Optimization search effort over the control landscapes for open quantum systems with Kraus-map evolution

Optimization search effort over the control landscapes for open quantum systems with Kraus-map evolution

Anand Oza[1]**, Alexander Pechen**[1]**, Jason Dominy**[2]**, Vincent Beltrani**[1]**, Katharine Moore**[1] **and Herschel Rabitz**[1]

[1] Department of Chemistry, Princeton University, Princeton, NJ 08544, USA
[2] Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

E-mail: aoza@princeton.edu, apechen@princeton.edu, jdominy@princeton.edu, beltrani@princeton.edu, kwmoore@princeton.edu and hrabitz@princeton.edu

Abstract

A quantum control landscape is defined as the expectation value of a target observable Θ as a function of the control variables. In this work, control landscapes for open quantum systems governed by Kraus map evolution are analyzed. Kraus maps are used as the controls transforming an initial density matrix ρ_i into a final density matrix to maximize the expectation value of the observable Θ. The absence of suboptimal local maxima for the relevant control landscapes is numerically illustrated. The dependence of the optimization search effort is analyzed in terms of the dimension of the system N, the initial state ρ_i and the target observable Θ. It is found that if the number of nonzero eigenvalues in ρ_i remains constant, the search effort does not exhibit any significant dependence on N. If ρ_i has no zero eigenvalues, then the computational complexity and the required search effort rise with N. The dimension of the top manifold (i.e., the set of Kraus operators that maximizes the objective) is found to positively correlate with the optimization search efficiency. Under the assumption of full controllability, incoherent control modeled by Kraus maps is found to be more efficient in reaching the same value of the objective than coherent control modeled by unitary maps. Numerical simulations are also performed for control landscapes with linear constraints on the available Kraus maps, and suboptimal maxima are not revealed for these landscapes.

PACS numbers: 02.30.Yy, 03.65.Yz, 02.60.Pn, 02.40.Vh

1. Introduction

The general goal of quantum control is to apply a suitable external field to a system in order to maximize the expectation value of a target operator. If the system under control is isolated from the environment, then the dynamics are coherent and described by a unitary transformation, as

appears in coherent control [1–9]. In practice, all real systems are open and interact with the environment in some fashion, so that the dynamics of the system will have some incoherent component. Such incoherent control through interaction with the environment [10–16] or through quantum measurements [17–23] can be beneficial in many cases (e.g., in creation of thermal beams of metastable noble gases [24], in quantum computing with mixed states [25] or in the modification of Grover's algorithm [26] to extend the capabilities of the original unitary scheme).

In this paper, we consider the most general class of physically allowed state transformations of controlled open quantum systems. These transformations are represented by Kraus maps [27] providing a kinematic description of incoherent control. Embedded in these maps is information about the system and environment, both of which may be subject to control. A control action determines the system's evolution with a Kraus map Φ, which transforms an initial system state ρ_i into the evolved final state $\rho_f = \Phi(\rho_i)$. The final state ρ_f determines the expectation value $J[\Phi] := \langle \Theta \rangle = \mathrm{Tr}[\Phi(\rho_i)\Theta] = \mathrm{Tr}[\rho_f \Theta]$ of a target Hermitian operator Θ representing a desired physical property to be optimized. The corresponding control goal is formulated as follows: given an initial state ρ_i and a target observable Θ, find a Kraus map Φ_{opt} that transforms ρ_i into a state maximizing the expectation value, i.e., such that $J[\Phi_{\mathrm{opt}}] = \max_\Phi J[\Phi]$. The set of all Kraus maps for a given quantum system forms a complex Stiefel manifold to formulate the control goal as a nonlinear problem of maximizing the objective function J over the Stiefel manifold.

As shown in [28], for any desired final state ρ_f there exists a Kraus map Φ_{ρ_f} that transforms all initial states ρ_i into ρ_f, i.e., such that $\Phi_{\rho_f}(\rho_i) = \rho_f$ for any ρ_i. If $|\psi\rangle$ is an eigenvector of the target operator Θ that corresponds to the maximal eigenvalue $\theta_{\max}$, then the expectation $\mathrm{Tr}[\rho_f \Theta]$ is maximized by the final state $\rho_f = \rho_{f,\psi} = |\psi\rangle\langle\psi|$ and therefore the objective $J[\Phi] = \mathrm{Tr}[\Phi(\rho_i)\Theta]$ is maximized, e.g., by the optimal map $\Phi = \Phi_{\rho_{f,\psi}}$. The corresponding maximum objective value is $J_{\max} = \theta_{\max}$. Thus, the ability to generate dynamically arbitrary Kraus maps for an open quantum system implies its complete state-to-state controllability and, in particular, complete controllability for the objectives of the form $J = \langle \Theta \rangle$. In contrast, a closed quantum system controlled by unitary dynamics has restricted state controllability; if ρ_i and ρ_f do not have the same eigenvalue spectrum, there does not exist a unitary transformation U such that $U\rho_i U^\dagger = \rho_f$. The maximum attained value for the objective $J = \langle \Theta \rangle$ in this case will generally be less than $\theta_{\max}$.

The quantum control landscape is defined by $J = \langle \Theta \rangle$ as a function of the control variables. The ability to successfully use a gradient or other local algorithms for maximization of the objective function depends on the existence or the absence of suboptimal local maxima. If local maxima exist, a local algorithm could get stuck at such points, and for this reason, we refer to suboptimal local maxima as 'traps;' the presence of local saddle points should not serve as traps. In the case of coherent laser control, the landscape is known to be trap free [29, 30].

A detailed analysis of the control landscapes for incoherent control of open two-level quantum systems was performed [31], where the absence of traps for these landscapes was proven. Arbitrary multi-level systems were considered in [32], where it was shown that no suboptimal traps exist for the control landscapes for any finite-level open quantum system. In addition, a high-dimensional submanifold of optimal controls was found. As in the case of coherent control, these results on the absence of traps and the multi-dimensionality of the global optimum manifold provide a theoretical foundation for the empirical fact that it is relatively easy to find optimal solutions even in the presence of an environment.

The absence of traps in control landscapes for both closed and open quantum systems implies that the search using a local algorithm will eventually reach a global optimum solution. However, the absence of traps does not specify the efficiency of the optimization procedure and

the search effort needed to reach the solution. The efficiency of the optimization procedure to find an optimal control, which is of practical importance due to limitations on computer time in simulations and laboratory resources in experiments, is determined by the local features of the control landscape as well as its topological characteristics. Different trap-free control landscapes can exhibit different degrees of search complexity. The prior relevant theoretical landscape analyses [31, 32] for incoherent control of open quantum systems did not describe the dependence of efficiency on the key parameters of the control problem: the dimension of the system N and the eigenvalues of the target operator Θ and the initial state ρ. For closed quantum systems, a theoretical analysis of the computational complexity of coherent control landscapes was performed [33, 34] along with a numerical analysis of the search effort using gradient, genetic and simplex algorithms [35, 36]. The results indicate that the search effort scales weakly, or possibly independently, with the dimension of the system N.

This paper presents a numerical analysis, with a gradient algorithm, of the search effort for incoherent control of open quantum systems. The analysis lends insight into the topological and structural characteristics of the corresponding quantum control landscapes. It shows that the search effort for driving a pure state into another pure state with Kraus maps remains relatively constant as the dimension N of the system increases, and this behavior is qualitatively similar to the scaling behavior of the search effort for closed systems [35, 36]. A more general result is established for arbitrary, not necessarily pure, initial states: the search effort is essentially determined by the number of nonzero eigenvalues of the initial state ρ, and not by the dimension of the system N. Thus, when the number of nonzero eigenvalues of the initial state remains constant, the search effort does not depend on N. At the extreme of driving a mixed state with no zero eigenvalues into a pure state, the search effort increases with the dimension of the system. The detailed analysis shows that the search effort is sensitive to the eigenstructure of the initial state ρ and the target operator Θ; specifically, the degeneracies of the zero eigenvalue of ρ and of the maximal eigenvalue of Θ positively correlate with the search efficiency, so that higher values of these degeneracies require less optimization search effort and correspond to a more efficient search. Further, a comparative analysis of incoherent and coherent control shows that incoherent control under the full controllability assumption is a more efficient process than coherent control, indicating that the additional control freedom afforded by incoherent control can decrease the complexity of the problem. Finally, an analysis of control landscapes with linear constraints on the control variables is performed, and it does not reveal the presence of suboptimal traps even for a large number of independent constraints. Use of Kraus maps for modeling the controlled evolution of the system in this paper greatly simplifies computations as it does not require solving the dynamical evolution equations. Analysis of the scaling properties of the search effort for dynamical optimization of open quantum systems remains as an issue for future study that can be performed using various specific models for the system and the environment [37, 38].

This paper is organized as follows. Section 2 describes the general theoretical framework for the kinematic analysis of incoherent control of multilevel open quantum systems. The expressions for the gradient and Hessian of the objective function J are derived in section 3. Section 4 contains the results of the numerical simulations. Section 4.1 describes the details of the optimization procedure, and section 4.2 discusses the distribution of the objective values for randomly generated controls. Section 4.3 computationally demonstrates the absence of traps in the control landscape for a five-level quantum system. Section 4.4 shows the dependence of optimization efficiency on the dimension of the quantum system, and section 4.5 examines the dependence of the computational complexity on the degeneracy structure of the eigenvalues of the initial state and target observable. Section 4.6 compares the computational efficiency

of coherent and incoherent control. Optimization over constrained landscapes is investigated in section 5. Concluding remarks are given in section 6.

2. Formulation of control for arbitrary *N*-level systems

In this section, the evolution of controlled *N*-level open quantum systems is modeled by Kraus maps. As background, first the common formulation of the objective function in terms of Kraus operators is provided. Then the control problem is reformulated as optimization over a suitable Stiefel manifold; this representation is used in the subsequent numerical analysis.

2.1. Kraus maps

Let $\mathcal{M}_N$ be the vector space of $N \times N$ complex matrices, with identity matrix $\mathbb{I}_N$. The density matrix $\rho \in \mathcal{M}_N$ of an *N*-level quantum system is a positive semidefinite (and therefore Hermitian) matrix with unit trace, $\mathrm{Tr}\,\rho = 1$. A linear map $\Phi : \mathcal{M}_N \to \mathcal{M}_N$ is *positive* if $\Phi(M) \geqslant 0$ for any $M \in \mathcal{M}_N$ such that $M \geqslant 0$. The most general evolution transformations of density matrices are given by linear Kraus maps $\Phi : \mathcal{M}_N \to \mathcal{M}_N$, which are defined by the following two properties:

- Complete positivity: For any integer *n*, the map $\Phi \otimes \mathbb{I}_n$ acting on $\mathcal{M}_N \otimes \mathcal{M}_n$ is positive, where $\otimes$ denotes the Kronecker product.
- Trace preserving: For any $M \in \mathcal{M}_N$, $\mathrm{Tr}\,\Phi(M) = \mathrm{Tr}\,M$.

