

# Characterization of the Critical Sets of Quantum Unitary Control Landscapes

Jason M. Dominy, Tak-San Ho, and Herschel A. Rabitz

**Abstract**—This work considers various families of quantum control landscapes (i.e. objective functions for optimal control) for obtaining target unitary transformations as the general solution of the controlled Schrödinger equation. We examine the critical point structure of the kinematic landscapes  $J_F(U) = \|(U - W)A\|^2$  and  $J_P(U) = \|A\|^4 - |\text{Tr}(AA^\dagger W^\dagger U)|^2$  defined on the unitary group  $U(\mathcal{H})$  of a finite-dimensional Hilbert space  $\mathcal{H}$ . The parameter operator  $A \in \mathcal{B}(\mathcal{H})$  is allowed to be completely arbitrary, yielding an objective function that measures the difference in the actions of  $U$  and the target  $W$  on a subspace of state space, namely the column space of  $A$ . The analysis of this function includes a description of the structure of the critical sets of these kinematic landscapes and characterization of the critical points as maxima, minima, and saddles. In addition, we consider the question of whether these landscapes are Morse-Bott functions on  $U(\mathcal{H})$ . Landscapes based on the intrinsic (geodesic) distance on  $U(\mathcal{H})$  and the projective unitary group  $PU(\mathcal{H})$  are also considered. These results are then used to deduce properties of the critical set of the corresponding dynamical landscapes.

**Index Terms**—Quantum control, quantum information, optimization.

## I. INTRODUCTION

AN important application of quantum optimal control theory is the generation of target quantum logic gates for quantum information processing. The goal of such optimal control is to arrange the dynamics such that the desired logical gate is realized as the final time unitary evolution operator, which is the general solution of the controlled Schrödinger equation. In most applications, the optimization goal is not a single unitary operator, but a family of logically equivalent operators. For example, since the global phase is not observable, the goal may be any unitary operator that is equivalent to the target gate up to global phase. Likewise, in some cases only a subspace of the Hilbert space  $\mathcal{H}$  of states may be used for the quantum register, so that any unitary propagator should be acceptable that acts as the target gate on that subspace. In

contrast to other quantum control problems, for example the maximization of a quantum mechanical observable, there is no unique or natural choice for the objective function against which the optimization is performed. Indeed, *any* smooth function  $J : U(\mathcal{H}) \rightarrow \mathbb{R}$  with a global minimum at the target unitary gate or gates is a candidate objective for the unitary problem. But, as we will see, some choices may exhibit more favorable convergence and other properties.

The theory of quantum control landscapes has been developed over a series of papers, including [1]–[9], as a way to think about the problem of finding optimal solutions within quantum control. This is pursued principally by building up a picture of the topography of the objective function as a landscape over the space of all admissible controls, typically through analysis of the structure of the set of critical points of the objective function. This provides direct information about the gradient flow associated with the landscape. For example, the presence of a local maximum or minimum can act as a “trap” for the gradient flow or its time-reversal, respectively. And although saddles do not trap the flow, the flow can be greatly slowed in close proximity to a saddle. While, for a given objective function, gradient ascent/descent may not be the most efficient method for finding optimal controls, the topography of the landscape and its impact on the behavior of the gradient flow offers insights into the expected performance of classes of algorithms (local deterministic algorithms versus more non-local stochastic algorithms, for example). As a consequence, a quantum control landscape analysis will typically begin with the identification of the set of critical points.

The critical points of the kinematic landscape having been identified, they may then be characterized as local maxima, local minima, and saddles. As has been demonstrated for other classes of kinematic quantum control landscapes, the landscapes considered in this work will turn out to have global maxima and minima, but no other local extrema capable of impeding optimization. Moreover, the critical sets will be shown generally to comprise disjoint submanifolds, and these submanifolds are nondegenerate in the Morse-Bott sense [10]–[12]. In other words, the null space of the Hessian of  $J$  and the tangent space of the critical submanifold coincide at each critical point  $U \in U(\mathcal{H})$ . This condition identifies the kinematic landscape as a Morse-Bott function, which is interesting for at least two reasons. First, certain results about the convergence of the gradient flow may be proved for Morse-Bott functions, in particular that (on a compact manifold) the gradient flow always converges to a critical point [13]. Second, the identification of the null space of the Hessian and the tangent space of the critical submanifold is important for

Manuscript received February 9, 2011; revised MMMMMM DD, 2013. This work was supported, in part, by U.S. Department of Energy (DOE) Contract No. DE-AC02-76-CHO-3073 through the Program in Plasma Science and Technology at Princeton. Partial support is also acknowledged from the DOE grant No. DE-FG02-02ER15344.

J. M. Dominy was with the Program in Applied and Computational Mathematics, Princeton University, Princeton NJ 08544, during the writing of this paper. He is currently with the Center for Quantum Information Science and Technology, University of Southern California, 90089 (e-mail: jdominy@usc.edu)

T.-S. Ho is with the Department of Chemistry, Princeton University, Princeton NJ 08544 (e-mail: tsho@princeton.edu)

H. A. Rabitz is with the Department of Chemistry and the Program in Applied and Computational Mathematics, Princeton University, Princeton NJ 08544 (e-mail: hrabitz@princeton.edu)

Document Object Identifier

certain numerical methods, such as second order D-MORPH [14], that are designed to explore the critical sets.

Let  $\mathbb{K}$  denote the space of admissible control functions. For the present analysis,  $\mathbb{K}$  will be  $L^2([0, T]; \mathbb{R})$ , the space of square integrable real-valued functions on the interval  $[0, T]$ , where  $T$  is some fixed final time over which the controlled dynamics take place. Let  $\mathcal{H}$  be a complex Hilbert space of dimension  $N < \infty$  and let  $\mathcal{B}(\mathcal{H})$  denote the space of bounded linear operators on  $\mathcal{H}$  endowed with the real Hilbert-Schmidt inner product  $\langle A, B \rangle = \Re \text{Tr}(A^\dagger B)$ .  $U(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$  will denote the unitary group on  $\mathcal{H}$  endowed with the Riemannian metric induced by the Hilbert-Schmidt inner product, and  $\mathfrak{u}(\mathcal{H})$  will denote the corresponding Lie algebra of skew-Hermitian operators acting on  $\mathcal{H}$ . Also let  $U_T : \mathbb{K} \rightarrow U(\mathcal{H})$  denote the map, defined implicitly by the Schrödinger equation in the dipole approximation

$$i\hbar \frac{dU}{dt}(t, t_0) = (H_0 - \mu \mathcal{E}(t))U(t, t_0), \quad U(t_0, t_0) = \mathbb{1}, \quad (1)$$

such that  $U_T(\mathcal{E}) = U(T, 0)[\mathcal{E}]$  is the unitary propagator at time  $T$  for the control field  $\mathcal{E} \in \mathbb{K}$ . Finally, for any candidate objective function  $J : U(\mathcal{H}) \rightarrow \mathbb{R}$  (the ‘‘kinematic landscape’’), let  $\tilde{J} : \mathbb{K} \rightarrow \mathbb{R}$  (the ‘‘dynamical landscape’’) be the composition  $\tilde{J} = J \circ U_T$ . Then

$$\text{grad } \tilde{J}(\mathcal{E}) = (d_{\mathcal{E}} U_T)^* \left( \text{grad } J(U_T(\mathcal{E})) \right), \quad (2)$$

where  $(d_{\mathcal{E}} U_T)^*$  is the operator adjoint of the differential  $d_{\mathcal{E}} U_T$ . Much of the important information about the nature of the gradient flow of  $\tilde{J}$  is embodied in the critical points of this landscape, i.e. those fields  $\mathcal{E} \in \mathbb{K}$  for which  $\text{grad } \tilde{J}(\mathcal{E}) = 0$ . Any  $\mathcal{E} \in \mathbb{K}$  such that  $d_{\mathcal{E}} U_T$  is full rank and  $\text{grad } J(U_T(\mathcal{E})) = 0$  (so-called ‘‘regular’’ critical points) will satisfy the condition. There may be other critical points where  $d_{\mathcal{E}} U_T$  is rank-deficient and  $\text{grad } J(U_T(\mathcal{E}))$  may or may not be zero (‘‘singular’’ points). Consideration of such singular points is important for a complete understanding on the dynamical control landscape. However, since  $U_T$  is a highly nonlinear map from an infinite-dimensional space to a finite-dimensional space, singular points are expected to be rare and will not be considered in the present analysis. Singular points and their role in quantum control landscapes have only recently begun to be studied [15]–[18].

Several classes of landscapes for generating target unitary transformations will be considered. They include  $J_F(U) = \|(U - W)A\|^2$  and  $J_P(U) = \|A\|^4 - |\text{Tr}(AA^\dagger W^\dagger U)|^2$  for some fixed target  $W \in U(\mathcal{H})$  and some arbitrary fixed  $A \in \mathcal{B}(\mathcal{H})$ , as well as the corresponding landscapes using the intrinsic (geodesic) distance on the unitary group  $U(\mathcal{H})$  and the projective unitary group  $\text{PU}(\mathcal{H})$ , rather than the norm (Euclidean) distance as in  $J_F$  and  $J_P$ . The parameter operator  $A$ , though arbitrary in our analysis, can have some implications for the landscape topography and should be chosen carefully in application.  $A$  might be chosen to be a partial isometry, for example, so that  $AA^\dagger$  is a projection onto a subspace of  $\mathcal{H}$ . Such a form may be desirable in quantum information when only a subspace of  $\mathcal{H}$  is designated as the quantum register. Alternatively,  $A$  might be chosen to be nondegenerate, which can have the effect of simplifying the landscape topography

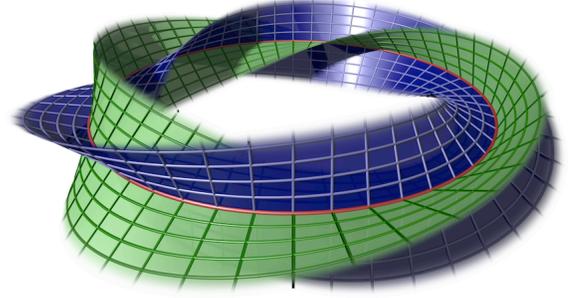


Fig. 1. Example of two orthogonal rank 1 vector bundles (i.e. line bundles) over  $S^1$  in  $\mathbb{R}^3$ . Analogously, each Hessian eigenbundle considered in this paper is a rank  $m$  vector bundle over a critical submanifold of  $U(\mathcal{H})$ , where  $m$  is the multiplicity of the corresponding eigenvalue. It is formed by ‘‘gluing’’ together the  $m$ -dimensional  $\lambda$ -eigenspaces at each point of the critical submanifold for a given Hessian eigenvalue  $\lambda$ .

by making all critical submanifolds zero-dimensional, and perhaps making optimization easier.

Landscapes of the form  $J_F$  and  $J_P$  have been studied in the past [2], [8], [9]. The present paper extends these various works by broadening the families of landscapes under consideration, describing the structures of the critical submanifolds and Hessian eigenbundles (vector bundles formed from the Hessian eigenspaces along critical submanifolds; see Figure 1), and directly addressing the issue of Morse-Bott nondegeneracy of the critical submanifolds.

The paper is organized as follows. Sections II and III describe the critical points of the kinematic landscapes  $J_F$  and  $J_P$ . The two additional landscapes based on geodesic distance are presented and analyzed in Section IV. In Section V, these results are related back to the dynamical landscapes. The overall results are summarized in Section VI. Two appendices are included which provide a proof of the infinite Fréchet differentiability of the control-to-propagator map  $U_T$  and a derivation of the gradient of the geodesic distance landscape on  $U(\mathcal{H})$ .

