangle^2 \tag{25}$$

**and are therefore bounded on the interval $[\frac{1}{N}, 1]$.**

**Unlike the Hessian, the $N^2 \times N^2$ Gramian matrix**

$$G_{\epsilon,T} = \int_0^T v[U(T)\mu(t)]\; v[\mu(t)U^\dagger(T)]^T\; dt \tag{26}$$

**(where $v[\cdot]$ represents the vectorization of an $N\times N$ complex matrix into a $N^2$-length complex vector) characterizes solely dynamic effects on the search trajectory. Its eigenvectors**

**are $N^2$ orthogonal directions in the tangent space to $U(N)$, and the corresponding eigenvalues are a measure of each direction's contribution to the variation $\delta U(T)$. If its eigenvalues are all far from zero then $G_{\epsilon,T}$ is *well-conditioned* and it follows that the gradient norm can only grow infinitesimally small at the global optimum of the control landscape. A well-conditioned Gramian corresponds to a search trajectory dominated by kinematic rather than dynamic effects and is expected to correlate to more rapid optimization.**

**All the previous metrics address local properties of the control lanscape, but its global character can be assessed by examining the Dyson expansion for the unitary propagator in the interaction representation:**

$$U_I(T) = I + \frac{i}{\hbar}\int_0^T V^\dagger(t)\mu V(t)\epsilon(t)dt \; + \; \left(\frac{i}{\hbar}\right)^2 \int_0^T V^\dagger(t)\mu V(t)\epsilon(t)\int_0^t V^\dagger(t')\mu V(t')\epsilon(t')dt'dt \; + \; \ldots \quad (27)$$

**where $V(t) = \exp(-iH_0t)$. The $n$th term in the expansion corresponds to all possible $n$-photon transitions between eigenstates over the interval $[0, T]$, and higher-order terms require higher field amplitudes in order to be nonnegligible. Since systems with sparser dipole matrices are more likely to require high-order terms to achieve an arbitrary unitary propagation $U(T)$, these systems are likely to be characterized by higher field fluences and greater computational effort.**

# 4 Control Systems

Our generic quantum systems are defined as flexibly as possible in order to permit full exploration of the variety possible in the structure of the observable control landscape. In order to relate these systems to common targets in quantum control experiments, three classes of quantum systems corresponding to spectroscopic objectives will also be studied. A rovibrational model combines the two approximations. The field-free Hamiltonian, transition dipole matrix, initial density operator, and target observable for each are described. In these latter cases the electric field is parameterized in a manner identical to that described in Section 4.1 for the generic dynamic simulations.

## 4.1 Generic Systems

We first study generic $N$-level quantum systems whose field-free Hamiltonian $H_0$ is diagonal:

$$H_0 = \frac{1}{10}\sum_{j=1}^{N}\left(1 + j(j+1)\right)|j\rangle\langle j| \tag{28}$$

Several distinct dipole coupling schemes were employed to illustrate the effect of Hamiltonian structure on local features of the control landscape. Remote states are coupled to one another less strongly than adjacent states in many physical systems and this is reflected in the first dipole structure selected: $\langle j|\mu|k\rangle = D^{|j-k|-1} \ \forall \ j \neq k$ and zero otherwise, where $D < 1$ is a scalar coupling factor. The effect of this "exponential" dipole on the landscape structure is contrasted with a dipole matix where all non-diagonal elements are equal: $\langle j|\mu|k\rangle = D \ \forall \ j \neq k$.

The control electric field is parameterized in the form

$$\epsilon(t) = A\exp\left(-\frac{(t-\frac{T}{2})^2}{2\sigma^2}\right)\sum_p cos(\omega_p t) \tag{29}$$

where different simulations vary the amplitude factor $A$ and the number and magnitude of the frequency components $\{\omega_p\}$.