Any Kraus map Φ can be written in the Kraus operator-sum representation (OSR) form

$$\Phi(\rho) = \sum_{i=1}^{l} K_i \rho K_i^\dagger, \tag{1}$$

and the trace preservation condition implies for the Kraus operators $K_i \in \mathcal{M}_N$ that the relation is satisfied:

$$\sum_{i=1}^{l} K_i^\dagger K_i = \mathbb{I}_N. \tag{2}$$

There exist many equivalent operator-sum representations of the same Kraus map. In particular, as shown in [39], for any OSR with $l > N^2$ Kraus operators there exists an equivalent OSR with no more than N^2 Kraus operators. Thus, we only need to consider the OSR with $l = N^2$ Kraus operators (some of the Kraus operators can be zero matrices). Even for $l = N^2$ the decomposition (1) is not unique. Indeed, let $\mathcal{U}(n)$ be the set of $n \times n$ unitary matrices and let $U \in \mathcal{U}(N^2)$ be a unitary matrix with matrix elements u_{ij}. Define a new set of Kraus operators by the relation

$$\tilde{K}_j = \sum_{i=1}^{N^2} u_{ji} K_i, \qquad 1 \leqslant j \leqslant N^2. \tag{3}$$

Then $\sum_{i=1}^{N^2} \tilde{K}_i^\dagger \tilde{K}_i = \mathbb{I}_N$ and $\Phi(\rho) = \sum_{i=1}^{N^2} K_i \rho K_i^\dagger = \sum_{i=1}^{N^2} \tilde{K}_i \rho \tilde{K}_i^\dagger$ for any ρ. Therefore, the two sets of Kraus operators $\{K_i\}$ and $\{\tilde{K}_i\}$ provide two equivalent representations of the same Kraus map.

2.2. The objective function: formulation in terms of Kraus operators

The optimization goal in quantum control is to maximize the objective function $J = \langle\Theta\rangle \equiv \text{Tr}[\Phi(\rho)\Theta] = \text{Tr}[\rho_{t_\text{f}}\Theta]$, where Θ is some target Hermitian operator, $\langle\cdot\rangle$ denotes the expectation value at the final time t_f and ρ_{t_f} is the state of the system at the final time, evolved under controls from some initial state $\rho = \rho_{t_0}$. The Kraus operators $\{K_i\} \equiv \{K_i(t_\text{f}, t_0)\}$ describe the generally non-unitary evolution Φ of the initial density matrix ρ at time t_0 into a density matrix ρ_{t_f} at time t_f, such that $\rho_{t_\text{f}} = \Phi(\rho) = \sum_{i=1}^{N^2} K_i\rho K_i^\dagger$. They contain the information about the system–environment interaction, all control field interactions and the state of the environment which can also be used as a control. Hence, J is a function of the Kraus operators

$$J[K_1, \ldots, K_{N^2}] = \text{Tr}\left[\sum_{i=1}^{N^2} K_i\rho K_i^\dagger\Theta\right], \tag{4}$$

and the control goal can be formulated as a constrained optimization problem: given ρ and Θ, maximize J over all sets of operators $\{K_i\}$ that satisfy the constraint (2).

For the remainder of this paper, we will take ρ and Θ to be simultaneously diagonal. Indeed, we can always choose a basis in which Θ is diagonal, and write ρ and $\{K_i\}$ in this basis. Since ρ is Hermitian, there exists a unitary matrix Ω such that $\rho = \Omega\sigma\Omega^\dagger$, where σ is a diagonal matrix. Then the objective function (4) takes the form $J = \text{Tr}\left[\sum_{i=1}^{N^2} \tilde{K}_i\sigma\tilde{K}_i^\dagger\Theta\right]$, where $\tilde{K}_i = K_i\Omega$. The new Kraus operators $\{\tilde{K}_i\}$ also satisfy the constraint (2), and the objective function is equivalently represented as a function of $\tilde{K}_i$ with simultaneously diagonal matrices σ and Θ.

2.3. The objective function: formulation in terms of Stiefel manifolds

The above formulation can be expressed more succinctly in terms of the Stiefel manifold [40]. Let $\mathcal{M}(n, k, \mathbb{F})$ be the set of $n \times k$ matrices with matrix elements in the field $\mathbb{F}$ of real or complex numbers (i.e., $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). The Stiefel manifold is defined as

$$V_k(\mathbb{F}^n) = \{S \in \mathcal{M}(n, k, \mathbb{F}) : S^\dagger S = \mathbb{I}_k\}.$$

The manifold $V_k(\mathbb{F}^n)$ is called a real (respectively, complex) Stiefel manifold if $\mathbb{F} = \mathbb{R}$ (respectively, $\mathbb{F} = \mathbb{C}$). Given a Kraus map Φ and a set of Kraus operators $\{K_i\}$, we form the corresponding $N^3 \times N$ Stiefel matrix S as follows:

$$S = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_{N^2} \end{pmatrix}. \tag{5}$$

The constraint (2) can be expressed as the equality $S^\dagger S = \mathbb{I}_N$, which defines the complex Stiefel manifold $\mathcal{S} = V_N(\mathbb{C}^{N^3})$. Furthermore, the objective function (4) can be written as a function of the Stiefel matrix S:

$$J(S) = \text{Tr}[S\rho S^\dagger(\mathbb{I}_{N^2}\otimes\Theta)]. \tag{6}$$

The control goal in this formulation is to maximize the objective function (6) over the Stiefel manifold $\mathcal{S}$. Note that the objective function (6) is by construction real valued for any initial density matrix ρ and for any Hermitian target operator Θ.

We now address the non-uniqueness of the Kraus operator parametrization in terms of the Stiefel manifold. Let $\mathcal{W} = \{U \otimes \mathbb{I}_N : U \in \mathcal{U}(N^2)\}$. It is straightforward to verify that

$\forall S \in \mathcal{S}$ and $\forall W \in \mathcal{W}$ holds $\tilde{S} \equiv WS \in \mathcal{S}$. If $\{K_i\}$ and $\{\tilde{K}_i\}$ are two sets of Kraus operators that determine two Stiefel matrices S and $\tilde{S}$ through (5), then they define the same Kraus map and are related by the equality (3) if and only if $\exists W \in \mathcal{W}$ such that $\tilde{S} = WS$. Thus, equivalent parametrizations of the same Kraus map correspond to Stiefel matrices related by $\tilde{S} = WS$ with some $W \in \mathcal{W}$. This property implies the invariance of the objective function under $\mathcal{W}$-transformations, $J(S) = J(WS)$ for any $W \in \mathcal{W}$, and will be used in section 5 for analysis of the search effort for optimization of J with additional constraints on the available Kraus operators.

The Stiefel manifold $V_k(\mathbb{F}^n)$ can also be defined as the set of orthonormal k-frames in $\mathbb{F}^n$ [41]. In this way, the Stiefel manifold $\mathcal{S}$ can be specified as the set of ordered N-tuples $X_1, \ldots, X_N \in \mathbb{C}^{N^3}$ such that $\langle X_i, X_j \rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta symbol and $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{C}^{N^3}$. In the remainder of the manuscript, the notation $\langle \cdot, \cdot \rangle$ will be used for inner products in several appropriate different spaces (namely, standard inner products in $\mathbb{C}^{N^3}$ and in $\mathbb{C}^{N^2}$, and real Hilbert–Schmidt inner products in S and in the tangent space $T_S \mathcal{S}$ at S). Vector X_i in this representation contains elements of the ith column of the Stiefel matrix (5) in certain order and can be decomposed into the direct sum

$$X_i = Y_1^i \oplus Y_2^i \oplus \cdots \oplus Y_N^i.$$

Here each $Y_j^i \in \mathbb{C}^{N^2}$, where $1 \leqslant i, j \leqslant N$, is a complex vector of length N^2 of the form $Y_j^i = \{(K_1)_{ji}, (K_2)_{ji}, \ldots, (K_{N^2})_{ji}\}$, i.e., components of the vector Y_j^i are the jith matrix elements of all the N^2 Kraus operators K_l. The orthogonality condition $\langle X_i, X_j \rangle = \delta_{ij}$ implies the relation

$$\sum_{k=1}^{N} \langle Y_k^i, Y_k^j \rangle = \delta_{ij}. \tag{7}$$

Here $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{C}^{N^2}$ and should be distinguished from the same notations used above to denote the inner product in $\mathbb{C}^{N^3}$.

The objective function for diagonal matrices $\rho = \sum_{i=1}^{N} \rho_i |i\rangle\langle i|$ and $\Theta = \sum_{j=1}^{N} \theta_j |j\rangle\langle j|$ can be written as

$$J[\{Y_j^i\}] = \sum_{i,j=1}^{N} \|Y_j^i\|^2 \rho_i \theta_j. \tag{8}$$

It is clear that $\theta_{\min} \leqslant J(S) \leqslant \theta_{\max}$, where $\theta_{\min}$ and $\theta_{\max}$ are the minimum and maximum eigenvalues of Θ, respectively. Indeed, we have

$$\theta_{\min} \sum_{i,j=1}^{N} \|Y_i^j\|^2 \rho_j \leqslant J \leqslant \theta_{\max} \sum_{i,j=1}^{N} \|Y_i^j\|^2 \rho_j.$$

Now, by first summing over i and using (7) and then summing over j and using $\mathrm{Tr}\,\rho = 1$, we have the desired inequalities.

Since the maximal value of the objective function J equals $\theta_{\max}$, the set of optimal controls (i.e., the set of all Stiefel matrices which maximize the objective function) is the manifold $\mathcal{M}_{\max} = \{S \in \mathcal{S} : J(S) = \theta_{\max}\}$. For the case $\Theta = |N\rangle\langle N|$ of special interest, it follows from (8) that

$$\mathcal{M}_{\max} = \left\{\{Y_j^i\}_{i,j=1}^{N} : \|Y_j^i\|^2 = \delta_{jN} \text{ for any } i \text{ such that } \rho_i \neq 0\right\}.$$

3. Gradient and Hessian of *J*

The numerical analysis in section 4 uses a gradient algorithm for the optimization of the objective function $J(S)$. This algorithm requires solving the equation

$$\frac{\mathrm{d}S}{\mathrm{d}\sigma} = \operatorname{grad} J(S). \tag{9}$$

Here grad J is the gradient of the objective function, which induces the corresponding gradient flow on the Stiefel manifold $\mathcal{S}$ via (9).