## II. KINEMATIC CRITICAL POINT ANALYSIS OF CERTAIN PHASE-DEPENDENT LANDSCAPES

For now we put aside the dynamical component of the map and focus just on the critical point analysis of the kinematic map

$$J_F(U) = \|(U - W)A\|^2 = 2\|A\|^2 - 2\Re \text{Tr}(\mathcal{A}^2 W^\dagger U), \quad (3)$$

defined as a function on the unitary group:  $J_F : U(\mathcal{H}) \rightarrow \mathbb{R}$ , where  $\mathcal{A}^2 := AA^\dagger$  is Hermitian and positive semi-definite. In contrast to the landscape  $J_P$  that will be the subject of the next section, the value of  $J_F(U)$  depends upon the global phase of  $U$ . However, if  $\mathcal{A}^2$  has a null space, then  $J_F(U)$  is invariant to the action of  $U$  on that null space. Generally speaking, the parameter operator  $A \in \mathcal{B}(\mathcal{H})$  may be chosen to vary the relative weights of different parts of the unitary operator. This freedom allows for the possibility that  $A$  is a projection operator as discussed in [19].

### A. Critical Point Identification

The operator  $\mathcal{A}^2$  induces a natural orthogonal decomposition of  $\mathcal{H}$  into the eigenspaces of  $\mathcal{A}^2$ ,  $\mathcal{H} = \bigoplus_{i=0}^{\kappa} \mathcal{V}_i$ , under which  $\mathcal{A}^2 = \bigoplus_{i=0}^{\kappa} \omega_i^2 \mathbb{1}_{\mathcal{V}_i}$ , where  $0 = \omega_0^2 < \omega_1^2 < \dots < \omega_{\kappa}^2$  are the eigenvalues of  $\mathcal{A}^2$ . Note that  $\omega_0^2 = 0$  is a special case, and that  $\mathcal{V}_0$ , the null space of  $\mathcal{A}^2$ , may be a trivial (zero-dimensional) subspace, while the other  $\mathcal{V}_i$  are assumed to be nontrivial. This is done because of the special significance of  $\mathcal{V}_0$  in the analysis that follows. The dimension of subspace  $\mathcal{V}_i$  (i.e. the multiplicity of eigenvalue  $\omega_i^2$ ) will be denoted by  $n_i$ .

The remainder of Section II will be concerned with proving the results encapsulated in the following theorem and illustrating them with examples.

**Theorem 1.** *The kinematic objective  $J_F$  is a Morse-Bott function on  $U(\mathcal{H})$  exhibiting  $M = \prod_{i=1}^{\kappa} (n_i + 1)$  connected critical submanifolds, each of which is isomorphic to a product of Grassmann manifolds and a unitary group:*

$$\begin{aligned} & \text{Crit}(J_F) \\ &= \bigsqcup_{\substack{0 \leq \nu_i \leq n_i \\ \forall i=1, \dots, \kappa}} \left\{ U = W \bigoplus_{i=0}^{\kappa} \mathcal{X}_i : \mathcal{X}_i \in U(\mathcal{V}_i) \text{ for } 0 \leq i \leq \kappa, \right. \\ & \quad \text{and } \mathcal{X}_i^2 = \mathbb{1}_{\mathcal{V}_i} \text{ and} \\ & \quad \left. \text{Tr}(\mathcal{X}_i) = n_i - 2\nu_i \text{ for } 1 \leq i \leq \kappa \right\} \end{aligned} \quad (4a)$$

$$\simeq \bigsqcup_{\substack{0 \leq \nu_i \leq n_i \\ \forall i=1, \dots, \kappa}} U(\mathcal{V}_0) \oplus \text{Gr}_{\nu_1}(\mathcal{V}_1) \oplus \dots \oplus \text{Gr}_{\nu_{\kappa}}(\mathcal{V}_{\kappa}), \quad (4b)$$

where  $\text{Gr}_{\nu_i}(\mathcal{V}_i)$  denotes the Grassmann manifold of all  $\nu_i$ -dimensional linear subspaces of  $\mathcal{V}_i$ . The critical submanifold described by a particular set of indices  $(\nu_1, \dots, \nu_{\kappa})$  corresponds to a critical value  $J_F = 4 \sum_{i=1}^{\kappa} \omega_i^2 \nu_i$  and has dimension

$$\mathcal{N}_0 = n_0^2 + 2 \sum_{i=1}^{\kappa} \nu_i (n_i - \nu_i), \quad (5)$$

while the ranks of the negative and positive Hessian eigenbundles (i.e. the numbers of negative and positive Hessian eigenvalues) on this submanifold are

$$\mathcal{N}_- = \sum_{i=1}^{\kappa} \nu_i^2 + 2 \sum_{0 \leq i < j} n_i \nu_j \quad (6)$$

$$\mathcal{N}_+ = \sum_{i=1}^{\kappa} (n_i - \nu_i)^2 + 2 \sum_{0 \leq i < j} n_i (n_j - \nu_j). \quad (7)$$

Of these critical submanifolds, exactly one (corresponding to  $\nu_i = n_i$  for all  $i = 1, \dots, \kappa$ ) is the set of global maxima and one (corresponding to  $\nu_i = 0$  for all  $i = 1, \dots, \kappa$ ) is the set of global minima, both being isomorphic to  $U(n_0)$ . The remaining  $M - 2$  critical submanifolds are all saddles, so that  $J_F$  admits no local traps.

The differential of  $J_F$ ,  $d_U J_F : T_U U(\mathcal{H}) \rightarrow \mathbb{R}$ , is given by

$$\begin{aligned} d_U J_F(\delta U) &= -2\Re \text{Tr}(\mathcal{A}^2 W^\dagger \delta U) \\ &= \langle U \mathcal{A}^2 W^\dagger U - W \mathcal{A}^2, \delta U \rangle, \end{aligned} \quad (8)$$

where the last step includes a projection of  $-2W \mathcal{A}^2$  into the tangent space  $T_U U(\mathcal{H})$ . Therefore,

$$\text{grad } J_F(U) = U \mathcal{A}^2 W^\dagger U - W \mathcal{A}^2 \quad (9)$$

and a critical point of  $J_F$  is a  $U$  such that  $U \mathcal{A}^2 W^\dagger U = W \mathcal{A}^2$ . Let  $\mathcal{X} = W^\dagger U \in U(\mathcal{H})$ . Then the critical point condition becomes  $\mathcal{X}^\dagger \mathcal{A}^2 = \mathcal{A}^2 \mathcal{X}$ . This result also implies that  $\mathcal{X} \mathcal{A}^4 = \mathcal{A}^4 \mathcal{X}$ , since  $\mathcal{X}^\dagger \mathcal{A}^4 \mathcal{X} = (\mathcal{X}^\dagger \mathcal{A}^2)(\mathcal{A}^2 \mathcal{X}) = (\mathcal{A}^2 \mathcal{X})(\mathcal{X}^\dagger \mathcal{A}^2) = \mathcal{A}^4$ . Because the eigendecomposition of  $\mathcal{H}$  induced by  $\mathcal{A}^4$  is the same as for  $\mathcal{A}^2$ ,  $\mathcal{H} = \bigoplus_{i=0}^{\kappa} \mathcal{V}_i$ , the fact that  $\mathcal{X}$  commutes with  $\mathcal{A}^4$  implies that  $\mathcal{X} = \bigoplus_{i=1}^{\kappa} \mathcal{X}_i$  with  $\mathcal{X}_i \in U(\mathcal{V}_i)$  for each  $i$ . Then  $\mathcal{X}^\dagger \mathcal{A}^2 = \mathcal{A}^2 \mathcal{X}$  implies that  $\omega_i^2 \mathcal{X}_i^\dagger = \omega_i^2 \mathcal{X}_i$  for each  $i = 0, \dots, \kappa$ , so that for  $i = 1, \dots, \kappa$ ,  $\mathcal{X}_i = \mathcal{X}_i^\dagger$  is both unitary and Hermitian, and therefore has eigenvalues  $\pm 1$ . In other words, each  $\mathcal{X}_i$  is a unitary involution:  $\mathcal{X}_i^2 = \mathbb{1}_{\mathcal{V}_i}$ . Consequently, for  $i = 1, \dots, \kappa$ , there exists a further orthogonal decomposition of  $\mathcal{V}_i$  into the positive and negative eigenspaces of  $\mathcal{X}_i$ , i.e.  $\mathcal{V}_i = \mathcal{V}_i^- \oplus \mathcal{V}_i^+$ , with respect to which  $\mathcal{X}_i = -\mathbb{1}_{\mathcal{V}_i^-} \oplus \mathbb{1}_{\mathcal{V}_i^+}$ . The dimension of  $\mathcal{V}_i^-$  will be denoted  $\nu_i$ , and the dimension of  $\mathcal{V}_i^+$  is then  $n_i - \nu_i$ . This is equivalent to the statement that  $\text{Tr}(\mathcal{X}_i) = n_i - 2\nu_i$ .

As a result, the set of critical points of  $J_F$  is given by

$$\begin{aligned} \text{Crit}(J_F) &= \left\{ U = W \bigoplus_{i=0}^{\kappa} \mathcal{X}_i : \mathcal{X}_i \in U(\mathcal{V}_i) \text{ for } 0 \leq i \leq \kappa \right. \\ & \quad \left. \text{and } \mathcal{X}_i^2 = \mathbb{1}_{\mathcal{V}_i} \text{ for } 1 \leq i \leq \kappa \right\} \end{aligned} \quad (10a)$$

$$= \bigsqcup_{\substack{0 \leq \nu_i \leq n_i \\ \forall i=1, \dots, \kappa}} C_{\{\nu_i\}} \quad (10b)$$

where the critical submanifold  $C_{\{\nu_i\}}$  is isomorphic to

$$U(\mathcal{V}_0) \oplus \text{Gr}_{\nu_1}(\mathcal{V}_1) \oplus \dots \oplus \text{Gr}_{\nu_{\kappa}}(\mathcal{V}_{\kappa}) \quad (11)$$

and where  $\mathcal{X}_i \in U(\mathcal{V}_i)$  such that  $\mathcal{X}_i^2 = \mathbb{1}_{\mathcal{V}_i}$  is uniquely identified by its  $-1$  eigenspace  $\mathcal{V}_i^-$ , so the space of all such  $\mathcal{X}_i$  with  $\nu_i$ -dimensional  $-1$  eigenspace is isomorphic to the space of all  $\nu_i$ -dimensional subspaces of  $\mathcal{V}_i$ , which is the Grassmannian

$$\text{Gr}_{\nu_i}(\mathcal{V}_i) \simeq \frac{U(\mathcal{V}_i)}{U(\mathcal{V}_i^-) \oplus U(\mathcal{V}_i^+)}. \quad (12)$$

For each  $\nu_i = 0, \dots, n_i$ , the Grassmannian of admissible  $\mathcal{X}_i$  forms a connected submanifold of  $U(\mathcal{V}_i)$ , and since the traces of the  $\mathcal{X}_i$  corresponding to different  $\nu_i$  are different [ $\text{Tr}(\mathcal{X}_i) = n_i - 2\nu_i$ ] the Grassmannians on  $\mathcal{V}_i$  corresponding to different  $\nu_i$  are disconnected. With  $n_i + 1$  disjoint choices for each  $i = 1, \dots, \kappa$ , it is clear that  $J_F$  admits exactly  $M = \prod_{i=1}^{\kappa} (n_i + 1)$  connected critical submanifolds.