Both $\rho \equiv \rho_0$ and $O$ are assumed to be mutually diagonalizable and are constructed as such. $Tr(\rho_0) = 1$ while $Tr(O) = n_O$. Recalling that $\rho$ and $O$ have $n_\rho$ and $n_O$ nonzero eigenvalues, respectively, we specify that the nonzero eigenvalues of $\rho$ correspond to the first $n_\rho$ levels of the system while the nonzero eigenvalues of $O$ are associated with levels $n_\rho + 1$ to $n_\rho + n_O$. Thus if both $\rho$ and $O$ have 2 nonzero eigenvalues:

$$\rho = \begin{pmatrix} \rho_1 & 0 & 0 & 0 & \cdots \\ 0 & \rho_2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad O = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & O_1 & 0 & \cdots \\ 0 & 0 & 0 & O_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{30}$$

## 4.2 Rigid Rotor

We'll treat the diatomic molecule as a rigid rotor with reduced mass $m$ and interatomic distance $r$. In the absence of the electric field, the system follows the Schrodinger equation for a free particle with mass $m$

$$\frac{-\hbar^2}{2m}\nabla^2\Psi(\vec{r}) + V\Psi(\vec{r}) = E\Psi(\vec{r}) \tag{31}$$

where the potential $V$ is zero due to the rigidity. Its eigenfunctions are the spherical harmonics $Y_J^{|M|}(\theta,\phi)$:

$$\Psi_J(\theta,\phi) = Y_J^{|M|}(\theta,\phi) = \sqrt{\frac{2J+1}{4\pi}\frac{(J-|M|)!}{(J+|M|)!}}\, e^{i|M|\phi}\, P_J^{|M|}(\cos\theta) \tag{32}$$

where $P_J^{|M|}$ is the associated Legendre polynomial.The field-free Hamiltonian $H_0$ is diagonal and its eigenvalues correspond to the energy levels of the system

$$H_0 = \frac{\hbar^2}{2I}\sum_{J=0}^{N-1} J(J+1)|J\rangle\langle J| \tag{33}$$

where $I = mr_e^2$ is the molecular moment of inertia. The selection rules for the system are $\Delta J = \pm 1$, $\Delta M^z = 0$, $\Delta M^{x,y} = \pm 1$. We assume that the applied electric field has only a $z$-component and that the $M$ states are degenerate in energy, i.e., that the Stark effect is negligible. Thus, these states need not be incorporated into our control model. The transition dipole matrix elements are

$$\langle JM|\mu_z|(J+1)M\rangle = \mu_{perm}\sqrt{\frac{(J+M+1)(J-M+1)}{(2J+1)(2J+3)}} \tag{34}$$

where $\mu_{perm}$ is the permanent dipole moment of the molecule.

**Both the target observable $O$ and the initial density matrix $\rho$ are in their diagonal bases. For the former we choose the simple form of a projector $O = |J'\rangle\langle J'|$ onto a target final**

**state $|J'\rangle$. The structure of the density matrix, however, reflects a thermal distribution of states dictated by the Boltzmann relation, where the elements of $\rho$ correspond to the initial fractional population of the state $|J\rangle$:**

$$\rho = \left[ \frac{g_J \exp\left(\frac{-E_J}{k_B T}\right)}{\sum_i g_J \exp\left(\frac{-E_l}{k_B T}\right)} \right] |J\rangle\langle J| \tag{35}$$

**The degeneracy $g_J = 2J + 1$ due to the $M$ states. It is important to note that since $E_J = \langle J|H_0|J\rangle$, the structures of the operators $H_0$ and $\rho$ are explicitly coupled in all spectroscopically motivated systems. Thus the magnitude of the $\rho$ elements varies significantly between different model molecules, for example CO and HCl:**

**The asymptotic behavior of Equation 35 indicates that in principle, all states possess at least some initial population even at low temperatures. However, the population of many high-energy states is small enough to be insignificant. The cutoff is determined by defining a fraction $\sigma$ of the population of the most populated state $J_{max}$; states with a smaller initial population than this are considered unpopulated for the purposes of the simulation. The number of significantly populated states depends on both $H_0$ and temperature, and constitutes a realistic lower bound on the Hilbert space dimension $N$ of the model system:**

**Here the black, red, and blue lines represent $\sigma = 10^{-3}$, $10^{-5}$, and $10^{-7}$ respectively. At modest temperatures, $N$ need be no larger than in the majority of generic simulations. In this paper we set $\sigma = 10^{-5}$.**