3.1. Gradient of J

We now derive an explicit expression for the gradient. Denote the differential of J at $S \in \mathcal{S}$ by $\mathrm{d}_S J : T_S \mathcal{S} \to \mathbb{R}$, where $T_S \mathcal{S}$ is the tangent space at S. By the product rule for derivatives

$$\mathrm{d}_S J(\delta S) = \Re \operatorname{Tr}[(\delta S)\rho S^\dagger (\mathbb{I}_{N^2} \otimes \Theta) + S\rho(\delta S)^\dagger (\mathbb{I}_{N^2} \otimes \Theta)], \tag{10}$$

where the real part $\Re$ is taken since the objective (6) is a real function. Since $\Re \operatorname{Tr} A = \Re \operatorname{Tr} A^\dagger$ for any matrix A, the second term on the right-hand side of (10) can be rewritten as $\Re \operatorname{Tr}[(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho S^\dagger]$ and we obtain

$$\begin{aligned}
\mathrm{d}_S J(\delta S) &= \Re \operatorname{Tr}[(\delta S)\rho S^\dagger (\mathbb{I}_{N^2} \otimes \Theta) + (\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho S^\dagger] \\
&= \Re \operatorname{Tr}[\rho S^\dagger (\mathbb{I}_{N^2} \otimes \Theta)(\delta S) + \rho S^\dagger (\mathbb{I}_{N^2} \otimes \Theta)(\delta S)] \\
&= 2 \Re \operatorname{Tr}[\rho S^\dagger (\mathbb{I}_{N^2} \otimes \Theta)(\delta S)] \\
&= \langle 2(\mathbb{I}_{N^2} \otimes \Theta)S\rho, \delta S \rangle,
\end{aligned} \tag{11}$$

where $\delta S \in T_S \mathcal{S}$, and $\langle A, B \rangle := \Re \operatorname{Tr}[A^\dagger B]$ is the inner product on $\mathcal{S}$ and $T_S \mathcal{S}$. By the Riesz representation theorem, there exists $X \in T_S \mathcal{S}$ such that $\mathrm{d}_S J(\delta S) = \langle X, \delta S \rangle$ for all $\delta S \in T_S \mathcal{S}$. The vector X is the *gradient* of J at S, denoted by grad $J(S)$.

Since grad $J(S)$ must lie in $T_S \mathcal{S}$, it is necessary to remove the component orthogonal to $T_S \mathcal{S}$ from the vector $2(\mathbb{I}_{N^2} \otimes \Theta)S\rho$ appearing in the last line of (11). Differentiation of the identity $S^\dagger S = \mathbb{I}_N$ gives $S^\dagger (\delta S) = -(\delta S)^\dagger S$, so $S^\dagger (\delta S)$ is skew-Hermitian. This can be rewritten as $\delta S = SB + (\mathbb{I}_{N^3} - SS^\dagger)D$, where $B \in \mathcal{M}(N, N, \mathbb{C})$ is a skew-Hermitian matrix and $D \in \mathcal{M}(N^3, N, \mathbb{C})$ is an arbitrary matrix. (Note that $S^\dagger (\delta S) = B$, since $S^\dagger S = \mathbb{I}_N$.) Any $A \in \mathcal{M}(N^3, N, \mathbb{C})$ can be decomposed as follows:

$$A = S\tfrac{1}{2}(S^\dagger A + A^\dagger S) + S\tfrac{1}{2}(S^\dagger A - A^\dagger S) + (\mathbb{I}_{N^3} - SS^\dagger)A.$$

Let $C = (S^\dagger A + A^\dagger S)/2$ and $B = (S^\dagger A - A^\dagger S)/2$, so that $A = SC + SB + (\mathbb{I}_{N^3} - SS^\dagger)A$. Clearly C is Hermitian and B is skew-Hermitian, so $\langle SC, SB - (\mathbb{I}_{N^3} - SS^\dagger)A \rangle = 0$. Therefore, SC is orthogonal to $T_S \mathcal{S}$, and hence $(A - SC) \in T_S \mathcal{S}$. As a result, $\mathcal{P}_S(A) = A - S(S^\dagger A + A^\dagger S)/2$ is an orthogonal projector from $\mathcal{M}(N^3, N, \mathbb{C})$ onto $T_S \mathcal{S}$, and

$$\begin{aligned}
\operatorname{grad} J(S) &= 2(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S[S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho + ((\mathbb{I}_{N^2} \otimes \Theta)S\rho)^\dagger S] \\
&= (2\mathbb{I}_{N^3} - SS^\dagger)(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S.
\end{aligned}$$

3.2. Hessian of J

In the analysis thus far, we have only considered grad $J(S)$, which gives first-order information about J. The Hessian gives useful second-order information about the minima, maxima and saddles of J (where grad $J(S) = 0$). At such points, the eigenvectors of the Hessian with positive (respectively negative) eigenvalues correspond to directions in which J increases (respectively decreases).

The Hessian of J at $S \in \mathcal{S}$ acting on $\delta S \in T_S \mathcal{S}$ is defined as the covariant derivative of grad $J(S)$ in the direction δS [42]:

$$\text{Hess } J(S) : T_S \mathcal{S} \to T_S \mathcal{S}, \qquad \text{Hess } J(S)(\delta S) = \nabla_{\delta S} \text{ grad } J(S).$$

Covariant differentiation of a function on a vector space is equivalent to taking the ordinary differential. However, $\mathcal{S}$ is not a vector space. In the following, the strategy will be to take the covariant derivative of grad $J(S)$ as a function on $\mathcal{M}(N^3, N, \mathbb{C})$, which is a vector space, and then project this onto $\mathcal{S}$. Since $\mathcal{S}$ inherits its inner product from $\mathcal{M}(N^3, N, \mathbb{C})$, this strategy gives the covariant derivative of grad $J(S)$ on $\mathcal{S}$.

We now calculate an expression for the eigenvalues and eigenvectors of the Hessian of J on the critical manifolds. By differentiating grad $J(S)$ in the direction of δS, we obtain

$$\begin{aligned}
\overline{\nabla}_{\delta S} \text{ grad } J(S) &= \text{d}_S \text{ grad } J(\delta S) \\
&= 2(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho - (\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S(\delta S)^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho \\
&\quad - SS^\dagger(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho - (\delta S)\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S - S\rho(\delta S)^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad - S\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)(\delta S),
\end{aligned}$$

where $\overline{\nabla}$ denotes the Riemannian connection on $\mathcal{M}(N^3, N, \mathbb{C})$. We now project this onto $T_S \mathcal{S}$. Letting $A = \overline{\nabla}_{\delta S} \text{ grad } J(S)$ gives

$$\text{Hess } J(S)(\delta S) = \nabla_{\delta S} \text{ grad } J(S) = P_S(\overline{\nabla}_{\delta S} \text{ grad } J(S)) = A - \tfrac{1}{2}S(S^\dagger A + A^\dagger S).$$

With some algebra, this expression can be reduced to

$$\begin{aligned}
\text{Hess } J(S)(\delta S) &= 2(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho - (\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho - (\delta S)\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad + \tfrac{1}{2}[SS^\dagger(\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho + SS^\dagger(\delta S)\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad - 2SS^\dagger(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho - 2S\rho(\delta S)^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad + S\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S(\delta S)^\dagger S + SS^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho(\delta S)^\dagger S].
\end{aligned}$$

Combining the first two terms in the square brackets gives

$$\begin{aligned}
SS^\dagger(\delta S)(S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho + \rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S) &= SS^\dagger(\delta S)(2S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S^\dagger \text{ grad } J(S)) \\
&= 2SS^\dagger(\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho,
\end{aligned}$$

since grad $J(S) = 0$ at a critical point. Combining the last two terms in the square brackets gives

$$\begin{aligned}
(S\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S + SS^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho)((\delta S)^\dagger S) &= (2(\mathbb{I}_{N^2} \otimes \Theta)S\rho - \text{grad } J(S))((\delta S)^\dagger S) \\
&= 2(\mathbb{I}_{N^2} \otimes \Theta)S\rho(\delta S)^\dagger S.
\end{aligned}$$

As a result, we have

$$\begin{aligned}
\text{Hess } J(S)(\delta S) &= 2(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho - (\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho - (\delta S)\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad - SS^\dagger(\mathbb{I}_{N^2} \otimes \Theta)(\delta S)\rho + SS^\dagger(\delta S)S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S\rho(\delta S)^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S \\
&\quad + (\mathbb{I}_{N^2} \otimes \Theta)S\rho(\delta S)^\dagger S.
\end{aligned}$$

4. Numerical assessment of optimization efficiency for landscapes without constraints

This section presents numerical simulations, including (a) an empirical demonstration of the absence of suboptimal traps in the control landscape, (b) an analysis of the dependence of optimization efficiency on the dimension of the system N, target operator Θ and initial state ρ, and (c) a comparison between coherent and incoherent control.

4.1. The optimization procedure

We now describe the procedure for the numerical analysis of the controlled excursions over the landscapes without constraints on the controls. First, an adapted version of the algorithm in [43] is used to randomly generate an initial Stiefel matrix S_0 with a uniform distribution on the Stiefel manifold $\mathcal{S}$. After the initial Stiefel matrix has been generated, the Runge–Kutta method built into MATLAB is used to solve (9) with the initial condition $S(0) = S_0$. The method relies on using a variable step size. The tolerances in the differential equation solver are set so that $\|S^\dagger S - \mathbb{I}_N\| < 2 \times 10^{-4}$ at any given point in the trajectory. Integration is terminated when $J(S) > (\theta_{\max} - 0.01)$.

The efficiency of the optimization procedure is measured by the two parameters: (1) the number τ of σ-steps taken by the differential equation solver in MATLAB to reach the objective value $J > (\theta_{\max} - 0.01)$ and (2) the path length λ taken to get there. A higher number τ of σ-steps corresponds to a more difficult optimization problem. Given the number τ of σ-steps, the path length λ is defined as

$$\lambda = \sum_{i=0}^{\tau-1} ||S(i+1) - S(i)||, \tag{12}$$

where $\|S\| = \sqrt{\langle S, S \rangle}$ is the norm on $\mathcal{S}$. Similarly, a large value of λ corresponds to a convoluted trajectory through $\mathcal{S}$ and indicates an inefficient optimization.

To ensure statistical uniformity for some simulations, an average was performed over the initial state with a uniform distribution. Uniform sampling on the space of diagonal density matrices is implemented as follows. Let $\mathcal{V}_n^+$ be the standard simplex, i.e., the set of all vectors $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ such that $z_i \geqslant 0$ and $\sum_{i=1}^n z_i = 1$. Let $x_i = -\log(a_i)$ where a_i is uniformly distributed on [0, 1], so x_i are exponentially distributed with parameter 1. Now let

$$y_i = \frac{x_i}{x_1 + \cdots + x_n}, \qquad 1 \leqslant i \leqslant n.$$

Then the random vector $y = (y_1, \ldots, y_n)$ is uniformly distributed on the simplex $\mathcal{V}_n^+$ [44] and the diagonal density matrix ρ with matrix elements $\rho_{ii} = y_i$ is uniformly distributed.