### B. Hessian Analysis

Turning to the question of the signatures of these critical points, we extend the gradient vector field  $\text{grad } J_F$  in the obvious way to all of  $\mathcal{B}(\mathcal{H})$  and differentiate to find  $d_U \text{grad } J_F(\delta U) = \delta U \mathcal{A}^2 W^\dagger U + U \mathcal{A}^2 W^\dagger \delta U$ . Projecting this onto the tangent bundle of  $U(\mathcal{H})$  gives the Hessian operator

TABLE I  
EIGENVALUES AND EIGENSPACES OF THE OPERATOR  
 $\mathcal{L}_{U^\dagger} \circ \text{Hess}_{J_F, U} \circ \mathcal{L}_U \in \mathcal{B}(\mathfrak{u}(\mathcal{H}))$  DESCRIBED IN (15), WHERE  $\mathcal{L}_U$   
DENOTES LEFT MULTIPLICATION BY  $U$ .

Subspace of $\mathfrak{u}(\mathcal{H})$	Eigenvalue $\lambda$
$\mathfrak{u}(\mathcal{V}_0)$	0
$\mathfrak{u}(\mathcal{V}_i^\pm)$	$\pm 2\omega_i^2$
$\mathfrak{u}(\mathcal{V}_i)/[\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+)]$	0
$\mathfrak{u}(\mathcal{V}_i^{s_i} \oplus \mathcal{V}_j^{s_j})/[\mathfrak{u}(\mathcal{V}_i^{s_i}) \oplus \mathfrak{u}(\mathcal{V}_j^{s_j})]$	$s_i\omega_i^2 + s_j\omega_j^2$

at  $U$ ,  $\text{Hess}_{J_F, U} \in \mathcal{B}(\text{T}_U \mathcal{U}(\mathcal{H}))$ ,

$$\text{Hess}_{J_F, U}(\delta U) := \nabla_{\delta U} \text{grad } J_F(U) \quad (13a)$$

$$= \frac{1}{2}(\delta U \mathcal{A}^2 W^\dagger U + W \mathcal{A}^2 U^\dagger \delta U + U \mathcal{A}^2 W^\dagger \delta U + \delta U U^\dagger W \mathcal{A}^2), \quad (13b)$$

where  $\nabla_{\delta U}$  denotes the covariant derivative in the direction  $\delta U$  [20], and where we have used the fact that any tangent vector  $\delta U \in \text{T}_U \mathcal{U}(\mathcal{H})$  satisfies  $\delta U^\dagger = -U^\dagger \delta U U^\dagger$ . At a critical point,  $\text{grad } J_F(U) = 0$ , so that  $W^\dagger U$  and  $U^\dagger W$  both commute with  $\mathcal{A}^2$ , and also  $W^\dagger U \mathcal{A}^2 = U^\dagger W \mathcal{A}^2$ . Then the Hessian becomes

$$\text{Hess}_{J_F, U}(\delta U) = \delta U \mathcal{A}^2 W^\dagger U + U \mathcal{A}^2 W^\dagger \delta U. \quad (14)$$

Suppose that  $U \in \mathcal{U}(\mathcal{H})$  is a critical point of  $J_F$ , let  $\mathcal{X} = W^\dagger U$  as before, and let  $Y = U^\dagger \delta U \in \mathfrak{u}(\mathcal{H})$ , where  $\mathfrak{u}(\mathcal{H})$  denotes the Lie algebra of skew-Hermitian operators on  $\mathcal{H}$ . Then  $\delta U$  is an eigenvector of  $\text{Hess}_{J_F, U}$ , i.e.  $\text{Hess}_{J_F, U}(\delta U) = \lambda \delta U$ , if and only if

$$Y \mathcal{A}^2 \mathcal{X} + \mathcal{A}^2 \mathcal{X} Y = \lambda Y. \quad (15)$$

We will use again the decomposition  $\mathcal{H} = \bigoplus_{i=0}^{\kappa} \mathcal{V}_i$  into eigenspaces of  $\mathcal{A}^2$  and the further decomposition at a critical point  $U = W \bigoplus \mathcal{X}_i$  of  $\mathcal{V}_i = \mathcal{V}_i^- \oplus \mathcal{V}_i^+$  into eigenspaces of  $\mathcal{X}_i$ . The space  $\mathfrak{u}(\mathcal{H})$  may similarly be decomposed into subspaces of skew-Hermitian operators supported on ‘‘diagonal’’ and ‘‘off-diagonal’’ blocks

$$\begin{aligned} \mathfrak{u}(\mathcal{H}) = & \mathfrak{u}(\mathcal{V}_0) \oplus \bigoplus_{i=1}^{\kappa} \mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+) \oplus \frac{\mathfrak{u}(\mathcal{V}_i)}{\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+)} \\ & \oplus \bigoplus_{0 \leq i < j}^{\kappa} \bigoplus_{\substack{s_i = \pm \\ s_j = \pm}} \frac{\mathfrak{u}(\mathcal{V}_i^{s_i} \oplus \mathcal{V}_j^{s_j})}{\mathfrak{u}(\mathcal{V}_i^{s_i}) \oplus \mathfrak{u}(\mathcal{V}_j^{s_j})}, \end{aligned} \quad (16)$$

where the sum over  $s_i$  is neglected for  $i = 0$  since  $\mathcal{V}_0$  is not decomposed into  $\pm$  subspaces. It is straightforward to see that the elements of these identified subspaces are eigenvectors of (15) with eigenvalues as in Table I.

Notice that for orthogonal subspaces  $Q$  and  $R$  of dimension  $q$  and  $r$ , respectively,  $\dim[\mathfrak{u}(Q)] = q^2$  and  $\dim[\mathfrak{u}(Q \oplus R)/(\mathfrak{u}(Q) \oplus \mathfrak{u}(R))] = (q+r)^2 - q^2 - r^2 = 2qr$ . Then the Hessian null space is

$$E_0 = U \left( \mathfrak{u}(\mathcal{V}_0) \oplus \bigoplus_{i=1}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i)}{\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+)} \right) \quad (17)$$

which is identical to the tangent space of the associated critical submanifold, and consequently has the same dimension

$$\mathcal{N}_0 = n_0^2 + 2 \sum_{i=1}^{\kappa} \nu_i (n_i - \nu_i). \quad (18)$$

Therefore  $J_F$  is a Morse-Bott function for all  $A$  matrices. The negative Hessian eigenspace (i.e. the negative Hessian eigenbundle at  $U$ , see Figure 1), spanned by the Hessian eigenspaces with strictly negative eigenvalues, is

$$E_- = U \left( \bigoplus_{i=1}^{\kappa} \mathfrak{u}(\mathcal{V}_i^-) \oplus \bigoplus_{0 \leq i < j}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i \oplus \mathcal{V}_j^-)}{\mathfrak{u}(\mathcal{V}_i) \oplus \mathfrak{u}(\mathcal{V}_j^-)} \right) \quad (19)$$

which has dimension

$$\mathcal{N}_- = \sum_{i=1}^{\kappa} \nu_i^2 + 2 \sum_{0 \leq i < j}^{\kappa} n_i \nu_j. \quad (20)$$

Finally, the positive Hessian eigenspace is

$$E_+ = U \left( \bigoplus_{i=1}^{\kappa} \mathfrak{u}(\mathcal{V}_i^+) \oplus \bigoplus_{0 \leq i < j}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i \oplus \mathcal{V}_j^+)}{\mathfrak{u}(\mathcal{V}_i) \oplus \mathfrak{u}(\mathcal{V}_j^+)} \right) \quad (21)$$

which has dimension

$$\mathcal{N}_+ = \sum_{i=1}^{\kappa} (n_i - \nu_i)^2 + 2 \sum_{0 \leq i < j}^{\kappa} n_i (n_j - \nu_j). \quad (22)$$

It is easy to see that  $\mathcal{N}_0 + \mathcal{N}_- + \mathcal{N}_+ = N^2 = \dim(\mathcal{U}(\mathcal{H}))$  as expected. Furthermore, we find that  $\mathcal{N}_+ = 0$  if and only if  $\nu_i = n_i$  for all  $i = 1, \dots, \kappa$ , i.e. only at the global maximum  $\{U = -W(\mathcal{X}_0 \oplus \mathbb{1}_{\mathcal{H}/\mathcal{V}_0}) : \mathcal{X}_0 \in \mathcal{U}(\mathcal{V}_0)\}$ . Likewise  $\mathcal{N}_- = 0$  if and only if  $\nu_i = 0$  for all  $i = 1, \dots, \kappa$ , i.e. only at the global minimum  $\{U = W(\mathcal{X}_0 \oplus \mathbb{1}_{\mathcal{H}/\mathcal{V}_0}) : \mathcal{X}_0 \in \mathcal{U}(\mathcal{V}_0)\}$ . So, there are no local traps in the kinematic landscape, and the remaining  $\prod_{i=1}^{\kappa} (n_i + 1) - 2$  critical submanifolds are all saddles.

### C. Examples

**Example 1** ( $\mathcal{A}^2$  is a projection). Suppose that  $\mathcal{A}^2 = A A^\dagger$  is a projection, or equivalently, that  $A$  is a partial isometry. Then  $\mathcal{H} = \mathcal{V}_0 \oplus \mathcal{V}_1$ ,  $\omega_0^2 = 0$ , and  $\omega_1^2 = 1$ . By Theorem 1,  $J_F$  admits exactly  $n_1 + 1$  critical submanifolds, each isomorphic to  $\mathcal{U}(\mathcal{V}_0) \oplus \text{Gr}_{\nu_1}(\mathcal{V}_1)$  and having dimension  $\mathcal{N}_0 = n_0^2 + 2\nu_1(n_1 - \nu_1)$  for  $\nu_1 = 0 \dots, n_1$ . The critical submanifold identified by index  $\nu_1$  has critical value  $J_F = 4\nu_1$  and the Hessian has eigenvalues in the set  $\{-2, -1, 0, +1, +2\}$ . At a critical point  $U \in C_{\nu_1}$ , the negative Hessian eigenspace

$$E_- = E_{-2} \oplus E_{-1} = U \left[ \mathfrak{u}(\mathcal{V}_1^-) \oplus \frac{\mathfrak{u}(\mathcal{V}_0 \oplus \mathcal{V}_1^-)}{\mathfrak{u}(\mathcal{V}_0) \oplus \mathfrak{u}(\mathcal{V}_1^-)} \right] \quad (23)$$

is of dimension  $\mathcal{N}_- = \nu_1^2 + 2n_0\nu_1$ , and the positive Hessian eigenspace

$$E_+ = E_{+2} \oplus E_{+1} = U \left[ \mathfrak{u}(\mathcal{V}_1^+) \oplus \frac{\mathfrak{u}(\mathcal{V}_0 \oplus \mathcal{V}_1^+)}{\mathfrak{u}(\mathcal{V}_0) \oplus \mathfrak{u}(\mathcal{V}_1^+)} \right] \quad (24)$$

is of dimension  $\mathcal{N}_+ = (n_1 - \nu_1)^2 + 2n_0(n_1 - \nu_1)$ . The full set of critical values for this problem is  $\{0, 4, 8, \dots, 4n_1\}$ .

$$\begin{array}{ccc}
\mathrm{U}(1) \xrightarrow{\phi \mapsto \phi \mathbb{1}} \mathrm{U}(\mathcal{H}) & & \mathrm{SU}(\mathcal{H}) \xleftarrow{e^{i2\pi k/N} \mathbb{1} \leftarrow k} \mathbb{Z}/N\mathbb{Z} \\
& \searrow \pi & \swarrow p \\
& & \mathrm{PU}(\mathcal{H})
\end{array}$$

Fig. 2.  $\mathrm{U}(\mathcal{H})$  and  $\mathrm{SU}(\mathcal{H})$  as fibre bundles over  $\mathrm{PU}(\mathcal{H})$ . The two compositions of maps depicted here are exact, i.e. each composition  $f \circ g$  is such that  $\ker(f) = \mathrm{im}(g)$ . Since  $\mathrm{U}(1)\mathbb{1} \subset \mathrm{U}(\mathcal{H})$  is the subgroup of global phase rotations,  $\mathrm{PU}(\mathcal{H})$  may be thought of as the unitary group modulo global phase: two unitary operators  $U, W \in \mathrm{U}(\mathcal{H})$  will be mapped by  $\pi$  to the same operator in  $\mathrm{PU}(\mathcal{H})$  if and only if  $W^\dagger U \in \mathrm{U}(1)\mathbb{1}$ , i.e., if and only if  $U = e^{i\theta}W$  for some  $\theta \in \mathbb{R}$ .