### 4.3 Morse Oscillator

The system again obeys Equation 31 except that the potential is no longer zero and only the internuclear coordinate $r$ is relevant. The Morse potential $V(r-r_e) = D_e\left(1-e^{-\beta(r-r_e)}\right)^2$, an excellent approximation of the ground electronic state behavior of a vibrating diatomic molecule, is employed. $D_e$ is the dissociation energy, $\omega_0$ is the characteristic angular frequency of the oscillator, and $\beta = \omega_o\sqrt{\frac{m}{2D_e}}$. Its eigenfunctions are

$$\Psi_\nu(r-r_e) = \sqrt{\frac{\beta(2K-2\nu)\nu!}{\Gamma(2K-\nu+1}}\, e^{-z/2}\, z^{K-\nu}\, L_\nu^{(2K-2\nu)}(z) \tag{36}$$

where $(K+1/2)^2 = 2mD_e/\hbar^2\beta^2$, $z = (2K+1)e^{-\beta(r-r_e)}$, and the $L_\nu^{(\alpha)}$ are the generalized Laguerre polynomials. $H_0$ is again in its diagonal basis with elements corresponding to the system energy levels:

$$H_0 = \sum_{\nu=0}^{N-1}\left[\hbar\omega_o\left(\nu+\frac{1}{2}\right) - \frac{\hbar^2\omega_o^2}{4D_e}\left(\nu+\frac{1}{2}\right)^2\right]|\nu\rangle\langle\nu| \tag{37}$$

Unlike the rigid rotor model, the Morse potential exhibits no strict selection rule; transitions $\Delta\nu = \pm2, \pm3, \ldots$ are only weakly forbidden. The simplest method to compute the dipole matrix elements is direct integration over the Morse eigenfunctions: $\langle\nu|\mu(r)|\nu'\rangle = \int \Psi_\nu\, \mu(r)\, \Psi_{\nu'}\, dr$. An equivalent analytical solution has been obtained [33] for elements of the dipole function $\mu(r) = q(r-r_e+x_0)\exp(-(r-r_e)/x_1)$:

$$\langle\nu|\mu(r)|\nu'\rangle = \frac{\mu_{perm}\mathcal{N}_\nu\mathcal{N}_{\nu'}}{\beta r_e}\frac{\Gamma(2K-2\nu'+1)\,\Gamma(2K-2\nu+1)}{(2K+1)^\xi\,\Gamma(2K-\nu'+1)}\sum_{n=0}^{\nu}A_n \tag{38}$$

$$\begin{aligned}A_n = {} & \frac{(-\nu)_n}{n!}\frac{\Gamma(2K-\nu-\nu'+n+\xi)\Gamma(\nu-n+1-\xi)}{\Gamma(2K-2\nu+n+1)\Gamma(\nu-\nu'-n+1-\xi)} \\ & \times\left[\psi(\nu-n+1-\xi)-\psi(2K-\nu-\nu'+n+\xi)-\psi(\nu-\nu'-n+1-\xi)+\ln(2K+1)+\beta x_0\right]\end{aligned} \tag{39}$$

$$\mathcal{N}_\nu = \sqrt{\frac{(2K-2\nu)\Gamma(2K-\nu+1)}{\nu!\,\Gamma(2K-2\nu+1)^2}} \tag{40}$$

where $\xi = 1/\beta x_1$ and $(-\nu)_n = \Gamma(-\nu + n)/\Gamma(-\nu)$ is the Pochhammer symbol. The parameters $x_0$ and $x_1$ can be fit to experimental data.

**As in Section 4.2, the observable $O = |\nu\rangle\langle\nu'|$ is a pure-state projector. $\rho$ is obtained using Equation 35, with the modification that $g_\nu = 1$. Due to the larger energy spacing in vibrational spectroscopy, far fewer vibrational states are thermally populated at a given temperature than in the rotational model. This disparity between the rank-deficiency of $\rho$ in the vibrational and rotational cases will be shown to directly influence the local landscape structure. Furthermore the Morse oscillator, unlike the rigid rotor, has a finite number of bound states equal to the integer part of $K$. As molecular dissocation is not incorporated into this model, $N \ll K$ to ensure that the unitary evolution of the system does not impinge on the dissociative regime. This will be verified numerically.**

## 4.4 Rotating Morse Oscillator

**In the approximation employed in this work, the eigenfunctions of the rotating oscillator are be separated into a product of the rotational and vibrational eigenfunctions previously derived for the rigid rotor and the Morse oscillator:**

$$\Psi_{\nu J}(r, \theta, \phi) = \frac{1}{r}\, \Psi_J(\theta, \phi)\, \Psi_\nu(r - r_e) \tag{41}$$