4.2. The statistical distribution of the objective for randomly generated controls

In practical optimization of the objective function, either in the laboratory or through simulations with a numerical algorithm, the initial control is usually randomly generated. As the Stiefel matrices serve as the controls, we first analyze the distribution of the objective value for randomly generated initial Stiefel matrices. Figure 1 shows the mean value $\bar{J}_0$ of the objective function $J_0(S_0, \rho) = \text{Tr}\left[S_0 \rho S_0^\dagger (\mathbb{I}_{N^2} \otimes \Theta)\right]$ for $\Theta = |N\rangle\langle N|$ as a function of the system dimension N for a uniform distribution of the initial Stiefel matrix S_0 and uniform distribution of the initial diagonal density matrices ρ. For this case, the mean value $\bar{J}_0$ equals $1/N$. To understand this result, let $|1\rangle, \ldots, |N\rangle$ be an orthonormal basis in the Hilbert space of the system such that $|N\rangle \equiv |\Psi\rangle$ is the target state. The uniform generation of the Stiefel matrix S_0 and initial density matrix ρ does not have a preferred state and thus preserves the symmetry between the states $|1\rangle, \ldots, |N\rangle$. Therefore, in the final density matrix ρ' obtained by applying to ρ the Kraus map associated with the Stiefel matrix S_0, the averaged (over uniform distributions of ρ and S_0) population p_i of each of these states will be the same for all i. Since $\sum_{i=1}^N p_i = 1$ and $p_1 = p_2 = \cdots = p_N$, we have $p_i = 1/N$ for each i. Hence, the mean value of $J_0 = \text{Tr}\left[S_0 \rho S_0^\dagger (\mathbb{I}_{N^2} \otimes \Theta)\right] \equiv p_N$ will be $1/N$ for $\Theta = |\Psi\rangle\langle\Psi|$ being a projector onto the target state $|\Psi\rangle$. Figure 1 shows the decrease in the expected initial value of the objective function along with a decrease in the standard deviation with increasing system
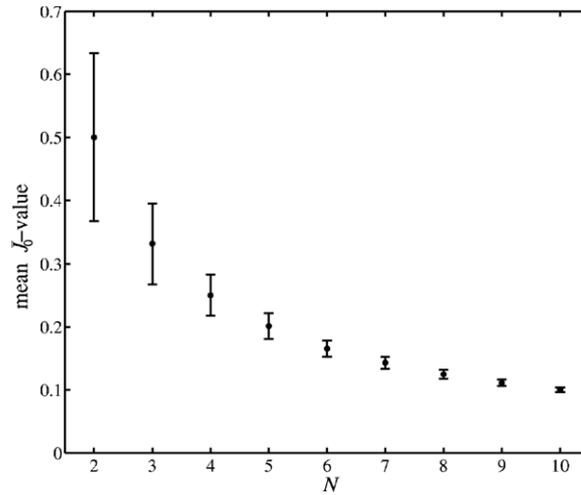
Figure 1. The mean value of the objective function $J_0(S_0, \rho) = \mathrm{Tr}[S_0 \rho S_0^\dagger (\mathbb{I}_{N^2} \otimes \Theta)]$, where S_0 has a uniform distribution on the Stiefel manifold, ρ has a uniform distribution over the set of diagonal density matrices and $\Theta = |N\rangle\langle N|$. Five hundred samples were taken for every point N. The error bars show the standard deviation for each N.
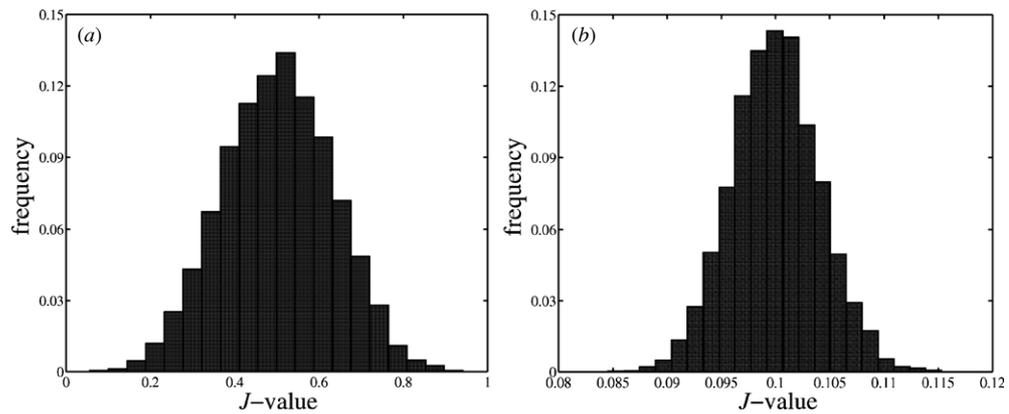

Figure 2. The distribution of the values of the objective function $J_0(S_0, \rho) = \mathrm{Tr}[S_0 \rho S_0^\dagger (\mathbb{I}_{N^2} \otimes \Theta)]$ for $N = 2$ (*a*) and for $N = 10$ (*b*). The initial Stiefel matrix S_0 is uniformly distributed on the Stiefel manifold, ρ is uniformly distributed on the set of diagonal density matrices and the target operator has the form $\Theta = |N\rangle\langle N|$. 10^4 samples were taken to produce the statistics for each plot.

dimension N. Figure 2 presents the detailed form of the distributions for cases $N = 2$ and $N = 10$, respectively shown in 2a and 2b, with a uniform distribution of S_0 on the Stiefel manifold and a uniform distribution of ρ on the set of diagonal matrices. In this figure, the distributions of the values of the objective function J_0 are produced using 10^4 randomly selected pairs of S_0 and ρ. The results agree with the natural expectation that the efficiency of a randomly chosen control decreases with increasing complexity of the system. The figure also shows that as N rises, the distribution of the objective values becomes more concentrated around the mean value. An open issue is to obtain an analytical expression for the distribution of the initial objective value $J_0(S_0, \rho)$.
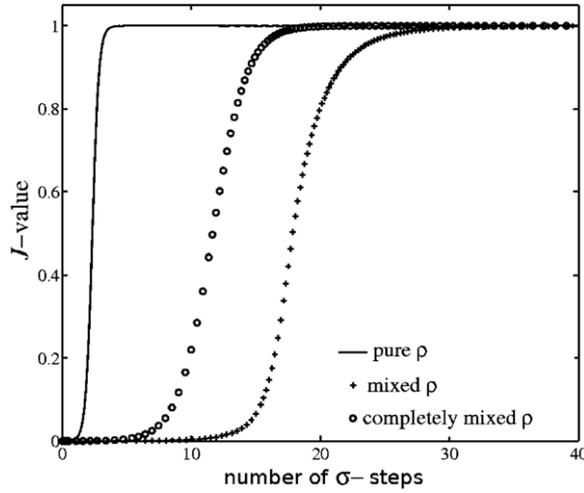
Figure 3. The figure shows the value of the objective function J at each step σ in the trajectory for a pure initial state $\rho = \mathrm{diag}(1, 0, 0, 0, 0)$, a mixed initial state and a completely mixed initial state $\rho = 1/5 \cdot \mathrm{diag}(1, 1, 1, 1, 1)$. All three cases correspond to $N = 5$ and $\Theta = |5\rangle\langle5| = \mathrm{diag}(0, 0, 0, 0, 1)$. Each trajectory reaches perfect control $J = 1$ at the top of the landscape.

4.3. Absence of suboptimal traps

Let X be a topological space and $f : X \to \mathbb{R}$. The function f is said to have a local maximum at $x_0 \in X$ if there exists an open neighborhood of x_0, $U_{x_0} \subset X$, such that $\forall x \in U_{x_0}$, $f(x) \leqslant f(x_0)$ and yet there exists some $x_1 \in X$ such that $f(x_1) > f(x_0)$. If X represents the space of all controls and $f : X \to \mathbb{R}$ is the objective function to be maximized on X, then a local maximum of f is called a suboptimal (or false) trap in the control landscape produced by f.

The control landscape for the objective function $J : \mathcal{S} \to \mathbb{R}$ defined by (6) is known to have no traps [32]. Figure 3 numerically demonstrates this general fact for a particular five-level quantum system. In the figure, three different initial density matrices are considered: a pure state $\rho = |1\rangle\langle1|$, a randomly generated mixed state and a completely mixed state $\rho = \mathbb{I}_5/5$. The control goal is to transform each of these states into the final state $\rho_{\mathrm{f}} = |5\rangle\langle5|$, which maximizes the expectation of the target operator $\Theta = |5\rangle\langle5|$. As shown in the figure, in each case the gradient algorithm is able to find the control corresponding to the maximum value $J = 1$ of the objective function. The algorithm was not impeded by suboptimal local maxima, for their presence would have caused the algorithm to terminate at $J < 1$. Many other cases showed the same trap-free behavior (not shown here).

4.4. Dependence of the search effort on the dimension of the system

We now analyze how the dimension N of the controlled system affects the optimization search effort. The goal is to numerically analyze the statistical dependence upon N of the number of steps τ to reach convergence and the path length λ. In this section, the target operator $\Theta = |N\rangle\langle N|$ is the projector onto the state $|N\rangle$. To obtain reasonable statistics, for each N we average over 50 simulations of the optimization procedure with randomly (uniformly across $\mathcal{S}$) generated $S(0)$ and randomly generated initial density matrices ρ. We also analyze how the number of zero eigenvalues of ρ (henceforth denoted as d_0) affects the scaling of optimization efficiency with N.
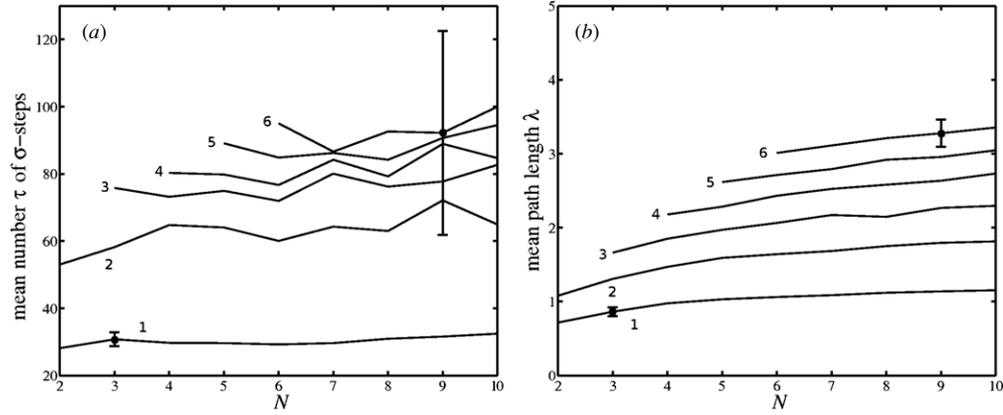
Figure 4. Dependence of search efficiency on the dimension of the system N for fixed numbers $N - d_0$ of nonzero eigenvalues of ρ. Fifty simulations were performed for each point, and the average values of the number τ of σ-steps and the path length λ are plotted in (*a*) and (*b*), respectively. On each subplot, the six lines, from bottom to top, correspond to the number of nonzero eigenvalues of the initial density matrix $N - d_0 = 1, 2, 3, 4, 5, 6$, respectively. The error bars indicate the typical standard deviation of the data for cases $N - d_0 = 1$ and $N - d_0 = 6$.
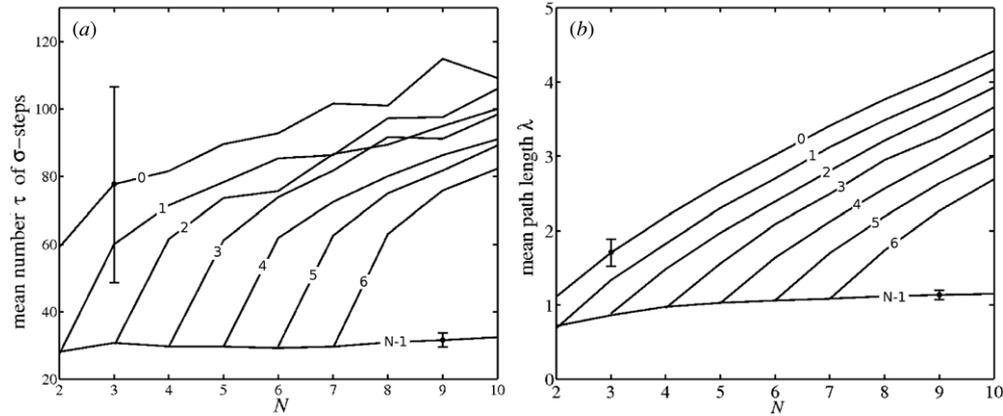

Figure 5. Dependence of search efficiency on N for different numbers d_0 of zero eigenvalues of ρ. Fifty simulations were performed for each point, and the average values of the number τ of σ-steps and the path length λ are plotted in (*a*) and (*b*), respectively. On each subplot, the eight lines, from top to bottom, correspond to the number of zero eigenvalues of the initial density matrix $d_0 = 0$, 1, 2, 3, 4, 5, 6, $N - 1$, respectively. The error bars indicate the typical standard deviation of the data for cases $d_0 = 0$ and $d_0 = N - 1$.