**Example 2** ( $\mathcal{A}^2$  is nondegenerate and nonsingular). Suppose that  $\mathcal{A}^2$  is nondegenerate and nonsingular, so that  $\mathcal{H} = \bigoplus_{i=1}^N \mathcal{V}_i$ ,  $n_i = 1$  for all  $i = 1, \dots, N$ , and  $\omega_1^2 < \dots < \omega_N^2$ . By Theorem 1, the critical set of  $J_F$  comprises exactly  $2^N$  isolated critical points (i.e. zero-dimensional critical submanifolds). For the critical point identified by indices  $(\nu_1, \dots, \nu_N)$ , the critical value is  $J_F = 4 \sum_{i=1}^N \nu_i \omega_i^2$ , the negative Hessian eigenbundle has dimension  $\mathcal{N}_- = \sum_{i=1}^N (2i - 1)\nu_i$ , and the positive eigenbundle has dimension  $\mathcal{N}_+ = N^2 - \sum_{i=1}^N (2i - 1)\nu_i$ .

### III. KINEMATIC CRITICAL POINT ANALYSIS OF CERTAIN PHASE-INVARIANT LANDSCAPES

We now turn our attention to the kinematic landscape

$$J_P(U) := \|A\|^4 - |\mathrm{Tr}(\mathcal{A}^2 W^\dagger U)|^2. \quad (25)$$

This function is phase-invariant, meaning that  $J_P(e^{i\theta}U) = J_P(U)$  for any  $\theta \in \mathbb{R}$ . Since the global phase of a state vector  $|\psi\rangle \in \mathcal{H}$  has no physical meaning, neither does the global phase of the unitary propagator, so that  $U$  and  $e^{i\theta}U$  are functionally equivalent. A phase-invariant objective function such as  $J_P$  which treats such equivalent operators as equally optimal may therefore be desirable as it may be expected to require optimization only with respect to the degrees of freedom that are physically relevant.

The remainder of Section III will be concerned with proving the following theorem.

**Theorem 2.** *The critical set of the kinematic objective function  $J_P$  comprises a global maximum set and  $M$  connected nondegenerate critical submanifolds. The global maximum set need not globally be a submanifold of  $\mathrm{U}(\mathcal{H})$ , but away from self-intersection points is a codimension 2 submanifold of  $\mathrm{U}(\mathcal{H})$ . The remaining critical submanifolds are of the form*

$$C_{\{\nu_i\}} = \left\{ U = \phi W Z : \phi \in \mathrm{U}(1), Z \in \bigoplus_{i=0}^{\kappa} \mathrm{U}(\mathcal{V}_i), \text{ and } Z_i^2 = \mathbb{1}_{\mathcal{V}_i} \text{ with } \mathrm{Tr}(Z_i) = n_i - 2\nu_i \text{ for } i \geq 1 \right\} \quad (26)$$

$$\simeq \mathrm{U}(\mathcal{V}_0) \oplus \left[ \mathrm{U}(1) \times \bigoplus_{i=1}^{\kappa} \mathrm{Gr}_{\nu_i}(\mathcal{V}_i) \right], \quad (27)$$

where  $0 \leq \nu_i \leq n_i$  are such that  $\sum_i (n_i - 2\nu_i)\omega_i^2 > 0$ . The number,  $M$ , of these critical submanifolds is equal to the number of choices of these integers  $\{\nu_i\}$  satisfying the above two conditions and therefore depends on the singular values  $\{\omega_i\}$  of the parameter operator  $A$ . The critical submanifold  $C_{\{\nu_i\}}$  described by a particular set of indices  $(\nu_1, \dots, \nu_\kappa)$  corresponds to a critical value of

$$J_P = 4 \left( \sum_{i=1}^{\kappa} \nu_i \omega_i^2 \right) \left( \sum_{i=1}^{\kappa} (n_i - \nu_i) \omega_i^2 \right) \quad (28)$$

and has dimension

$$\mathcal{N}_0 = 1 + n_0^2 + 2 \sum_{i=1}^{\kappa} \nu_i (n_i - \nu_i), \quad (29)$$

while the dimensions of the negative and positive Hessian eigenbundles on this submanifold are

$$\mathcal{N}_- = \sum_{i=1}^{\kappa} \nu_i^2 + 2 \sum_{0 \leq i < j} n_i \nu_j \quad (30)$$

$$\mathcal{N}_+ = -1 + \sum_{i=1}^{\kappa} (n_i - \nu_i)^2 + 2 \sum_{0 \leq i < j} n_i (n_j - \nu_j). \quad (31)$$

Consequently, of these submanifolds  $C_{\{\nu_i\}}$ , exactly one (corresponding to the case  $\nu_i = 0$  for all  $i = 1, \dots, \kappa$ ) is the set of global minima  $\{U = \phi W (Z_0 \oplus \mathbb{1}_{\mathcal{H}/\mathcal{V}_0}) : \phi \in \mathrm{U}(1) \text{ and } Z_0 \in \mathrm{U}(\mathcal{V}_0)\}$ , and the remaining  $M - 1$  critical submanifolds are all saddles, so that  $J_P$  admits no local traps. For an open, dense set of  $A$  operators in  $\mathcal{B}(\mathcal{H})$ , the global maximum set of  $J_P$  is a nondegenerate submanifold of  $\mathrm{U}(\mathcal{H})$ , in which case  $J_P$  is a Morse-Bott function.

#### A. Distance Metric on $\mathrm{PU}(\mathcal{H})$

There are various ways of deriving a phase-invariant landscape like  $J_P$  from one that is phase-dependent like  $J_F$ . A simple approach is to observe that

$$\min_{\phi \in \mathrm{U}(1)} \|(U - \phi W)A\|^2 = 2\|A\|^2 - 2|\mathrm{Tr}(\mathcal{A}^2 W^\dagger U)|. \quad (32)$$

This provides a means to define a quotient metric on the projective unitary group  $\mathrm{PU}(\mathcal{H})$  (see Figure 2) from the metric  $d(U, W) = \|(U - W)A\|^2$  on  $\mathrm{U}(\mathcal{H})$ .

Another approach involves the adjoint representation of the unitary group,  $\mathrm{Ad} : \mathrm{U}(\mathcal{H}) \rightarrow \mathrm{Aut}(\mathfrak{u}(\mathcal{H})) \subset \mathrm{GL}(\mathfrak{u}(\mathcal{H})) \cong \mathrm{GL}(N^2; \mathbb{R})$ , which is given by  $\mathrm{Ad}(U)A = UAU^\dagger$  for any  $A \in \mathfrak{u}(\mathcal{H})$ , and where  $\mathrm{Aut}(\mathfrak{u}(\mathcal{H}))$  is the group of Lie algebra automorphisms on  $\mathfrak{u}(\mathcal{H})$  [21], [22]. With  $\mathfrak{u}(\mathcal{H})$  given the Hilbert-Schmidt inner product,  $\langle \mathrm{Ad}(U)A, \mathrm{Ad}(U)B \rangle = \langle UAU^\dagger, UBU^\dagger \rangle = \langle A, B \rangle$ , so that for each  $U \in \mathrm{U}(\mathcal{H})$ ,  $\mathrm{Ad}(U)$  is an orthogonal operator on  $\mathfrak{u}(\mathcal{H})$ , i.e.  $\mathrm{Ad} : \mathrm{U}(\mathcal{H}) \rightarrow \mathrm{SO}(\mathfrak{u}(\mathcal{H}))$ . Furthermore, if  $W$  and  $U$  differ only by a global phase, i.e.  $W = \phi U$ , then  $\mathrm{Ad}(U) = \mathrm{Ad}(W)$ . Moreover, the kernel of  $\mathrm{Ad}$ , i.e.  $\mathrm{Ad}^{-1}(\mathrm{id})$ , is the center of  $\mathrm{U}(\mathcal{H})$  [23, Cor. 5.2, pg. 129] [21, Thm. 3.50] which is  $Z(\mathrm{U}(\mathcal{H})) = \{\phi \mathbb{1}\} = \mathrm{U}(1)\mathbb{1}$ . Then  $\mathrm{Ad}(U) = \mathrm{Ad}(W)$  if and only if  $W = \phi U$ , so the image of  $\mathrm{Ad}$  is a faithful representation of  $\mathrm{PU}(\mathcal{H})$  and  $\mathrm{Ad}$  may be thought of as playing a role similar to the projection  $\pi : \mathrm{U}(\mathcal{H}) \rightarrow \mathrm{PU}(\mathcal{H})$ .

Consider some target  $W \in U(\mathcal{H})$  and some  $B \in GL(\mathfrak{u}(\mathcal{H}))$  and define

$$J(U) = \frac{1}{2} \|(\text{Ad}(U) - \text{Ad}(W)) \circ B\|_{\text{HS}}^2, \quad (33)$$

where  $\|\cdot\|_{\text{HS}}$  is the Hilbert-Schmidt norm on  $\text{End}(\mathfrak{u}(\mathcal{H})) \cong \mathbb{R}^{N^2 \times N^2}$ , the space of all linear operators acting on  $\mathfrak{u}(\mathcal{H})$ . Then,

$$J(U) = \text{Tr}(B^*B) - \text{Tr}(B^* \text{Ad}(W)^* \text{Ad}(U)B). \quad (34)$$

Let  $A \in \mathcal{B}(\mathcal{H})$  be an arbitrary linear operator on  $\mathcal{H}$ , and let  $B$  be defined by  $B(\Omega) = A\Omega A^\dagger$ . It follows from (34) that the kinematic landscape  $J_P(U) = \|A\|^4 - |\text{Tr}(\mathcal{A}^2 W^\dagger U)|^2$  on  $U(\mathcal{H})$  is equivalent to the weighted Hilbert-Schmidt distance function on the subgroup of  $\text{SO}(\mathfrak{u}(\mathcal{H}))$  given by  $\text{Im}(\text{Ad}) \simeq \text{PU}(\mathcal{H})$ . In other words,  $J_P$  is completely equivalent to  $J_F$ , but applied to  $\text{PU}(\mathcal{H})$ , rather than  $U(\mathcal{H})$ .