**Consequently the energy of the rotating Morse oscillator can be approximated as the sum of the rotational and vibrational energies. The model quantum system will incorporate multiple vibrational states and their rotational substates. The number of rotational sublevels is infinite, but the strict selection rule described in Section 4.2 indicates that states whose rotational quantum number is much larger than that of the highest initially-populated state are very unlikely to become populated during the unitary system evolution at each search step. Therefore we approximate the infinite rotational sublevels with the $N_J$ lowest-energy sublevels, where $N_J \gg n_\rho$. The field-free Hamiltonian is**

$$H_0 = \sum_{\nu=0}^{N_\nu - 1} \sum_{J=0}^{N_J - 1} \left[ \frac{\hbar^2}{2I} J(J+1) + \hbar\omega_o \left(\nu + \frac{1}{2}\right) - \frac{\hbar^2 \omega_o^2}{4D_e} \left(\nu + \frac{1}{2}\right)^2 \right] |J\nu\rangle\langle J\nu| \tag{42}$$

**where our model includes $N_\nu$ vibrational states and the overall Hilbert space dimension of the system is $N = N_J \times N_\nu$.**

**Equation 41 implies that the transition dipole moment is also separable, so both selection rules separately observed observed for the rotor and oscillator models ($\Delta J = \pm 1$, $\Delta\nu = \pm 1, \pm 2, \pm 3, \ldots$) are obeyed. As in the case of the rigid rotor, we assume that the electric field has only a $z$ component:**

$$\langle \nu J | \mu(r) | \nu' J' \rangle = \int \Psi^*_{\nu J}\ \mu_z(r,\theta)\ \Psi_{\nu' J'}\ d\tau \tag{43}$$

$$= \int \frac{1}{r^2}\ \Psi_\nu\ \Psi^*_J\ \mu_z(r,\theta)\ \Psi_{\nu'}\ \Psi_{J'}\ d\tau \tag{44}$$

**And since $\mu_z(r,\theta) = \mu(r)\cos\theta$ and $d\tau = r^2 \sin\theta\ d\theta\ d\phi$:**

$$\langle \nu J | \mu(r) | \nu' J' \rangle = \int \mu(r)\ \Psi_\nu\ \Psi_{\nu'} dr \quad \iint \sin\theta\ \cos\theta\ \Psi^*_J\ \Psi_{J'}\ d\theta\ d\phi \tag{45}$$

**Thus the rovibrational dipole matrix elements can be directly computed using the rigid rotor and Morse eigenfunctions described in previous sections.**

**In this model, our objective is population transfer into a target vibrational state $|\nu'\rangle$ without regard to the ultimate rotational excitation; correspondingly, the observable operator has the form**

$$O = \sum_{J=0}^{N_J - 1} |J\nu'\rangle\langle J\nu'| \tag{46}$$

**while the density matrix again follows Equation 35 with $g_{J\nu} = 2J + 1$.**

## 4.5 Spectral Notation

Optimizations employing the generic Hamiltonians require a choice of structure for both $\rho$ and $O$, while in the spectroscopically-motivated quantum systems the structure of $\rho$ depends on the Hamiltonian itself. However, in both cases the spectral details of these operators can be complex. In this paper, therefore, we employ a shorthand notation to denote the number and degeneracy of the eigenvalues of $\rho$ and $O$. Only the nonzero elements of $\rho$ and $O$ are specified, with the remainder of each spectrum implicitly understood to be zero. In the $(m_i : n_j)$ notation, the index $m$ represents the number of distinct nonzero eigenvalues of the density matrix, and $i$ is the degree of degeneracy of these eigenvalues. Similarly, $n$ and $j$ respectively represent the number and degeneracy of the distinct nonzero eigenvalues of the observable. For example, a 10-level system with a $(3_3 : 4_2)$ structure has a density matrix with three distinct nonzero eigenvalues (each of degeneracy 3) and one zero eigenvalue; the target observable has four nonzero eigenvalues (each of degeneracy 2), and one doubly-degenerate zero eigenvalue. Although this notation only encompasses spectral structures in which all nonzero eigenvalues are of identical degeneracy, it can be generalized to include all possible spectral structures. In a $(1_6 2_2 : 2_1 1_4)$ system, for example, the density matrix possesses one six-fold degenerate nonzero eigenvalue and two doubly degenerate nonzero eigenvalues, while the target observable possesses two nonzero, nondegenerate eigenvalues and one four-fold degenerate nonzero eigenvalue. For the great majority of simulations, however, the behavior of the $(m_i : n_j)$ subclass was found to capture the essential features of the search effort scaling.