In the case of mixed ρ, we change ρ at the start of each individual simulation. Figures 4 and 5 plot τ and λ in two different ways in order to illustrate the issues driving the scaling efficiency. For each of the six curves in figure 4, the number of nonzero eigenvalues $N - d_0$ of the initial state ρ remains fixed. Each curve labeled by $N - d_0$ corresponds, for example, to the control of a sequence of quantum systems prepared initially in a state at a relatively low temperature, with no population in d_0 high eigenstates of the density matrix. Both τ and λ do not show any significant dependence upon N. It is clear that for fixed N, the search

efficiency is greater for larger values of d_0. However, increasing d_0 for fixed $N - d_0$ does not change the slope of the curves in figure 4, showing that the complexity of the search remains relatively insensitive to N. The most efficient control problem considered in the figure is the transformation of a pure initial state $\rho = |j\rangle\langle j|$ ($j \neq N$) into a pure final state $\rho_f = |N\rangle\langle N|$.

In contrast, for the simulations in figure 5, d_0 is held fixed for all N. This corresponds, for example, to the control of a sequence of quantum systems with the initial state at ever higher temperature as N rises, producing a large number of populated energy states. Both τ and λ increase quite sharply as N increases, as shown in figure 5. It is clear that for fixed N, the efficiency of optimization increases as d_0 increases. However, as with figure 4, increasing d_0 does not change the slope of the curves in figure 5, showing that the efficiency remains sensitive to N. The most inefficient search corresponds to the control goal of transforming a maximum entropy initial state with $d_0 = 0$ into a pure final state with $d_0 = N - 1$, which agrees with simple intuition.

The conclusion from figures 4 and 5 is that when Θ is a projector, the search efficiency decreases with increasing numbers of *nonzero* eigenvalues of ρ. The overall dimension of the quantum system N has little effect upon the search efficiency, provided that the number of nonzero eigenvalues of ρ remains fixed. The large standard deviations in both figures are most likely caused by fluctuations of the initial Stiefel matrix S and of the parameters of the initial density matrix ρ not included in the number of zero eigenvalues d_0.

The results in figure 4 have practical relevance. In the laboratory, a sequence of quantum systems with increasing N and a roughly fixed small number of populated energy levels can be arranged. Under these conditions, the results shown in figure 4 indicate that the search effort in the laboratory should not be very sensitive to the dimension of the quantum system under control. This behavior is generally consistent with the broad findings that system and environmental complexity appear to have little effect on the number of iterations to reach successful control in the laboratory.

4.5. Dependence of the search effort on the degeneracy structure of ρ and Θ

In this section, we analyze the dependence of the optimization search effort on the degeneracy structure of ρ and Θ. Recall that $\mathcal{M}_{max} = \{S \in \mathcal{S} : J(S) = \theta_{max}\}$, where θ_{max} is the maximal eigenvalue of Θ. As shown in [32], the dimension of $\mathcal{M}_{max}$ is

$$\dim(\mathcal{M}_{max}) = 2(d_0 + e_1)N^3 - (2d_0e_1 + 1)N^2, \tag{13}$$

where d_0 and e_1 are the degeneracies of the zero eigenvalue of ρ and maximal eigenvalue θ_{max} of Θ, respectively. The dimension of the maximum manifold as a function of d_0 and e_1 is plotted on figure 6(a). If d_0 is close to N, then the initial state ρ is close to a pure state, and for e_1 close to N, the target operator Θ is close to a constant multiple of the identity operator. Equation (13) and figure 6(a) show that large values of d_0 and e_1 correspond to higher-dimensional maximum submanifolds (note that the dimension of the maximum manifold on the vertical axis of figure 6(a) increases in the downward direction).

Figures 6(b) and (c) show the dependence of efficiency of optimization upon d_0 and e_1. As d_0 and e_1 approach N, the efficiency of optimization increases rapidly. Comparison with figure 6(a) shows a strong positive correlation between the dimension of the maximum manifold and the efficiency of optimization. This result is expected, since an increase in $\dim(\mathcal{M}_{max})$ corresponds to a larger target submanifold of optimal solutions. The presence of the positive correlation is illustrated in a more explicit way in figure 7, where the two parameters τ and λ characterizing the efficiency of optimization are plotted versus the dimension of the maximum submanifold. The dimension of the maximum manifold is determined by the pair
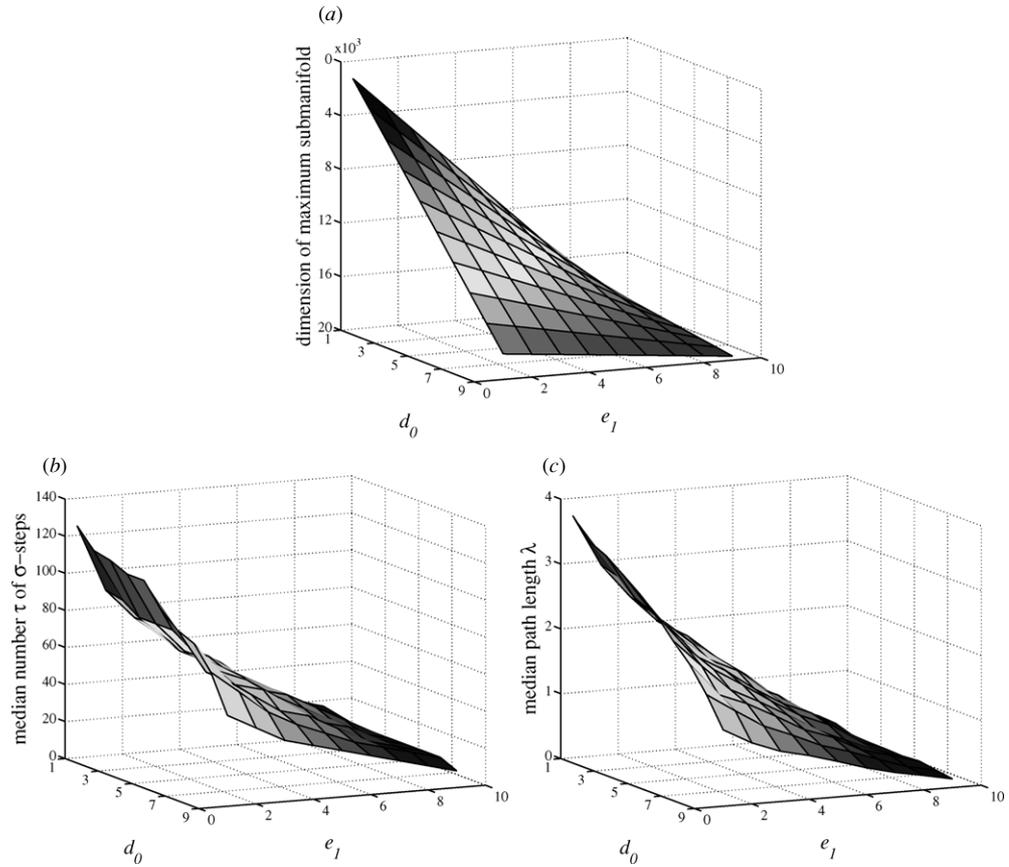
Figure 6. Dependence of search efficiency on the degeneracy d_0 of the zero eigenvalue of ρ and the degeneracy e_1 of the maximal eigenvalue of Θ for $N = 10$. (*a*) shows the dimension of the maximum submanifold as a function of these two parameters. (*b*) shows the median number τ of σ-steps and (*c*) shows the median path length λ. (*b*) and (*c*) show that the search efficiency increases as the dimension of the maximum submanifold increases.

(d_0, e_1), and different pairs can produce the same dimension of the maximum manifold. Each point in figure 7 corresponds to a pair (d_0, e_1). The figure shows the general trend that an increase in the dimension of the maximum manifold decreases the required optimization search effort; however, the correlation is not perfect and different pairs (d_0, e_1) and (d'_0, e'_1) with the same or almost the same dimensions of their respective maximum manifolds can have different values of the parameters τ and λ.

4.6. Comparison of coherent and incoherent control

We now compare the efficiencies of coherent and incoherent control. The coherent control mechanism is implemented as follows. Let $S(0)$ be defined by $K_i = U/N$ for some $U \in \mathcal{U}(N)$. It is shown in appendix A that unitary Kraus maps form an invariant submanifold of $\mathcal{S}$ with respect to grad $J(S)$. That is, if $\mathcal{S}_{\mathcal{U}} = \{S \in \mathcal{S} : \exists U \in \mathcal{U}(N) : K_i = U/N\}$, then the solution to $\mathrm{d}S/\mathrm{d}\sigma = \mathrm{grad}\, J(S)$ with the initial condition $S(0) = S_0 \in \mathcal{S}_{\mathcal{U}}$ will lie entirely in $\mathcal{S}_{\mathcal{U}}$. Hence,
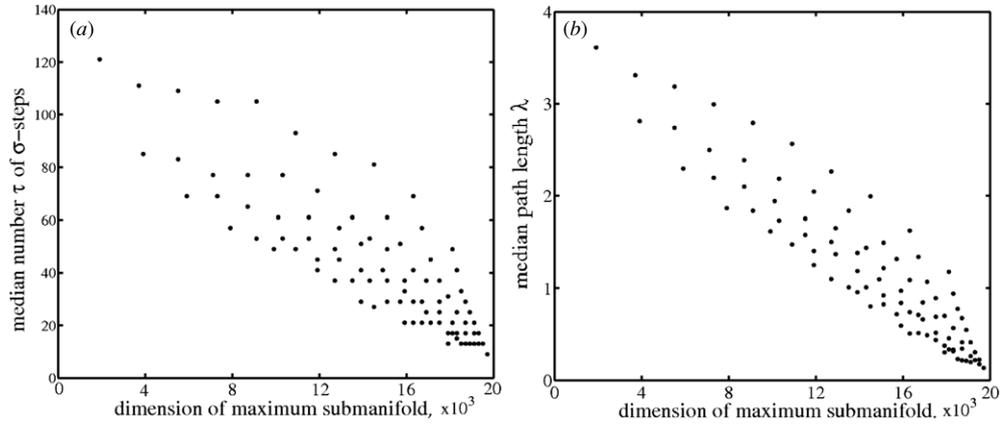
Figure 7. Dependence of search efficiency on the dimension of the maximum submanifold for $N = 10$. (*a*) and (*b*) show the dependence of the median number τ of σ-steps and median path length λ, respectively, on the dimension of the maximum submanifold. The dimension of the maximum submanifold is determined by the pair (d_0, e_1); each point on the plot corresponds to a (d_0, e_1) pair. These figures show the increase in the search efficiency as the dimension of the maximum submanifold increases.

solving the differential equation allows us to simulate density matrix evolution by coherent unitary control. Indeed, then $\rho_t = \sum_{i=1}^{N^2} K_i(t)\rho K_i^{\dagger}(t) = U(t)\rho U^{\dagger}(t)$.