### B. Critical Point Identification

Now, the differential of  $J_P$  at  $U \in U(\mathcal{H})$ ,  $d_U J_P : T_U U(\mathcal{H}) \rightarrow \mathbb{R}$  is given by

$$\begin{aligned} d_U J_P(\delta U) &= -\text{Tr}(\mathcal{A}^2 W^\dagger \delta U) \text{Tr}(U^\dagger W \mathcal{A}^2) \\ &\quad - \text{Tr}(\mathcal{A}^2 W^\dagger U) \text{Tr}(\delta U^\dagger W \mathcal{A}^2) \end{aligned} \quad (35a)$$

$$\begin{aligned} &= \langle \text{Tr}(U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger U \\ &\quad - \text{Tr}(\mathcal{A}^2 W^\dagger U) W \mathcal{A}^2, \delta U \rangle \end{aligned} \quad (35b)$$

so that

$$\text{grad } J_P(U) = \text{Tr}(U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger U - \text{Tr}(\mathcal{A}^2 W^\dagger U) W \mathcal{A}^2. \quad (36)$$

For  $J_P(U) < \|A\|^4$ ,  $\text{Tr}(U^\dagger W \mathcal{A}^2) \neq 0$ , so  $\text{grad } J_P(U) = 0$  if and only if  $\mathcal{A}^2 Z = Z^\dagger \mathcal{A}^2$ , where

$$Z = \frac{\text{Tr}(U^\dagger W \mathcal{A}^2)}{|\text{Tr}(U^\dagger W \mathcal{A}^2)|} W^\dagger U \in U(\mathcal{H}). \quad (37)$$

This same condition was considered in section II (and [9]), where it was shown to imply that, under the orthogonal decomposition  $\mathcal{H} = \bigoplus_{i=0}^{\kappa} \mathcal{V}_i$  of  $\mathcal{H}$  into the eigenspaces of dimensions  $\{n_i\}$  of  $\mathcal{A}^2$ ,  $Z = \bigoplus_{i=0}^{\kappa} Z_i$  with  $Z_i \in U(\mathcal{V}_i)$  for  $i = 0, \dots, \kappa$  and  $Z_i^2 = \mathbb{1}_{\mathcal{V}_i}$  for  $i \geq 1$ . Since the involutions  $Z_i$  for  $i \geq 1$  have eigenvalues  $\pm 1$ , they induce a further orthogonal decomposition of  $\mathcal{V}_i$  into  $\mathcal{V}_i = \mathcal{V}_i^- \oplus \mathcal{V}_i^+$  into the  $\pm 1$  eigenspaces of  $Z_i$  of dimensions  $\nu_i$  and  $n_i - \nu_i$ , respectively.

Drawing on the material above, we find that any critical point  $U$  of  $J_P$  with  $J_P(U) < \|A\|^4$  can be written as

$$U = \frac{\text{Tr}(\mathcal{A}^2 W^\dagger U)}{|\text{Tr}(U^\dagger W \mathcal{A}^2)|} W Z \quad (38)$$

with  $Z = \bigoplus_{i=0}^{\kappa} Z_i$  and  $Z_i^2 = \mathbb{1}_{\mathcal{V}_i}$  for  $i \geq 1$ . This characterization is complicated by the presence of  $U$  on both sides of the equation, especially with regard to the phase factor on the right hand side. However, it may be observed for any  $U = \phi W Z$  with  $\phi \in U(1)$ ,  $Z \in \bigoplus_{i=0}^{\kappa} U(\mathcal{V}_i)$ , and  $Z_i^2 = \mathbb{1}_{\mathcal{V}_i}$  for  $i \geq 1$ , that  $[U^\dagger W, \mathcal{A}^2] = [W^\dagger U, \mathcal{A}^2] = 0$  and  $U^\dagger W \mathcal{A}^2 = (\mathcal{A}^2 W^\dagger U)^\dagger = \phi^{-2} \mathcal{A}^2 W^\dagger U$ , so that  $\text{grad } J_P(U) = 0$ . Hence,

every such  $U$  is a critical point of  $J_P$ , and they comprise connected critical sets

$$\begin{aligned} C_{\{\nu_i\}} := \left\{ U = \phi W Z : \phi \in U(1), Z \in \bigoplus_{i=0}^{\kappa} U(\mathcal{V}_i), \text{ and} \right. \\ \left. Z_i^2 = \mathbb{1}_{\mathcal{V}_i} \text{ with } \dim(\mathcal{V}_i^-) = \nu_i \text{ for } i \geq 1 \right\} \end{aligned} \quad (39)$$

for all  $0 \leq \nu_i \leq n_i$ . However, it may be observed that for any such set of indices  $\{\nu_i\}$ ,  $C_{\{\nu_i\}} = C_{\{n_i - \nu_i\}}$ , since  $U = \phi W Z \in C_{\{\nu_i\}}$  if and only if  $U = (-\phi) W (-Z) \in C_{\{n_i - \nu_i\}}$ . It suffices then to only consider  $C_{\{\nu_i\}}$  for which  $\text{Tr}(\mathcal{A}^2 Z) = \sum_{i=1}^{\kappa} \omega_i^2 (n_i - 2\nu_i) > 0$  to avoid identifying the same critical submanifold twice. Such a critical submanifold  $C_{\{\nu_i\}}$  has the critical value

$$J_P(U) = (\text{Tr}(\mathcal{A}^2))^2 - |\text{Tr}(\mathcal{A}^2 W^\dagger U)|^2 \quad (40a)$$

$$= \left( \sum_{i=1}^{\kappa} n_i \omega_i^2 \right)^2 - \left( \sum_{i=1}^{\kappa} (n_i - 2\nu_i) \omega_i^2 \right)^2 \quad (40b)$$

$$= 4 \left( \sum_{i=1}^{\kappa} \nu_i \omega_i^2 \right) \left( \sum_{i=1}^{\kappa} (n_i - \nu_i) \omega_i^2 \right) \quad (40c)$$

for every  $U \in C_{\{\nu_i\}}$ .

### C. Hessian Analysis

Given the form of the gradient of  $J_P$  in (36), by again extending the gradient vector field to  $\mathcal{B}(\mathcal{H})$  and differentiating, it is found that

$$\begin{aligned} d_U \text{grad } J_P(\delta U) &= \text{Tr}(\delta U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger U + \text{Tr}(U^\dagger W \mathcal{A}^2) \delta U \mathcal{A}^2 W^\dagger U \\ &\quad + \text{Tr}(U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger \delta U - \text{Tr}(\mathcal{A}^2 W^\dagger \delta U) W \mathcal{A}^2, \end{aligned} \quad (41)$$

whence, by projection onto the tangent bundle of  $U(\mathcal{H})$ ,

$$\begin{aligned} \text{Hess}_{J_P, U}(\delta U) &= \nabla_{\delta U} \text{grad } J_P \\ &= -\text{Tr}(\mathcal{A}^2 W^\dagger \delta U) W \mathcal{A}^2 + \text{Tr}(\delta U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger U \\ &\quad + \frac{1}{2} \left\{ \text{Tr}(U^\dagger W \mathcal{A}^2) \delta U \mathcal{A}^2 W^\dagger U + \text{Tr}(\mathcal{A}^2 W^\dagger U) W \mathcal{A}^2 U^\dagger \delta U \right. \\ &\quad \left. + \text{Tr}(U^\dagger W \mathcal{A}^2) U \mathcal{A}^2 W^\dagger \delta U + \text{Tr}(\mathcal{A}^2 W^\dagger U) \delta U U^\dagger W \mathcal{A}^2 \right\}. \end{aligned} \quad (42)$$

On one of the critical submanifolds  $C_{\{\nu_i\}}$ , the Hessian is given by

$$\begin{aligned} \text{Hess}_{J_P, U}(\delta U) &= \text{Tr}(\mathcal{A}^2 W^\dagger U) \delta U U^\dagger W \mathcal{A}^2 \\ &\quad + \text{Tr}(\mathcal{A}^2 W^\dagger U) W \mathcal{A}^2 U^\dagger \delta U \\ &\quad - 2 \text{Tr}(\mathcal{A}^2 W^\dagger \delta U) W \mathcal{A}^2. \end{aligned} \quad (43)$$

Writing a critical  $U \in C_{\{\nu_i\}}$  as  $U = \phi W Z$  and letting  $Y = U^\dagger \delta U$ , the Hessian eigenvalue problem  $\text{Hess}_{J_P, U}(\delta U) = \lambda \delta U$  can be written as an eigenvector problem on  $\mathfrak{u}(\mathcal{H})$  as

$$\text{Tr}(\mathcal{A}^2 Z) [Y Z \mathcal{A}^2 + Z \mathcal{A}^2 Y] - 2 \text{Tr}(\mathcal{A}^2 Z Y) Z \mathcal{A}^2 = \lambda Y. \quad (44)$$

Observe that  $\text{Tr}(\mathcal{A}^2 Z Y) = 0$  for any  $Y$  in  $\mathfrak{u}(\mathcal{V}_0)$ ,  $\mathfrak{su}(\mathcal{V}_i^\pm)$  for  $i = 1, \dots, \kappa$ ,  $\mathfrak{u}(\mathcal{V}_i) / (\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+))$  for  $i = 1, \dots, \kappa$ , or  $\mathfrak{U}(\mathcal{V}_i^{s_i} \oplus \mathcal{V}_j^{s_j}) / [\mathfrak{U}(\mathcal{V}_i^{s_i}) \oplus \mathfrak{U}(\mathcal{V}_j^{s_j})]$  for  $0 \leq i < j$  and

TABLE II  
SOME EIGENVALUES AND EIGENSPACES OF THE OPERATOR  
 $\mathcal{L}_{U^\dagger} \circ \text{Hess}_{J_{P,U}} \circ \mathcal{L}_U \in \mathcal{B}(\mathfrak{u}(\mathcal{H}))$  DESCRIBED IN (44), WHERE  $\mathcal{L}_U$   
DENOTES LEFT MULTIPLICATION BY  $U$ .

Subspace of $\mathfrak{u}(\mathcal{H})$	Eigenvalue $\lambda$
$\mathfrak{u}(\mathcal{V}_0)$	0
$\mathfrak{su}(\mathcal{V}_i^\pm)$	$\pm 2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z)$
$\mathfrak{u}(\mathcal{V}_i)/[\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+)]$	0
$\mathfrak{u}(\mathcal{V}_i^{s_i} \oplus \mathcal{V}_j^{s_j})/[\mathfrak{u}(\mathcal{V}_i^{s_i}) \oplus \mathfrak{u}(\mathcal{V}_j^{s_j})]$	$(s_i\omega_i^2 + s_j\omega_j^2) \text{Tr}(\mathcal{A}^2 Z)$

$s_i, s_j \in \{\pm\}$ . So for  $Y$  in any of these subspaces of  $\mathfrak{u}(\mathcal{H})$ , the eigenvalue problem (44) becomes

$$\text{Tr}(\mathcal{A}^2 Z)[YZ\mathcal{A}^2 + Z\mathcal{A}^2 Y] = \lambda Y, \quad (45)$$

which means, as in Section II-B, that each element of these subspaces is an eigenvector as in Table II. The only subspace of  $\mathfrak{u}(\mathcal{H})$  not covered by these cases is the subspace  $S \subset \mathfrak{u}(\mathcal{H})$  spanned by elements  $Y \in \mathfrak{u}(\mathcal{H})$  of the form  $Y = \bigoplus_{i=1}^{\kappa} (\alpha_i^- \mathbb{1}_{\mathcal{V}_i^-} \oplus \alpha_i^+ \mathbb{1}_{\mathcal{V}_i^+})$  for imaginary numbers  $\{\alpha_i^\pm\}$ . For such a  $Y$ , (44) is block diagonal with  $\mathcal{V}_i^\pm$  diagonal block

$$\pm \omega_i^2 (2 \text{Tr}(\mathcal{A}^2 Z) \alpha_i^\pm - 2 \text{Tr}(\mathcal{A}^2 Z Y)) \mathbb{1}_{\mathcal{V}_i^\pm} = \lambda \alpha_i^\pm \mathbb{1}_{\mathcal{V}_i^\pm}. \quad (46)$$

Solving for  $\alpha_i^\pm$ , we find that, for  $\lambda \neq \pm 2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z)$  for all  $\mathcal{V}_i^\pm$  such that  $\dim(\mathcal{V}_i^\pm) > 0$ ,

$$\alpha_i^\pm = \frac{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z Y)}{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z) \mp \lambda}. \quad (47)$$