## 4.6 Classification of Control Systems

Many of the landscape features that directly affect the behavior of a search trajectory over the observable control landscape are related to the spectral details of the operators $\rho$ and $O$. The generic and molecular control systems described in the previous sections are selected to sample a wide range of physically relevant spectral structures from the near-infinite set of possible systems. For the purposes of the present analysis, we divide the observable simulations into three categories according to their predicted behavior with respect to the required computational resources for optimizations as well as the metrics

in Section 3. The simplest observable simulations involve $\rho$ and $O$ that are both pure-state projectors, corresponding to optimizing a transition between two system states. The generic $1_1 : 1_1$ systems fall into this category, as does the Morse model system at temperatures low enough to limit all initial population to the ground state. Since the nonzero spectral components are invariant to $N$, computational effort is not expected to scale with system dimension. The second category of observable control system include all $\rho$ and $O$ that possess some degree of degeneracy but are not pure-state projectors. Since multiple rotational levels are thermally populated at all but the lowest temperatures, the rigid rotor and rotating Morse models fall into this category, as well as $(m_i : n_j)$ systems where $m, n < N$. The third category correspond to full rank $\rho$ and $O$; since no common spectroscopic model involves such spectral structures, generic $N_1 : N_1$ systems were introduced. Scaling of computational effort with $N$ is expected for these systems. Subsequently these three categories are referred to as A, B, and C respectively.

As described in Section 3, a well-conditioned Gramian matrix is expected to correspond to optimization trajectories in which dynamic effects are minimized and convergence is achieved quickly. This hypothesis is tested by examining systems from class C with dipole matrices of differing sparsity since this is connected to Gramian condition number. However, this correlation may be attenuated by the presence of significant degeneracy in the operators $\rho$ and $O$. Quantum systems in which $J$ depends on only a portion of the propagator $U(T)$ may be partially or entirely insensitive to the spectral structure of the Gramian. Study of class A control systems should indicate whether this is also true, while study of the various class B systems may show whether specific spectral characteristics of $\rho$ and $O$ cause the search trajectories to approach either extreme.

# 5 Computational Effort and Resource Scaling

## 5.1 Generic Systems

The mean search efforts reported arise from sample sizes of 20 optimizations; among the optimizations in each sample, $H_0$, $\mu$, field fluence and bandwidth, and the spectral structure of $\rho$ and $O$ are the same. The initial field amplitudes and frequencies and the eigenvalues of $\rho$ and $O$ are generated quasi-randomly

in order to abide by the above constraints.

### 5.1.1 Degenerate Cases

The dynamic optimizations of systems in which $\rho$ and $O$ possess significant degeneracy took place in several steps, over which the algorithm and physical parameters of the optimization were varied. Initial optimizations were performed using a simple fixed step RK4 algorithm with $T = 300$, 2048 time points, initial field fluence = 10, bandwidth = [0,1], and dipole dropoff rate $D = 0.9$ on $(m_1 : n_1)$ spectral structures for $N =$ multiples of 5 up to $N = 40$, and only for $1 \leq m, n \leq 3$. The limitations of the fixed step-size algorithm made it impractical to proceed to more complex spectral structures. However, flat scaling with respect to $N$ is clearly observed and the correlation between search effort and the parameter $\min[m, n]$ (also observed in kinematic simulations) is apparent as well. As the vertical axis is logarithmic, the search effort appears to grow quadratically or exponentially with $\min[m, n]$.

Optimizations were also performed using the variable step RK4 algorithm, similar to those above but for $1 \leq m, n \leq 6$, and $N \leq 20$. The results are comparable to the fixed step simulations except that the correlation between search effort and $\min[m, n]$ appears approximately linear.

The most recent variable step simulations vary from those above in two significant respects; the bandwidths of the initial fields correspond to the largest transition frequencies possible given the spectral structure of $\rho$ and $O$, and the trace of the observable is normalized to $n$ (the number of nonzero eigenvalues in $O$) rather than to 1. The final time for each optimization was also reduced to $T = 100$ to better resolve the higher field frequencies and permit larger systems to be optimized. However, they exhibit the same scaling and relative effort as the other variable step studies.