In all the simulations here, $\Theta = |N\rangle\langle N|$. For unitary control, the maximal value of the objective function $J[U] = \text{Tr}[U\rho U^{\dagger}\Theta]$ is the maximal eigenvalue $\rho_{\max}$ of the initial state ρ. Thus, to ensure a fair comparison between the coherent and incoherent control, the target observable value is set to $J = \rho_{\max}$ for both incoherent and coherent control, and the algorithm stops as soon as the value $J = \rho_{\max} - 0.01$ is attained. This stopping criteria is the reason for the difference between the curve corresponding to incoherent control in figure 5(*a*) and the curve in figure 8; in the simulations displayed in the prior figure, the target observable value was $J = 1$ rather than $J = \rho_{\max}$. Figure 8 shows that with the ability to generate arbitrary Kraus maps, incoherent control can be a far more efficient process than coherent control for both pure and mixed ρ, especially for large values of N. The greater freedom allowed by incoherent control decreases the complexity of the problem and allows for a more efficient search.

5. Control under linear constraints on the Kraus operators

This section considers control under additional constraints on the available Kraus maps, which produce constraints on the Stiefel manifold. The target operator is assumed to have the form $\Theta = |N\rangle\langle N|$.

Let $h : \mathcal{S} \rightarrow \mathbb{R}^q$ be a set of q real-valued constraints. Recall from section 2.3 that the objective function J is invariant under $\mathcal{W}$-transformations. Since $\mathcal{W}$-transformations correspond to different parametrizations of the same physical evolution Kraus map, any reasonable constraint should be $\mathcal{W}$-invariant, and thus we impose the requirement that $h(S) = h(WS)$ for any $S \in \mathcal{S}$ and any $W \in \mathcal{W}$.

We restrict our attention to *affine* constraints, which are of the form $h(S) = g(S) - \gamma$, where g is linear over $\mathbb{R}$ and $\gamma \in \mathbb{R}^q$ is a constant. Specifically, for a given set of matrices $\{B_1, \ldots, B_n\}$ we consider $\mathcal{W}$-invariant affine constraints of the form $\text{Tr}(B_l^{\dagger} K_j) = 0$ for each
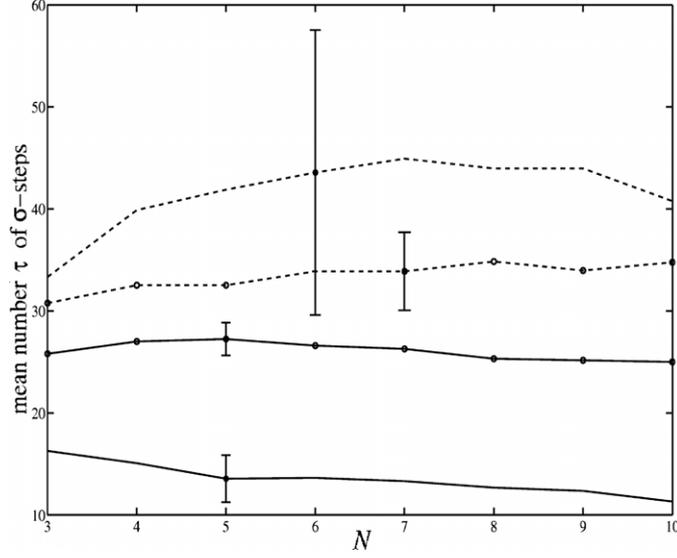
15

Figure 8. Comparison of the optimization efficiency for incoherent control by Kraus maps and coherent control by unitary transformations, for both pure and mixed initial states ρ. Fifty simulations were performed to generate each point. The mean number τ of σ-steps with typical standard deviations indicated by error bars is plotted on the vertical axis. The solid lines correspond to control by Kraus maps, and the dashed lines correspond to control by unitary maps. The lines marked by circles correspond to pure state ρ, and the unmarked lines correspond to mixed ρ. Similar behavior is observed for path length (not shown).

$l = 1, \ldots, n$, $j = 1, \ldots, N^2$ and $B_l \in \mathcal{M}_N$. Since $\mathrm{Tr}\big(B_l^\dagger \tilde{K}_j\big) = \mathrm{Tr}\big(\sum_{i=1}^{N^2} u_{ji} B_l^\dagger K_i\big) = 0$ by linearity of the trace operation, this constraint is $\mathcal{W}$-invariant and the set of Kraus matrices satisfying this constraint forms a $\mathcal{W}$-invariant subset of the Stiefel manifold.

5.1. The numerical procedure

The constraints $\mathrm{Tr}\big(B_i^\dagger K_j\big) = 0$ with $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant N^2$ can be rewritten as a set of $2nN^2$ constraints $h_k : \mathcal{S} \to \mathbb{R}$ defined as follows. Let $\tilde{G}_l \in \mathcal{M}(N^3, N, \mathbb{C})$ for $l = n(j-1) + i$ be the matrix with B_i occupying rows $(j-1)N + 1$ through jN and with other matrix elements set to zero. Then define $G_k = \tilde{G}_k$ for $k = 1, \ldots, nN^2$, $G_k = i\tilde{G}_{k-nN^2}$ for $k = nN^2 + 1, \ldots, 2nN^2$ and set $h_k(S) = \langle G_k, S \rangle$. The control goal is to maximize J over $h^{-1}(0)$.

First, we need to find a matrix $\tilde{S} \in \mathcal{S}_h$ which represents an initial control satisfying the constraint. To do this, define

$$f(S) = \sum_{k=1}^{2nN^2} \langle G_k, S \rangle^2. \tag{14}$$

We see that

$$\mathrm{d}_S f(\delta S) = \sum_{k=1}^{2nN^2} 2\langle G_k, \delta S \rangle \langle G_k, S \rangle = \left\langle \sum_{k=1}^{2nN^2} 2\langle G_k, S \rangle G_k, \delta S \right\rangle. \tag{15}$$

Hence, grad $f(S) = \sum_{k=1}^{2nN^2} 2\langle G_k, S\rangle G_k$. Now generate an arbitrary $S_0 \in \mathcal{S}$ and solve the equation

$$\frac{\mathrm{d}S}{\mathrm{d}\sigma} = -\mathcal{P}_S(\mathrm{grad}\, f(S)),$$

with the initial condition $S(0) = S_0$, where $\mathcal{P}_S$ is the orthogonal projector from $\mathcal{M}(N^3, N, \mathbb{C})$ onto $T_S\mathcal{S}$ (see section 3.1). Then, if the landscape of f on $\mathcal{S}$ is trap free, the algorithm will always find a global minimum $\tilde{S}$ of f, which will satisfy the constraint $h(\tilde{S}) = 0$. It is unknown whether this constrained landscape is trap free.

After producing the initial Stiefel matrix $\tilde{S}$, we maximize the objective function J on $\mathcal{S}_h$ by solving the differential equation

$$\frac{\mathrm{d}S}{\mathrm{d}\sigma} = \mathcal{P}_{h,S}(\mathrm{grad}\, J(S)),$$

with the initial condition $S(0) = \tilde{S}$. Here grad $J(S)$ is the gradient of J on $\mathcal{S}$ and $\mathcal{P}_{h,S}$ is a projector from $T_S\mathcal{S}$ onto $T_S\mathcal{S}_h$. The explicit expression for $\mathcal{P}_{h,S}$ is derived in appendix B.

5.2. Numerical results: general linear constraint

It is difficult to derive a general analytical expression for the maximum value of J on the constrained manifold $h^{-1}(0)$ due to the complicated nature of the constraints. For this reason, we cannot determine that the gradient algorithm is stuck at a false trap $\tilde{S}$ (where grad $J(\tilde{S}) = 0$) by simply calculating $J(\tilde{S})$. Therefore, for a fixed constraint h, we performed the optimization procedure ten times using a different initial condition $\tilde{S}(0)$ each time and compared the resultant ten maximal values of the objective function. Let S_i^* be the optimal control on the ith run (where $1 \leqslant i \leqslant 10$), with the corresponding maximum value $J_i^* = J(S_i^*)$. If $J_k^* < J_l^*$ for some k and l, then S_k^* is a false trap. Note that $J_k^* = J_l^*$ for all k and l does not guarantee that the landscape is trap free; the only conclusion is that the ten runs of the algorithm have not found a false trap.

We performed simulations for $N = 2, 3, 4$. For each N, five different initial states ρ were generated, and for each ρ five different collections of matrices $\{B_1, \ldots, B_n\}$ corresponding to five constraints were produced. We consider $n \leqslant N^2 - N - 1$, where $N^2 - N - 1$ represents the maximum number of constraints of special form corresponding to fixing to zero individual matrix elements of the Kraus operators. As a result of the numerical optimization, each of the ten runs performed with initial controls $\tilde{S}_i(0)$ produced the same maximal value $J^* = J_i^*$, and therefore we did not find a false trap. Although this result does not prove the absence of false traps for linear constraints, it indicates that it is surprisingly difficult to find such traps, if they exist.

5.3. Numerical results: fixing to zero individual matrix elements

We now consider a special case of the $\mathcal{W}$-invariant linear constraints such that $h : \mathcal{S} \to \mathbb{R}^{2N^2}$ is the constraint $(K_l)_{ij} = 0$ for all $l = 1, \ldots, N^2$ and for some pair (i, j). The constraint corresponds to setting the jith element in each of the N^2 Kraus operators to zero; we consider the real and imaginary parts separately, hence there are $2N^2$ constraints. Since $\tilde{K}_n = \sum_{m=1}^{N^2} u_{nm} K_m$ defines the $\mathcal{W}$-transformation, $(\tilde{K}_n)_{ij} = 0$ for all n as well. Hence, $h(WS) = h(S)$, and the constraint is $\mathcal{W}$-invariant. More generally, we consider $\mathcal{W}$-invariant constraints of the form

$$(K_l)_{i_q, j_q} \equiv \left(Y_{j_q}^{i_q}\right)_l = 0, \qquad l = 1, \ldots, N^2, \qquad \forall j_q \in I_1, \qquad \forall i_q \in I_2, \qquad q = 1, \ldots, n,$$
(16)

where I_1 and I_2 are two subsets of the set $\{1, 2, \ldots, N\}$ each with n elements.