Then for  $Y = \bigoplus_{i=1}^{\kappa} (\alpha_i^- \mathbb{1}_{\mathcal{V}_i^-} \oplus \alpha_i^+ \mathbb{1}_{\mathcal{V}_i^+}) \in S$ , it follows that

$$\text{Tr}(\mathcal{A}^2 Z Y) = \sum_{i=1}^{\kappa} \omega_i^2 (-\nu_i \alpha_i^- + (n_i - \nu_i) \alpha_i^+) \quad (48a)$$

$$= \text{Tr}(\mathcal{A}^2 Z Y) \sum_{i=1}^{\kappa} 2\omega_i^4 \left( -\frac{\nu_i}{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z) + \lambda} + \frac{n_i - \nu_i}{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z) - \lambda} \right) \quad (48b)$$

which implies that either  $\text{Tr}(\mathcal{A}^2 Z Y) = 0$  or  $f(\lambda) = 1$  where

$$f(\lambda) := \sum_{i=1}^{\kappa} \left( -\frac{2\omega_i^4 \nu_i}{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z) + \lambda} + \frac{2\omega_i^4 (n_i - \nu_i)}{2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z) - \lambda} \right). \quad (49)$$

It may be observed that  $f(\lambda) \rightarrow 0$  as  $\lambda \rightarrow \pm\infty$ , that  $f$  is an increasing function away from its poles, and that  $f$  has a simple pole at  $-2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z)$  for any  $i = 1, \dots, \kappa$  for which  $\nu_i > 0$ , and a simple pole at  $2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z)$  for any  $i$  for which  $\nu_i < n_i$  (see Fig. 3). The number of distinct poles is then equal to the dimension of the subspace  $S \in \mathfrak{u}(\mathcal{H})$  under consideration; furthermore, if  $\nu_i = n_i$  for all  $i$ , then  $\text{Tr}(\mathcal{A}^2 Z) < 0$  which was disallowed by convention, so  $f$  must have at least one positive pole. It is then clear that  $f(\lambda) = 1$  has  $\dim(S)$  distinct solutions: one less than the smallest pole, and one between each pair of adjacent poles. Moreover, it may be seen that  $\lambda = 0$  is one of these solutions, corresponding to the eigenvector  $Y^{(0)} = i \bigoplus_{i=1}^{\kappa} \mathbb{1}_{\mathcal{V}_i}$  of (44). The number of solutions  $f(\lambda) = 1$  with  $\lambda < 0$  is then equal to the number

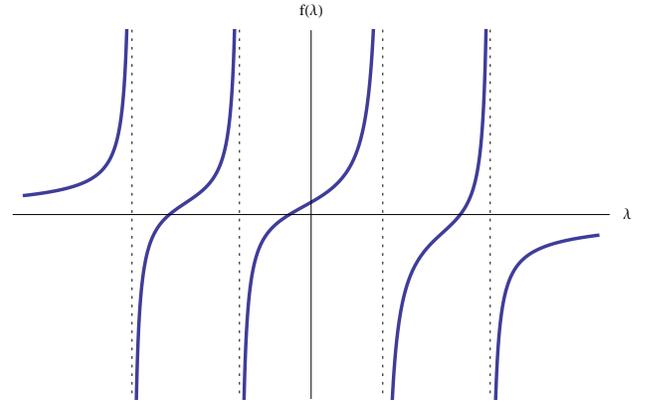


Fig. 3. A depiction of the general properties of  $f(\lambda)$  defined in Eq.(49), namely that  $f(\lambda) \rightarrow 0$  as  $\lambda \rightarrow \pm\infty$ ,  $f(0) = 1$ ,  $f$  exhibits simple poles at some subset of  $\{\pm 2\omega_i^2 \text{Tr}(\mathcal{A}^2 Z)\}$ , and that  $f$  is strictly increasing away from these poles.

$\nu_+$  of  $i \in \{1, \dots, \kappa\}$  for which  $\nu_i > 0$ , and the number of solutions  $f(\lambda) = 1$  with  $\lambda > 0$  is one less than the number  $\nu_-$  of  $i \in \{1, \dots, \kappa\}$  for which  $\nu_i < n_i$ . To each of these solutions  $\lambda$ , we may associate the eigenvector

$$Y^{(\lambda)} = \bigoplus_{i=1}^{\kappa} (\alpha_i^- \mathbb{1}_{\mathcal{V}_i^-} \oplus \alpha_i^+ \mathbb{1}_{\mathcal{V}_i^+}) \in S \quad (50)$$

with

$$\alpha_j^\pm = \frac{2i\omega_j^2 \text{Tr}(\mathcal{A}^2 Z)}{2\omega_j^2 \text{Tr}(\mathcal{A}^2 Z) \mp \lambda}. \quad (51)$$

So, to summarize, at a critical point  $U = \phi W Z \in C_{\{\nu_i\}}$ , the Hessian null space is

$$E_0 = U \left( \mathfrak{u}(\mathcal{V}_0) \oplus \left[ \mathfrak{u}(1) \mathbb{1}_{\mathcal{H}/\mathcal{V}_0} \times \bigoplus_{i=1}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i)}{\mathfrak{u}(\mathcal{V}_i^-) \oplus \mathfrak{u}(\mathcal{V}_i^+)} \right] \right) \quad (52)$$

which is readily seen to be identical to the tangent space of  $C_{\{\nu_i\}}$  and is of dimension

$$\mathcal{N}_0 = 1 + n_0^2 + 2 \sum_{i=1}^{\kappa} \nu_i (n_i - \nu_i). \quad (53)$$

The negative Hessian eigenspace is

$$E_- = U \left( \bigoplus_{\substack{\lambda < 0 \\ f(\lambda)=1}} \mathbb{R} Y^{(\lambda)} \oplus \bigoplus_{i=1}^{\kappa} \mathfrak{su}(\mathcal{V}_i^-) \oplus \bigoplus_{0 \leq i < j}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i \oplus \mathcal{V}_j^-)}{\mathfrak{u}(\mathcal{V}_i) \oplus \mathfrak{u}(\mathcal{V}_j^-)} \right), \quad (54)$$

having dimension

$$\mathcal{N}_- = \sum_{i=1}^{\kappa} \nu_i^2 + 2 \sum_{0 \leq i < j}^{\kappa} n_i \nu_j, \quad (55)$$

and the positive Hessian eigenspace is

$$E_+ = U \left( \bigoplus_{\substack{\lambda > 0 \\ f(\lambda)=1}} \mathbb{R} Y^{(\lambda)} \oplus \bigoplus_{i=1}^{\kappa} \mathfrak{su}(\mathcal{V}_i^+) \oplus \bigoplus_{0 \leq i < j}^{\kappa} \frac{\mathfrak{u}(\mathcal{V}_i \oplus \mathcal{V}_j^+)}{\mathfrak{u}(\mathcal{V}_i) \oplus \mathfrak{u}(\mathcal{V}_j^+)} \right), \quad (56)$$

having dimension

$$\mathcal{N}_+ = -1 + \sum_{i=1}^{\kappa} (n_i - \nu_i)^2 + 2 \sum_{0 \leq i < j}^{\kappa} n_i (n_j - \nu_j). \quad (57)$$

As a result,  $\mathcal{N}_- = 0$  if and only if  $\nu_i = 0$  for all  $i = 1, \dots, \kappa$ , i.e. only at the global minimum

$$\{U = \phi W (Z_0 \oplus \mathbb{1}_{\mathcal{H}/\mathcal{V}_0}) : \phi \in U(1), Z_0 \in U(\mathcal{V}_0)\}. \quad (58)$$

Furthermore,  $\mathcal{N}_+$  can be zero for such a critical point only if  $\sum_{i=1}^{\kappa} (n_i - \nu_i)^2 = 1$  and  $\sum_{0 \leq i < j}^{\kappa} n_i (n_j - \nu_j) = 0$ , i.e. only if  $n_0 = 0$ ,  $\nu_1 = n_1 - 1$ , and  $\nu_i = n_i$  for  $i = 2, \dots, \kappa$ . However, in this case, the constraint  $\text{Tr}(\mathcal{A}^2 Z) > 0$  implies that  $2\omega_1^2 > \sum_{i=1}^{\kappa} n_i \omega_i^2$ , which can only happen in the trivial case  $\kappa = N = 1$  where global phase rotations are the only dynamics and every point is critical with respect to  $J_P$ . Therefore, there are no maxima among the  $C_{\{\nu_i\}}$ . The only maximal points of  $J_P$  belong to the global maximum set considered presently.

#### D. Global Maximum Set

Finally, consider the global maximum set  $\{U : J_P(U) = \|A\|^4\} = \{U : \text{Tr}(\mathcal{A}^2 W^\dagger U) = 0\}$ , which is the intersection of  $U(\mathcal{H})$  with the (complex) hyperplane in  $\mathcal{B}(\mathcal{H})$  orthogonal to  $W\mathcal{A}^2$ . This set does not admit analysis by the methods used thus far, so a different approach is required. Let  $F : U(\mathcal{H}) \rightarrow \mathbb{R}^2$  be given by  $F_1 := \Re \text{Tr}(\mathcal{A}^2 W^\dagger U)$  and  $F_2 := \Im \text{Tr}(\mathcal{A}^2 W^\dagger U)$ . Then,

$$\begin{aligned} d_U F_1(\delta U) &= \Re \text{Tr}(\mathcal{A}^2 W^\dagger \delta U) \\ &= \frac{1}{2} \langle W\mathcal{A}^2 - U\mathcal{A}^2 W^\dagger U, \delta U \rangle \end{aligned} \quad (59a)$$

$$\begin{aligned} d_U F_2(\delta U) &= \Im \text{Tr}(\mathcal{A}^2 W^\dagger \delta U) \\ &= \frac{1}{2} \langle iW\mathcal{A}^2 + iU\mathcal{A}^2 W^\dagger U, \delta U \rangle \end{aligned} \quad (59b)$$

so that the gradients are given by

$$\text{grad } F_1(U) = \frac{1}{2} (W\mathcal{A}^2 - U\mathcal{A}^2 W^\dagger U) \quad (60a)$$

$$\text{grad } F_2(U) = \frac{i}{2} (W\mathcal{A}^2 + U\mathcal{A}^2 W^\dagger U). \quad (60b)$$

Thus,  $dF$  is surjective except where there exists  $(\alpha, \beta) \neq 0 \in \mathbb{R}^2$  such that  $\alpha \text{grad } F_1(U) = \beta \text{grad } F_2(U)$ , i.e. where

$$\phi U^\dagger W\mathcal{A}^2 = \bar{\phi} \mathcal{A}^2 W^\dagger U, \quad (61)$$

where  $\phi = (\alpha - i\beta)/|\alpha + i\beta|$ . As we have already seen, this equation implies that  $U = \phi W \bigoplus_{i=0}^{\kappa} Z_i$  with  $Z_i^2 = \mathbb{1}_{\mathcal{V}_i}$  for  $i \geq 1$ .

For such a  $U$ ,

$$\text{Tr}(\mathcal{A}^2 W^\dagger U) = \phi \text{Tr}(\mathcal{A}^2 Z) = \phi \sum_{i=1}^{\kappa} \omega_i^2 (n_i - 2\nu_i), \quad (62)$$

so the only way that  $\text{Tr}(\mathcal{A}^2 W^\dagger U)$  can be zero is if the vector  $(\omega_1^2, \dots, \omega_\kappa^2)$  is orthogonal to one of the possible vectors  $(n_1 - 2\nu_1, \dots, n_\kappa - 2\nu_\kappa)$ , i.e.  $(\omega_1^2, \dots, \omega_\kappa^2)$  must lie in the union of the  $2^{N-n_0-1}$  hyperplanes which are the orthogonal spaces of the vectors  $(n_1 - 2\nu_1, \dots, n_\kappa - 2\nu_\kappa)$ . Consequently, for a given orthogonal decomposition  $\mathcal{H} = \bigoplus_{i=0}^{\kappa} \mathcal{V}_i$ , the collection of all  $A$ 's for which  $\mathcal{A}^2 = AA^\dagger = \bigoplus_{i=1}^{\kappa} \omega_i^2 \mathbb{1}_{\mathcal{V}_i}$  and  $dF$  is surjective

at  $F(U) = 0$  [hence  $\{U : F(U) = 0\}$  is a codimension 2 submanifold of  $U(\mathcal{H})$ ] comprises an open dense subset of the  $A$ 's in  $\mathcal{B}(\mathcal{H})$  for which  $AA^\dagger = \bigoplus_{i=1}^{\kappa} \omega_i^2 \mathbb{1}_{\mathcal{V}_i}$ . It follows that the set of all  $A$  operators for which  $dF$  is surjective at  $F(U) = 0$  is open and dense in  $\mathcal{B}(\mathcal{H})$ .