A number of variantions on this last series of simulations have yielded the same qualitative dependence of search effort on the structure of $\rho$ and $O$. Imparting a random sign to each of the dipole elements:

and rendering all dipole matrix elements identical (rather than allowing them to decrease exponentially with distance from the main diagonal):

all yielded virtually the same scaling behavior and relative search effort. Limiting the bandwidth of the initial electric field to the interval [0,1] rather than an interval of frequencies equal in energy to the largest transition between nonzero eigenvalues of $\rho$ and $O$:

yielded similar qualitative behavior, but unsurprisingly the search effort for the less-degenerate cases was somewhat increased as the initial field lacked frequencies capable of stimulating key system transitions. Similar overall increases in search effort were observed when a varying proportion of dipole elements were reduced to zero; however, elimination of 25% (left) or 50% (right) of the dipole elements

do not greatly alter the overarching dependence of search effort on the spectra of $\rho$ and $O$. Doubling the number of nonzero eigenvalues in the spectra of $\rho$ and $O$ by rendering each eigenvalue twofold degenerate increases the search effort:

However, it appears from this limited sample of simulations that the resulting effort is less than what would be observed when optimizing systems whose spectra include an equivalent number of nondegenerate eigenvalues.

All the simulations indicate that invariance of the nonzero parts of the $\rho$ and $O$ spectra to $N$ eliminates all scaling, and that the qualitative behavior of these degenerate cases is quite robust with respect to changes in the physical parameters involved. Thus, the conclusions drawn in this section will focus on the factors that influence the relative search effort of these cases. The spectral structure of $\rho$ and $O$ plays a role very similar to the kinematics, but the nature of the correlation depends on the algorithm used. Other physical parameters such as the bandwidth of the initial field and the magnitude of the $O$ eigenvalues have no noticeable effect. In particular, the parameters specific to dynamic simulations play a minor role at most, suggesting that the search effort corresponding to these systems is almost wholly dictated by kinematic factors.

### 5.1.2 Nondegenerate Cases

In light of the above simulations' similarity to the previous kinematic study, it seems likely that systems in which the non-trivial spectral elements of $\rho$ and $O$ *do* depend on $N$ will exhibit search effort scaling of some kind. Simulations in which both $\rho$ and $O$ are fully nondegenerate were performed to explore this issue. As in the previous section, the dependence of this scaling on various physical parameters will be explored. No optimizations employing the fixed step algorithm were performed due to the greater computational demands of the nondegenerate spectral structures.

In order to determine the effect of the dipole structure on the scaling of these nondegenerate systems, optimizations using both the flat and exponential dipole types described earlier were performed, varying the dipole parameter $D$ in both cases. $N$ ranged from 4 to 20 in multiples of 4 with an initial field fluence = 10 and bandwidth $[0, f_{1\to N}]$, $f_{1\to N}$ being the frequency of the largest possible system transition. Results so far indicate that the flat dipole (solid lines) induces roughly linear scaling while the exponential dipole (dotted line) results in scaling of quadratic order or above.

Simulations on systems with different size $N$ and final time $T$ were performed for initial field fluences ranging from $10^{-8}$ to $10^{3}$. They indicate that a large initial fluence increases the search effort, but that

very low initial fluence doesn't increase the effort substantially.

However, the optimization must ultimately achieve a certain minimum fluence in order to reach the landscape maximum. This required minimum fluence $f_{min}$ grows as $N$ increases *and* as $T$ decreases, but is remarkably constant as a function of the initial fluence.

Overall, the nondegenerate simulations present a less "kinematic" picture than the degenerate ones. The dipole structure has a large effect on scaling, though the initial field fluence only has an effect when it's very large.

# 6 Landscape Analysis

??

# 7 Conclusion

It appears that many of the effects of the $\rho$ and $O$ spectral structure on search effort are similar to the kinematic results, so many of these conclusions will mirror those in the kinematic manuscript. So far the dependence on the dynamic physical parameters is intuitive also. The scaling behavior of the degenerate generic systems with respect to the spectra of $\rho$ and $O$ is consistent over a wide range of physical parameters.