For such a constraint, equation (8) can be used to determine analytically the optimal value of the objective function J on the constrained set $h^{-1}(0)$:

$$
J_{\text{max}} = \begin{cases} 1 & \text{if} \quad N \notin I_1 \\ 1 - \sum_{j \in I_2} \rho_{jj} & \text{if} \quad N \in I_1. \end{cases}
$$

For each $N = 2, 3, 4, 5$, we fix to zero n matrix elements of every Kraus operator, with n between N and $N^2 - N - 1$ (the maximum possible number of matrix elements which can simultaneously be fixed to zero). For a given n, the optimization procedure was performed 25 times, and a different collection of matrix elements was fixed to zero during each run (i.e., different sets I_1 and I_2 were chosen). The gradient algorithm was able to reach the maximal value J_{max} each time, showing that there appear no false traps in this landscape. If suboptimal maxima were encountered, the algorithm would have got stuck at $J < J_{\text{max}}$, and global optimization could not have been performed. Thus the optimization procedure did not discover any false traps for 25 randomly generated constraints. Again, this could not be taken as conclusive proof of the absence of false traps. Evidently, more complex or demanding constraints are called for to find traps.

6. Conclusion

This paper analyzes the efficiency of optimization over control landscapes for open quantum systems governed by Kraus map evolution. Several conclusions stem from the findings. When Θ is a rank-one projector, which corresponds to the control goal of transforming an initial state ρ into a pure state, the search efficiency primarily depends on the number of nonzero eigenvalues of the initial state. The efficiency is relatively insensitive to the dimension of the quantum system N, provided that the number of populated energy states in the initial density matrix remains constant. As the number of nonzero eigenvalues of ρ rises with N, the search for an optimal control becomes less efficient. This result agrees with the expectation that transforming a high-entropy initial state into a low-entropy final state is a more difficult control problem than controlled transformations between states with similar entropy.

The analysis also reveals that for fixed N, the search efficiency positively correlates with the number of zero eigenvalues of ρ. This result can be extended to a more general principle: when the dimension of the quantum system is fixed, the dimension of the maximum submanifold (the set of Kraus operators that correspond to optimal control) positively correlates with the efficiency of the optimization procedure. This statement agrees with the common intuition that a 'larger' target results in an easier and more efficient search. The scaling behavior with N found in this work is also consistent with that identified with unitary evolution, both dynamically and kinematically [35, 36].

We then showed that incoherent control modeled by Kraus map evolution, under the assumption that any Kraus map can be generated, is more efficient than coherent control modeled by unitary evolution. The larger number of control variables available in incoherent control actually decreases the complexity of the search effort. While the influence of the environment makes the total system ostensibly more complicated, the results show that the ability to control the environment can decrease the search effort.

We also analyzed control landscapes with linear constraints on the Kraus maps. Even with the maximum possible number of linear constraints, false traps were not found. While this result does not prove the absence of false traps, it is nonetheless surprising. In the future work, we would like to investigate the control landscapes for constrained Kraus maps in more detail both numerically and theoretically.

18

The kinematic analysis needs to be extended by a more detailed investigation of the role of the critical structure of the control landscapes on the search effort. In particular, the possible influence of saddle manifolds on the required search effort should be analyzed. This analysis may reveal more subtle structural details about the quantum control landscapes. Also non-topological properties of quantum control landscapes may affect the optimization efficiency. In general, it is necessary to find all essential characteristics of the initial state ρ and the target operator Θ that affect the efficiency of the search. Another important problem is to study the dynamics of controlled open quantum systems with regard to topological and non-topological characteristics of the corresponding dynamical control landscapes. Various specific model systems can be used to study the dependence of search efficiency upon the parameters characterizing the system and environment. The presence or absence of false traps in the dynamical control landscapes should be investigated, including situations with constraints on the dynamical controls.

Acknowledgments

Appendix A. Invariance of the submanifold $\mathcal{S}_\mathcal{U}$ for $\mathrm{d}S/\mathrm{d}\sigma = \mathrm{grad}\, J(S(\sigma))$

Here we show that the submanifold $\mathcal{S}_\mathcal{U} := \{S_U \in \mathcal{S} | \exists U \in \mathcal{U}(N) \text{ such that } K_i = \frac{1}{N}U \text{ for } i = 1, \ldots, N^2\}$ (i.e., all of the Kraus matrices determining a point $S_U \in \mathcal{S}_\mathcal{U}$ are equal to the same constant multiple of some unitary matrix) is invariant for the differential equation $\mathrm{d}S/\mathrm{d}\sigma = \mathrm{grad}\, J(S(\sigma))$.

Let X be a manifold with tangent bundle TX. Consider the differential equation

$$\mathrm{d}x/\mathrm{d}\sigma = f(x(\sigma)), \tag{A.1}$$

where $f : X \to TX$ is a smooth function and $x : [0, 1] \to X$ is a path through X parametrized by the real variable σ. A manifold $Y \subset X$ is called an *invariant submanifold* for the differential equation (A.1) if $x(0) \in Y$ implies that $x(\sigma) \in Y$ for all $\sigma \in [0, 1]$. A compact manifold $Y \subset X$ is an invariant submanifold for (A.1) if and only if $f(x) \in T_x Y$ for each $x \in Y$ [45].

It was shown in section 3.1 that $\delta S \in T_S \mathcal{S}$ if and only if $S^\dagger(\delta S)$ is skew-Hermitian. Therefore, writing δS as a stack of N^2 $N \times N$ matrices $\delta S_1, \ldots, \delta S_{N^2}$, we see that for any $S_U \in \mathcal{S}_\mathcal{U}, \delta S \in T_{S_U}\mathcal{S}_\mathcal{U}$ if the matrix $U^\dagger \sum_{k=1}^{N^2} \delta S_k$ is skew-Hermitian.

Theorem 1. *Let $\mathbf{E}$ be the $N^2 \times 1$ matrix (column vector) with all elements equal to 1 (i.e., $\mathbf{E}(i) = 1$ for all i). Then for any $S = \frac{1}{N}(\mathbf{E} \otimes U) \in \mathcal{S}_\mathcal{U}$, the matrix $Z := U^\dagger[(\mathbf{E}^\dagger \otimes \mathbb{I}_N)\, \mathrm{grad}\, J(S)]$ is skew-Hermitian.*

Proof. Recall that $\mathrm{grad}\, J(S) = (2\mathbb{I}_{N^3} - SS^\dagger)(\mathbb{I}_{N^2} \otimes \Theta)S\rho - S\rho S^\dagger(\mathbb{I}_{N^2} \otimes \Theta)S$. Then

$$
\begin{aligned}
Z = {}& U^\dagger(\mathbf{E}^\dagger \otimes \mathbb{I}_N)\left[\frac{2}{N}\mathbb{I}_{N^3} - \frac{1}{N^3}(\mathbf{E} \otimes U)(\mathbf{E}^\dagger \otimes U^\dagger)\right](\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U)\rho \\
& - U^\dagger \frac{1}{N^3}(\mathbf{E} \otimes U)\rho(\mathbf{E}^\dagger \otimes U^\dagger)(\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U) \\
= {}& U^\dagger\left[\frac{2}{N}(\mathbf{E}^\dagger \otimes \mathbb{I}_N)\mathbb{I}_{N^3} - \frac{1}{N}U(\mathbf{E}^\dagger \otimes U^\dagger)\right](\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U)\rho \\
& - U^\dagger \frac{1}{N}U\rho(\mathbf{E}^\dagger \otimes U^\dagger)(\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U)
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\frac{2}{N} U^\dagger (\mathbf{E}^\dagger \otimes \mathbb{I}_N) \mathbb{I}_{N^3} - \frac{1}{N} (\mathbf{E}^\dagger \otimes U^\dagger) \right] (\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U)\rho \\
&\quad - \frac{1}{N} \rho (\mathbf{E}^\dagger \otimes U^\dagger)(\mathbb{I}_{N^2} \otimes \Theta)(\mathbf{E} \otimes U) \\
&= \left[\frac{2}{N} U^\dagger (\mathbf{E}^\dagger \otimes \mathbb{I}_N) \mathbb{I}_{N^3} - \frac{1}{N} (\mathbf{E}^\dagger \otimes U^\dagger) \right] (\mathbf{E} \otimes \Theta U)\rho - N\rho U^\dagger \Theta U \\
&= 2N U^\dagger \Theta U \rho - N U^\dagger \Theta U \rho - N\rho U^\dagger \Theta U \\
&= N[U^\dagger \Theta U, \rho], \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (\text{A.2})
\end{aligned}
$$

which is skew-Hermitian for Hermitian matrices ρ and Θ. As a result, grad $J(S_U) \in T_{S_U}\mathcal{S}_\mathcal{U}$ for $S_U \in \mathcal{S}_\mathcal{U}$, so $\mathcal{S}_\mathcal{U}$ is an invariant submanifold for the differential equation $\mathrm{d}S/\mathrm{d}\sigma =$ grad $J(S(\sigma))$. $\quad\square$

Appendix B. Derivation of the projector $\mathcal{P}_{h,S}$

Let $\tilde{h} : \mathcal{M}(N^3, N, \mathbb{C}) \to \mathbb{R}^q$ define a constraint on the Stiefel matrices, which restricts the set of admissible controls to $\mathcal{S}_h = \mathcal{S} \cap \tilde{h}^{-1}(0)$. The goal is to find a projector $\mathcal{P}_{h,S} : T_S\mathcal{S} \to T_S\mathcal{S}_h$, such that the gradient of J on $\mathcal{S}_h$ will be $\mathcal{P}_{h,S}(\text{grad } J(S))$.

We will use the following lemma.

Lemma 1. *Let X and Y be Riemannian manifolds and $F : X \to Y$. Suppose that $\mathrm{d}_x F$ is surjective for all $x \in X$. Let P_x be the operator on $T_x X$ defined as $P_x = I - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F$. Then* (**a**) *$P_x$ is a projection (that is, $P_x{}^2 = P_x$) and* (**b**) *$P_x : T_x X \to T_x F^{-1}(F(x))$.*

Proof.

(a) It is straightforward to see that $P_x^2 = P_x$:

$$
\begin{aligned}
P_x^2 &= (I - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F)^2 \\
&= I - 2(\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F \\
&\quad + (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F \circ (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F \\
&= I - 2(\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F + (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F \\
&= I - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F = P_x. \quad\quad\quad\quad\quad\quad (\text{B.1})
\end{aligned}
$$

(b) It is clear that if $\mathrm{d}_x F(z) = 0$, then $z \in T_x F^{-1}(F(x))$. Note that any vector $v \in T_x X$ can be written as $v = z + \mathrm{d}_x F^*(w)$, where w is arbitrary and $z \in T_x F^{-1}(F(x))$. Indeed, let $w = (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F(v)$. Then $\mathrm{d}_x F(z) = \mathrm{d}_x F(v) - \mathrm{d}_x F \circ (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F(v) = 0$, and therefore $z \in T_x F^{-1}(F(x))$.