Now, at a point  $U$  such that  $\text{Tr}(\mathcal{A}^2 W^\dagger U) = 0$ ,

$$\begin{aligned} \text{Hess}_{J_P, U}(\delta U) &= -\text{Tr}(\mathcal{A}^2 W^\dagger \delta U) W\mathcal{A}^2 + \text{Tr}(\delta U^\dagger W\mathcal{A}^2) U\mathcal{A}^2 W^\dagger U \quad (63a) \\ &= -\frac{1}{2} \left\{ (W\mathcal{A}^2 - U\mathcal{A}^2 W^\dagger U, \delta U) (W\mathcal{A}^2 - U\mathcal{A}^2 W^\dagger U) \right. \\ &\quad \left. + \langle iW\mathcal{A}^2 + iU\mathcal{A}^2 W^\dagger U, \delta U \rangle (iW\mathcal{A}^2 + iU\mathcal{A}^2 W^\dagger U) \right\} \quad (63b) \end{aligned}$$

so that the Hessian is rank 2 except where there exists  $(\alpha, \beta) \neq 0 \in \mathbb{R}^2$  such that  $\alpha(W\mathcal{A}^2 - U\mathcal{A}^2 W^\dagger U) = \beta(iW\mathcal{A}^2 + iU\mathcal{A}^2 W^\dagger U)$ . This is exactly the condition just considered for the surjectivity of  $dF$ , so that the Hessian is rank 2 if and only if  $dF$  is surjective. So, on the open dense set of  $A$ 's for which this happens, the maximum set is a nondegenerate (in the Morse-Bott sense), codimension 2 submanifold of  $U(\mathcal{H})$  (this is similar to a classical result on such distance functions [24, Thm. 6.6]). Since the other critical points also comprise nondegenerate submanifolds, we conclude that for these  $A$ 's,  $J_P$  is a Morse-Bott function.

Let  $P := U^\dagger W\mathcal{A}^2 - \mathcal{A}^2 W^\dagger U \in \mathfrak{u}(\mathcal{H})$  and  $Q := iU^\dagger W\mathcal{A}^2 + i\mathcal{A}^2 W^\dagger U \in \mathfrak{u}(\mathcal{H})$ . Then letting  $Y = U^\dagger \delta U$ , the Hessian eigenvalue equation at the global maximum may be written

$$\lambda Y = -\frac{1}{2} (\langle P, Y \rangle P + \langle Q, Y \rangle Q). \quad (64)$$

Let  $\gamma_\pm$  and  $v^\pm$  be the eigenvalues and eigenvectors of the real, symmetric, positive semidefinite Gram matrix

$$G = \begin{bmatrix} \langle P, P \rangle & \langle P, Q \rangle \\ \langle Q, P \rangle & \langle Q, Q \rangle \end{bmatrix}. \quad (65)$$

Then the eigenvalues of (64) are  $\lambda_\pm = -\gamma_\pm/2$  and the eigenvectors are  $Y^\pm = v_1^\pm P + v_2^\pm Q$ .

#### E. Examples

**Example 3** ( $\mathcal{A}^2$  is a projection). Suppose that  $\mathcal{A}^2$  is a projection, or equivalently,  $A$  is a partial isometry. Then  $\mathcal{H} = \mathcal{V}_0 \oplus \mathcal{V}_1$ ,  $\omega_0^2 = 0$ , and  $\omega_1^2 = 1$ . By Theorem 2, aside from the global maximum set,  $J_P$  admits exactly  $\lceil n_1/2 \rceil$  critical submanifolds  $C_{\nu_1}$  for  $\nu_1 = 0, \dots, \lceil n_1/2 \rceil - 1$ , each isomorphic to  $U(\mathcal{V}_0) \oplus [U(1) \times \text{Gr}_{\nu_1}(\mathcal{V}_1)]$  and having dimension  $\mathcal{N}_0 = 1 + n_0^2 + 2\nu_1(n_1 - \nu_1)$ . The critical submanifold  $C_{\nu_1}$  has critical value  $J_P = 4\nu_1(n_1 - \nu_1)$ , and the Hessian operator at a critical point  $U \in C_{\nu_1}$  can have (depending on  $\nu_1$  and  $n_0$ ) eigenvalues  $-2n_1, -2n_1 + 4\nu_1, -n_1 + 2\nu_1, 0, n_1 - 2\nu_1$ , and  $2n_1 - 4\nu_1$ . The negative Hessian eigenspace at  $U$  is

$$\begin{aligned} E_- &= E_{-2n_1} \oplus E_{-2n_1+4\nu_1} \oplus E_{-n_1+2\nu_1} \\ &= U \left[ \mathbb{R}Y^{(-2n_1)} \oplus \mathfrak{su}(\mathcal{V}_1^-) \oplus \frac{\mathfrak{u}(\mathcal{V}_0 \oplus \mathcal{V}_1^-)}{\mathfrak{u}(\mathcal{V}_0) \oplus \mathfrak{u}(\mathcal{V}_1^-)} \right] \end{aligned} \quad (66)$$

of dimension  $\mathcal{N}_- = (2n_0 + \nu_1)\nu_1$ , where

$$Y^{(-2n_1)} = \begin{cases} \frac{i}{\nu_1} \mathbb{1}_{\mathcal{V}_1^-} \oplus \frac{i}{n_1 - \nu_1} \mathbb{1}_{\mathcal{V}_1^+} & \nu_1 > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (67)$$

The positive Hessian eigenspace at  $U$  is

$$\begin{aligned} E_+ &= E_{2n_1-4\nu_1} \oplus E_{n_1-2\nu_1} \\ &= U \left[ \text{su}(\mathcal{V}_1^+) \oplus \frac{\mathfrak{u}(\mathcal{V}_0 \oplus \mathcal{V}_1^+)}{\mathfrak{u}(\mathcal{V}_0) \oplus \mathfrak{u}(\mathcal{V}_1^+)} \right] \end{aligned} \quad (68)$$

of dimension  $\mathcal{N}_+ = (2n_0 + n_1 - \nu_1)(n_1 - \nu_1) - 1$ . The full set of critical values for this problem is  $\{0, 4(n_1 - 1), 8(n_1 - 2), \dots, 4[(n_1^2 - 1)/4], n_1^2\}$ , where the final value  $n_1^2 = \|A\|^4$  is the globally maximal value. In the particular case that  $\mathcal{A}^2$  is fully degenerate (e.g.,  $A = \mathbb{1}$ ), it is found that the maximum set of  $J_P$  is a nondegenerate submanifold if and only if  $N = \dim(\mathcal{H})$  is odd. However, when  $N$  is even, arbitrarily small perturbations of  $A$  about  $\mathbb{1}$  are sufficient to obtain a Morse-Bott function.

**Example 4** ( $\omega_i^2 = 2^i$ ). Suppose that  $\mathcal{A}^2$  is nondegenerate and nonsingular, so that  $\mathcal{H} = \bigoplus_{i=1}^N \mathcal{V}_i$ ,  $n_i = 1$  for all  $i = 1, \dots, N$ , and  $\omega_1^2 < \dots < \omega_N^2$ . Suppose further that  $\omega_i^2 = 2^i$  for  $i = 1, \dots, N$ . Then  $\omega_N^2 > \sum_{i=1}^{N-1} \omega_i^2$ , so  $\sum_{i=1}^N \omega_i^2 (n_i - 2\nu_i) > 0$  if and only if  $\nu_N = 0$ . Therefore, by Theorem 2, the critical set of  $J_P$  comprises 1 global maximal set and exactly  $2^{N-1}$  critical submanifolds  $C_{\{\nu_i\}}$ , each isomorphic to  $U(1)$  and equal to a global phase orbit  $\{\phi U : \phi \in U(1)\}$  of a single unitary operator. For the critical submanifold identified by indices  $(\nu_1, \dots, \nu_N)$ , the critical value is

$$\begin{aligned} J_P &= 4 \left( \sum_{i=1}^{N-1} 2^i \nu_i \right) \left( \sum_{i=1}^N 2^i (n_i - \nu_i) \right) \\ &= 4 \left( \sum_{i=1}^{N-1} 2^i \nu_i \right) \left[ 2^{N+1} - 2 - \sum_{i=1}^{N-1} 2^i \nu_i \right], \end{aligned} \quad (69)$$

the negative Hessian eigenbundle has dimension  $\mathcal{N}_- = \sum_{i=1}^{N-1} (2i - 1)\nu_i$ , and the positive eigenbundle has dimension  $\mathcal{N}_+ = N^2 - 1 - \sum_{i=1}^{N-1} (2i - 1)\nu_i$ . The full set of critical values for this problem is  $\{16(2^N - 2), 32(2^N - 3), 48(2^N - 4), \dots, 2^{N+1}(2^{N+1} - 4), (2^{N+1} - 2)^2\}$ . The last of these critical values corresponds to the global maximum.

#### IV. LANDSCAPES BASED ON INTRINSIC DISTANCE

The kinematic landscapes  $J_F$  and  $J_P$  considered above are based on the Euclidean (or norm) distance on  $U(\mathcal{H})$  and  $\text{PU}(\mathcal{H})$ , respectively. We now describe two additional distance measures based on the intrinsic distance between operators in  $U(\mathcal{H})$  and  $\text{PU}(\mathcal{H})$  under the Riemannian metric induced by the real Hilbert-Schmidt inner product on  $\mathcal{B}(\mathcal{H})$ .

The first of these distance measures is quite simple to define. Since the chosen Riemannian metric is bi-invariant on  $U(\mathcal{H})$ , any geodesic starting at  $U \in U(\mathcal{H})$  is of the form  $\gamma(s) = Ue^{As}$  for some  $A \in \mathfrak{u}(\mathcal{H})$ . To find a geodesic joining  $U$  to some target  $W \in U(\mathcal{H})$ , let  $Ue^A = \gamma(1) = W$ , so that  $e^A = U^\dagger W$  and  $A = \log(U^\dagger W)$ . This matrix logarithm is not uniquely defined, but the length of the geodesic  $\gamma$  defined on the interval  $[0, 1]$  is given by  $L[\gamma] = \int_0^1 \|\gamma'(s)\| ds = \|A\|$ . The minimum such length is obtained by taking  $A = \log(U^\dagger W)$  from the principal branch of the logarithm so that

all eigenvalues lie in  $(-i\pi, i\pi]$ . We then define the landscape as

$$J_G(U) := \frac{1}{2} \|\log(U^\dagger W)\|^2. \quad (70)$$

Then the gradient of  $J_G$  is given by (see Appendix B)

$$\text{grad } J_G(U) = -U \log(U^\dagger W). \quad (71)$$

As most numerical matrix logarithm routines (e.g., the `logm` function in MATLAB) compute the principal branch, they provide a ready means to obtain both the landscape value and the gradient. Since the norm of  $\text{grad } J_G$  is the distance to the target, this vector field is only zero at the target, i.e. the global minimum of the landscape. Hence, there are no traps or saddles. The gradient field has the property that it is discontinuous and multiply defined at the cut loci of  $U(\mathcal{H})$  (where the spectrum of  $U^\dagger W$  contains  $-1$ ), but this is not a problem for an optimal control algorithm since the matrix logarithm routine will have to choose one from among the possible solutions, all of which describe minimal geodesics to the target that are equally satisfactory.