# References


[1] M. Artamanov, H. Rabitz, *Chem. Phys.* **305**, 213-222 (2004)

[2] W. Jakubetz, E. Kades, J. Manz, *J. Phys. Chem.* **97**, 12609-12619 (1993)

[3] W. Zhu, H. Rabitz, *J. Chem Phys.* **108**, 1953 (1998)

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

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

[6] K. Young, M. Sarovar, R. Kosut, B. Whaley, *Phys. Rev. A* **79**, 062301 (2009)

[7] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, G. Gerber, *Science* **282**, 919-922 (1998)

[8] M. Bergt, T. Brixner, B. Kiefer, M. Strehle, G. Gerber, *J. Phys. Chem. A* **103**, 10381-10387 (1999)

[9] S. Vajda, A. Bartelt, E. Kaposta, T. Leisner, C. Lupulescu, S. Minemoto, P. Rosendo-Francisco, L. Woste, *J. Chem. Phys.* **267**, 231-239 (2001)

[10] V. Prokhorenko, A. Nagy, S. Waschuk, L. Brown, R. Birge, R. Miller, *Science* **313**, 1257 (2006)

[11] R. Bartels, S. Backus, E. Zeek, L. Misoguti, G. Vdovin, I. Christov, M. Murnane, H. Kapteyn, *Nature* **406**, 164-166 (2000)

[12] D. Reitze, S. Kazamias, F. Weihe, G. Mullot, D. Douillet, F. Auge, O. Albert, V. Ramanathan, J. Chambaret, D. Hulin, and P. Balcou, *Opt. Lett.* **29**, 86-88 (2004)

[13] T. Pfeifer, R. Kemmer, R. Spitzenpfeil, D. Walter, C. Winterfeldt, G. Gerber, C. Spielmann, *Opt. Lett.* **30**, 1497-1499 (2005)

[14] C. Bardeen, V. Yakovlev, K. Wilson, S. Carpenter, P. Weber, W. Warren, *Chem. Phys. Lett.* **280**, 151-158 (1997)

[15] T. Hornung, R. Meier, M. Motzkus, *Chem¿ Phys. Lett.* **326**, 445-453 (2000)

[16] T. Brixner, N. Damrauer, P. Niklaus, G. Gerber, *Nature* **414**, 57-59 (2001)

[17] J. Herek, W. Wohlleben, R. Cogdell, D. Zeidler, M. Motzkus, *Nature* **417**, 533-535 (2002)

[18] A. Efimov, M. Moores, N. Beach, J. Krause, and D. Reitze, *Opt. Lett.* **23**, 1915-1917 (1998)

[19] T. Weinacht, J. Ahn, P. Bucksbaum, *Nature* **397**, 233-235 (1999)

[20] T. Hornung, R. Meier, D. Zeidler, K. Kompa, D. Proch, M. Motzkus, *Appl. Phys. B* **71**, 277-284 (2000)

[21] R. Bartels, S. Backus, I. Christov, H. Kapteyn, M. Murnane, *Chem. Phys.* **267**, 277-289 (2001)

[22] F. Omenetto, A. Taylor, M. Moores, D. Reitze, *Opt. Lett.* **26**, 938-940 (2001)

[23] T. Brixner, N. Damrauer, G. Krampert, P. Niklaus, G. Gerber, *J. Opt. Soc. Am. B* **20**, 878-881 (2003)

[24] A. Efimov, A. Taylor, F. Omenetto, E. Vanin, *Opt. Lett.* **29**, 271-273 (2004)

[25] V. Ramakrishna, M.V. Salapaka, M. Dahleh, H. Rabitz, *Phys. Rev. A* **51**, 960-966 (1995)

[26] V. Jurdjevic, *Geometric Control Theory* (Cambridge University Press, Cambridge, 1997)

[27] A. Bryson and Y. Ho, *Applied Optimal Control: Optimization, Estimation and Control* (Taylor and Francis, Boca Raton, FL, 1975)

[28] T. Ho, J. Dominy, H. Rabitz, *Phys. Rev. A* **79**, 013422 (2009)

[29] G. Maday, G. Turinici, *J. Chem. Phys.* **118**, 8191-8196 (2003)

[30] Y. Ohtsuki, G. Turinici, H. Rabitz, *J. Chem. Phys.* **120**, 5509 (2004)

[31] MathWorks, MATLAB, The MathWorks, Natick, MA, 1994.

[32] N. Nieto and L. Simmons, *Phys. Rev. A* **19**, 438-444 (1979).

[33] E. de Lima, J. Hornos, *J. Chem. Phys.* **125**, 164110 (2006).