Now we will show that the image of P_x lies in $T_x F^{-1}(F(x))$. For any $v \in T_x X$, write $v = z + \mathrm{d}_x F^*(w)$, where $z \in T_x F^{-1}(F(x))$. Then $P_x(v) = P_x(z) + P_x \circ (\mathrm{d}_x F)^*(w) = z - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F(z) + (\mathrm{d}_x F)^*(w) - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F \circ (\mathrm{d}_x F)^*(w) = z - (\mathrm{d}_x F)^* \circ (\mathrm{d}_x F \circ (\mathrm{d}_x F)^*)^{-1} \circ \mathrm{d}_x F(z) = z$ since $z \in T_x F^{-1}(F(x))$ by assumption. Hence, the image of P_x lies in $T_x F^{-1}(F(x))$. This proves the lemma. $\quad\square$

Recall now that $\mathcal{P}_S$ is the projector from $\mathcal{M}(N^3, N, \mathbb{C})$ to $T_S\mathcal{S}$. Then $\mathrm{d}_S h = \mathrm{d}_S \tilde{h}|_{T_S\mathcal{S}}$ and $\mathrm{d}_S h^* = \mathcal{P}_S \circ \mathrm{d}_S \tilde{h}^*$. If $\mathrm{d}_S h$ is full-rank, then according to lemma 1 we have a projector $\mathcal{P}_{h,S} : T_S\mathcal{S} \to T_S\mathcal{S}_h$:

$$
\begin{aligned}
\mathcal{P}_{h,S}(\delta S) &= \delta S - \mathrm{d}_S h^* \circ (\mathrm{d}_S h \circ \mathrm{d}_S h^*)^{-1} \circ \mathrm{d}_S h(\delta S) \\
&= \delta S - \mathcal{P}_S \circ \mathrm{d}_S \tilde{h}^* \circ (\mathrm{d}_S \tilde{h} \circ \mathcal{P}_S \circ \mathrm{d}_S \tilde{h}^*)^{-1} \circ \mathrm{d}_S \tilde{h}(\delta S).
\end{aligned}
$$

In what follows, we will restrict our attention to affine maps defined by a set $\tilde{g} = (\tilde{g}_1, \ldots, \tilde{g}_q)$ of bounded linear functionals $\tilde{g}_i : \mathcal{M}(N^3, N, \mathbb{C}) \to \mathbb{R}$. By the Riesz representation theorem, there exist unique matrices $G_i \in \mathcal{M}(N^3, N, \mathbb{C})$ such that $\tilde{g}_i(A) = \langle G_i, A \rangle$ for all $A \in \mathcal{M}(N^3, N, \mathbb{C})$. For a constraint of the form $h_k(S) = \Re\left(Y_j^i\right)_k = 0$ (respectively $\Im\left(Y_j^l\right)_k = 0$) as considered in section 5.3, these matrices have the form $G_k = |d\rangle\langle l|$ (respectively $G_k = i|d\rangle\langle l|$), where $d = (k-1)N^2 + (j-1)N$.

Since γ is constant and $\tilde{g}$ is linear, $\mathrm{d}_S\tilde{h} = \tilde{g}$ and $\mathrm{d}_S\tilde{h}^* = \tilde{g}^*$. To determine a formula for $\tilde{g}^*$, note that for any $y \in \mathbb{R}^q$

$$\langle \tilde{g}^*(y), \delta S \rangle = \langle y, \tilde{g}(\delta S) \rangle = \sum_{i=1}^q y_i \tilde{g}_i(\delta S) = \sum_{i=1}^q y_i \langle G_i, \delta S \rangle = \left\langle \sum_{i=1}^q y_i G_i, \delta S \right\rangle.$$

Therefore, $\tilde{g}^*(y) = \sum_{i=1}^q y_i G_i$. Putting these expressions together gives

$$\mathrm{d}_S\tilde{h} \circ \mathrm{d}_S\tilde{h}^*(y) = \tilde{g}\left(\sum_{i=1}^q y_i \mathcal{P}_S(G_i)\right) = \begin{pmatrix} \langle G_1, \sum_{i=1}^q y_i \mathcal{P}_S(G_i) \rangle \\ \vdots \\ \langle G_q, \sum_{i=1}^q y_i \mathcal{P}_S(G_i) \rangle \end{pmatrix}$$

$$= \sum_{i=1}^q y_i \begin{pmatrix} \langle G_1, \mathcal{P}_S(G_i) \rangle \\ \vdots \\ \langle G_q, \mathcal{P}_S(G_i) \rangle \end{pmatrix} = Zy,$$

where Z has matrix elements $Z_{ij} = \langle G_i, \mathcal{P}_S(G_j) \rangle$. We finally obtain

$$\mathcal{P}_{h,S}(\delta S) = \delta S - \mathcal{P}_S \circ \tilde{g}^*(Z^{-1}\tilde{g}(\delta S)) = \delta S - \sum_{i=1}^q \mathcal{P}_S(G_i)(Z^{-1}\tilde{g}(\delta S))_i$$

$$= \delta S - \sum_{i=1}^q \sum_{j=1}^q \mathcal{P}_S(G_i)(Z^{-1})_{ij}\tilde{g}_j(\delta S) = \delta S - \sum_{i=1}^q \sum_{j=1}^q \mathcal{P}_S(G_i)(Z^{-1})_{ij}\langle G_i, \delta S \rangle.$$

References

[1] Walmsely I and Rabitz H 2003 *Phys. Today* **56** 43
[2] Butkovskiy A G and Samoilenko Yu I 1984 *Control of Quantum-Mechanical Processes and Systems* (Moscow: Nauka)
    Butkovskiy A G and Samoilenko Yu I 1990 *Control of Quantum-Mechanical Processes and Systems* (Dordrecht: Kluwer) (Engl. Transl.)
[3] Tannor D and Rice S A 1985 *J. Chem. Phys.* **83** 5013
[4] Judson R S and Rabitz H 1992 *Phys. Rev. Lett.* **68** 1500
[5] Warren W S, Rabitz H and Dahleh M 1993 *Science* **259** 1581
[6] Rice S A and Zhao M 2000 *Optical Control of Molecular Dynamics* (New York: Wiley)
[7] Rabitz H, de Vivie-Riedle R, Motzkus M and Kompa K 2000 *Science* **288** 824
[8] Shapiro M and Brumer P 2003 *Principles of the Quantum Control of Molecular Processes* (Hoboken, NJ: Wiley-Interscience)
[9] Dantus M and Lozovoy V V 2004 *Chem. Rev.* **104** 1813
[10] Pechen A and Rabitz H 2006 *Phys. Rev.* A **73** 062102
[11] Pechen A and Rabitz H 2008 *QP–PQ Quantum Probability and White Noise Analysis Proc. 28th Conf. on Quantum Probability and Related Topics* ed vol 23 J C Garcia, R Quezada and S B Sontz (Singapore: World Scientific) pp 197–211 (arXiv:0801.3467)
[12] Pechen A and Rabitz H 2009 Vestnik of Samara State University, Mathematical Series, *Proc. Int. Conf. Mathematical Physics and its Applications* (Samara, Russia, 2008)
[13] Romano R and D'Alessandro D 2006 *Phys. Rev.* A **73** 022323
[14] Accardi L and Imafuku K 2006 *QP–PQ: Quantum Probability and White Noise Analysis* vol 19 ed L Accardi *et al* (Singapore: World Scientific) pp 28–45

[15] Linington I E and Garraway B M 2008 *Phys. Rev.* A **77** 033831
[16] Fu H C, Dong H, Liu X F and Sun C P 2009 *J. Phys. A: Math. Theor.* **42** 045303
[17] Vilela Mendes R and Man'ko V I 2003 *Phys. Rev.* A **67** 053404
[18] Mandilara A and Clark J W 2005 *Phys. Rev.* A **71** 013406
[19] Pechen A, Il'in N, Shuang F and Rabitz H 2006 *Phys. Rev.* A **74** 052102
[20] Roa L, Delgado A, Ladron de Guevara M L and Klimov A B 2006 *Phys. Rev.* A **73** 012322
[21] Shuang F, Pechen A, Ho T-S and Rabitz H 2007 *J. Chem. Phys.* **126** 134303
[22] Sugny D and Kontz C 2008 *Phys. Rev.* A **77** 063420
[23] Shuang F, Zhou M, Pechen A, Wu R, Shir O M and Rabitz H 2008 *Phys. Rev.* A **78** 063422
[24] Ding Y *et al* 2007 *Rev. Sci. Instrum.* **78** 023103
[25] Tarasov V E 2002 *J. Phys. A: Math. Gen.* **35** 5207
[26] Mizel A 2008 arXiv:0810.0470
[27] Kraus K 1983 *States, Effects, and Operations* (New York: Springer)
[28] Wu R, Pechen A, Brif C and Rabitz H 2007 *J. Phys. A: Math. Theor.* **40** 5681
[29] Rabitz H, Hsieh M and Rosenthal C M 2004 *Science* **303** 1998
[30] Rabitz H, Hsieh M and Rosenthal C M 2006 *J. Chem. Phys.* **124** 204107
[31] Pechen A, Prokhorenko D, Wu R and Rabitz H 2008 *J. Phys. A: Math. Theor.* **41** 045205
[32] Wu R, Pechen A, Rabitz H, Hsieh M and Tsou B 2008 *J. Math. Phys.* **49** 022108
[33] Chakrabarti R, Wu R and Rabitz H 2007 arXiv:0708.3513
[34] Chakrabarti R and Rabitz H 2007 *Int. Rev. Phys. Chem.* **26** 671
[35] Moore K, Hsieh M and Rabitz H 2008 *J. Chem. Phys.* **128** 154117
[36] Riviello G, Moore K and Rabitz H (at press)
[37] Grace M, Brif C, Rabitz H, Walmsley I A, Kosut R L and Lidar D A 2007 *J. Phys. B: At. Mol. Opt. Phys.* **40** 5103
[38] Accardi L, Lu Y G and Volovich I V 2002 *Quantum Theory and Its Stochastic Limit* (Berlin: Springer)
[39] Choi M-D 1975 *Linear Algebra Appl.* **10** 285
[40] Stiefel E 1935–1936 *Comment. Math. Helv.* **8** 305
[41] Hatcher A 2002 *Algebraic Topology* (Cambridge: Cambridge University Press)
[42] do Carmo M P 1992 *Riemannian Geometry* (Boston: Birkhauser)
[43] Mezzadri F 2007 *Not. Am. Math. Soc.* **54** 592
[44] Devroye L 1986 *Non-Uniform Random Variate Generation* (New York: Springer) p 207
[45] Chicone C 1999 *Ordinary Differential Equations with Applications* (New York: Springer)