A phase-invariant version of  $J_G$  may be constructed analogously by considering minimal geodesics on the projective unitary group  $\text{PU}(\mathcal{H}) \simeq U(\mathcal{H})/U(1)$ , or equivalently by defining  $J_{GP}(U) := \min_{\phi \in U(1)} \frac{1}{2} \|\log(\phi U^\dagger W)\|^2$  on  $U(\mathcal{H})$ . It may be shown that

$$J_{GP} = \min_{k \in \mathbb{Z}_N} \frac{1}{2} \left\| \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right) \right\|^2 \quad (72a)$$

$$\begin{aligned} \text{grad } J_{GP}(U) &= -U \left\{ \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right) \right. \\ &\quad \left. - \text{Tr} \left[ \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right) \right] \frac{\mathbb{1}}{N} \right\}, \end{aligned} \quad (72b)$$

where  $k$  in (72b) is the minimizer from (72a). With this minimizing  $k$ , the trace in (72b) will be zero, so that

$$\text{grad } J_{GP}(U) = -U \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right). \quad (73)$$

As with  $J_G$ , the norm of  $\text{grad } J_{GP}$  is the distance to the target, and this vector field is only zero at the target, i.e. the global minimum of the landscape. Hence, there are no traps or saddles. One downside to this landscape is that it appears that all  $N$  possible values of  $k$  must be tried in order to find the minimizer of (72a). This behavior has a topological interpretation on  $\text{PU}(\mathcal{H})$ . Since the fundamental group of  $\text{PU}(\mathcal{H})$  is  $\pi_1(\text{PU}(\mathcal{H})) \cong \mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ , there are exactly  $N$  homotopy classes of paths connecting  $\pi(U)$  to the target  $\pi(W)$ . Within each of these classes is a unique minimal geodesic, and these  $N$  minimal geodesics are identified by the vectors

$$\begin{aligned} &U \left\{ \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right) \right. \\ &\quad \left. - \frac{1}{N} \text{Tr} \left[ \log \left( e^{\frac{2\pi ik}{N}} \det(U^\dagger W)^{-\frac{1}{N}} U^\dagger W \right) \right] \mathbb{1} \right\} \end{aligned} \quad (74)$$

indexed by  $k$ .

A distance metric based on intrinsic distance could in principle be applied to the case where only some of the

states are important, analogous to  $J_F$  and  $J_P$  where  $A$  is rank deficient (e.g. where  $A$  is a projector). This is equivalent to computing the geodesic distance between points on the Stiefel manifold  $V_{\mathcal{H}/\mathcal{V}_0}(\mathcal{H}) \simeq \text{U}(\mathcal{H})/(\mathbb{1}_{\mathcal{V}_0} \oplus \text{U}(\mathcal{H}/\mathcal{V}_0))$  or on its projective cousin  $V_{\mathcal{H}/\mathcal{V}_0}(\mathcal{H})/\text{U}(1)$ . However, the two-point geodesics on these spaces are non-trivial to compute. The calculation requires solution of a boundary value problem or an optimization problem to find each minimal geodesic. For that reason, these intrinsic distance metrics may not be practical for this scenario.

## V. DYNAMICAL CRITICAL POINT ANALYSIS

Now that we have elucidated the structure of the critical sets of the kinematic landscapes  $J_F$  and  $J_P$ , we return to the problem of characterizing the critical set of the dynamical landscapes  $\tilde{J} = J \circ U_T$ . Let  $\mathcal{M} \subset \text{U}(\mathcal{H})$  be one of the critical submanifolds identified in the previous sections. It can be proved (see Appendix A) that  $U_T : \mathbb{K} \rightarrow \text{U}(\mathcal{H})$  is  $C^\infty$  (i.e., infinitely Fréchet differentiable). In addition, since  $\text{U}(\mathcal{H})$  is finite-dimensional, if  $U_T(\mathcal{E}) \in \mathcal{M}$  then  $(d_{\mathcal{E}}U_T)^{-1}(\text{T}_{U_T(\mathcal{E})}\mathcal{M})$  has finite codimension, so is closed and has a closed complement (i.e., it “splits”). Therefore, away from singular points of  $U_T$  (i.e., those  $\mathcal{E} \in \mathbb{K}$  such that  $d_{\mathcal{E}}U_T$  is rank-deficient),  $U_T$  is transversal to  $\mathcal{M}$  and by the transversal mapping theorem [25],  $U_T^{-1}(\mathcal{M})$  is a Hilbert submanifold of  $\mathbb{K}$ ,  $\text{T}_{\mathcal{E}}(U_T^{-1}(\mathcal{M})) = (d_{\mathcal{E}}U_T)^{-1}(\text{T}_{U_T(\mathcal{E})}\mathcal{M})$ , and  $\text{codim}(U_T^{-1}(\mathcal{M})) = \text{codim}(\mathcal{M})$ .

Let  $\mathcal{E} \in \mathbb{K}$  be a regular critical point of  $\tilde{J}$ , i.e. such that  $\text{grad}\tilde{J}(\mathcal{E}) = 0$  and  $d_{\mathcal{E}}U_T$  is full rank. It may be seen that at such a point, the Hessian of  $\tilde{J}$  is given by  $\text{Hess}_{\tilde{J},\mathcal{E}} = (d_{\mathcal{E}}U_T)^* \circ \text{Hess}_{J,U_T(\mathcal{E})} \circ (d_{\mathcal{E}}U_T)$ . Let  $\mathcal{A}_{\mathcal{E}}$  be the linear operator on  $\text{T}_{U_T(\mathcal{E})}\text{U}(\mathcal{H})$  given by

$$\mathcal{A}_{\mathcal{E}} = (d_{\mathcal{E}}U_T \circ (d_{\mathcal{E}}U_T)^*)^{\frac{1}{2}} \circ \text{Hess}_{J,U_T(\mathcal{E})} \circ (d_{\mathcal{E}}U_T \circ (d_{\mathcal{E}}U_T)^*)^{\frac{1}{2}}. \quad (75)$$

Since  $d_{\mathcal{E}}U_T$  is assumed to have full rank, we may invoke Sylvester’s law of inertia [26] to conclude that  $\mathcal{A}_{\mathcal{E}}$  and  $\text{Hess}_{J,U_T(\mathcal{E})}$  have the same numbers of positive, negative, and zero eigenvalues. Let  $\{(\eta_j, Q_j)\}$  for  $j = 1, \dots, N^2$  be the eigenvalues and eigenvectors of  $\mathcal{A}_{\mathcal{E}}$ , and let  $Z_j = (d_{\mathcal{E}}U_T)^* \circ (d_{\mathcal{E}}U_T \circ (d_{\mathcal{E}}U_T)^*)^{-\frac{1}{2}} Q_j$ . Then

$$\begin{aligned} \text{Hess}_{\tilde{J},\mathcal{E}} Z_j &= (d_{\mathcal{E}}U_T)^* \circ \text{Hess}_{J,U_T(\mathcal{E})} \circ (d_{\mathcal{E}}U_T \circ (d_{\mathcal{E}}U_T)^*)^{\frac{1}{2}} Q_j \\ &= \eta_j (d_{\mathcal{E}}U_T)^* \circ (d_{\mathcal{E}}U_T \circ (d_{\mathcal{E}}U_T)^*)^{-\frac{1}{2}} Q_j \\ &= \eta_j Z_j \end{aligned} \quad (76)$$

so that  $\{(\eta_j, Z_j)\}$  for  $j = 1, \dots, N^2$  are eigenvalues and eigenvectors of  $\text{Hess}_{\tilde{J},\mathcal{E}}$ . Because  $\text{Hess}_{\tilde{J},\mathcal{E}}$  is self-adjoint, any other eigenvector  $Z$  must be orthogonal to the  $\{Z_j\}$ . Also, note that since the  $\{Q_j\}$  span  $\text{T}_{U_T(\mathcal{E})}\text{U}(\mathcal{H})$ , the  $\{Z_j\}$  span  $\text{Range}((d_{\mathcal{E}}U_T)^*)$ . Then, for any  $X \in \text{T}_{U_T(\mathcal{E})}\text{U}(\mathcal{H})$ ,  $0 = \langle Z, (d_{\mathcal{E}}U_T)^*(X) \rangle = \langle d_{\mathcal{E}}U_T(Z), X \rangle$ , so that  $d_{\mathcal{E}}U_T(Z) = 0$  and therefore  $\text{Hess}_{\tilde{J},\mathcal{E}} Z = 0$ . Thus,  $\text{Hess}_{\tilde{J},\mathcal{E}}$  has infinitely many eigenvalues;  $N^2$  of them are identical to the eigenvalues of  $\mathcal{A}_{\mathcal{E}}$ , and the remaining infinite number of eigenvalues are all zero. Since  $J$  has no local traps, we can conclude that  $\tilde{J}$

has no local traps among the regular critical points. From the transversal mapping theorem we find that  $\text{T}_{\mathcal{E}}(U_T^{-1}(\mathcal{M})) = (d_{\mathcal{E}}U_T)^{-1}(\text{T}_{U_T(\mathcal{E})}\mathcal{M})$ , implying that for any  $f \in \mathbb{K}$ , we have

$$\begin{aligned} \text{Hess}_{\tilde{J},\mathcal{E}}(f) &= (d_{\mathcal{E}}U_T)^* \circ \text{Hess}_{J,U_T(\mathcal{E})} \circ (d_{\mathcal{E}}U_T)(f) = 0 \\ \iff d_{\mathcal{E}}U_T(f) &\in \ker \text{Hess}_{J,U_T(\mathcal{E})} \end{aligned} \quad (77a)$$

$$\iff d_{\mathcal{E}}U_T(f) \in \text{T}_{U_T(\mathcal{E})}\mathcal{M} \quad (77b)$$

$$\iff f \in (d_{\mathcal{E}}U_T)^{-1}(\text{T}_{U_T(\mathcal{E})}\mathcal{M}) \quad (77c)$$

$$\iff f \in \text{T}_{\mathcal{E}}(U_T^{-1}(\mathcal{M})). \quad (77d)$$

Hence, the null space of  $\text{Hess}_{\tilde{J},\mathcal{E}}$  is identical to  $\text{T}_{\mathcal{E}}(U_T^{-1}(\mathcal{M}))$ , the tangent space to the critical submanifold.

In the case where the Hamiltonian takes the dipole form  $H(t) = H_0 - \mathcal{E}(t)\mu$  for any  $\mathcal{E} \in L^2(\mathbb{R}_+; \mathbb{R})$ , the Fréchet derivative of  $U_T$  is given by [27]

$$d_{\mathcal{E}}U_T(\delta\mathcal{E}) = \frac{i}{\hbar} U_T(\mathcal{E}) \int_0^T U_t^\dagger(\mathcal{E}) \mu U_t(\mathcal{E}) \delta\mathcal{E}(t) dt. \quad (78)$$

Then the adjoint operator of the derivative is

$$d_{\mathcal{E}}U_T^*(A)(t) = -\Im \text{Tr} (A^\dagger U_T(\mathcal{E}) U_t^\dagger(\mathcal{E}) \mu U_t(\mathcal{E})) \quad (79)$$

for any  $A \in \text{T}_{U_T(\mathcal{E})}\text{U}(\mathcal{H})$ , and the operator norm of this adjoint is uniformly bounded by  $\|d_{\mathcal{E}}U_T^*\| \leq \sqrt{T}\|\mu\|$ . For any smooth “kinematic” function  $g : \text{U}(\mathcal{H}) \rightarrow \mathbb{R}$ , let  $\tilde{g} = g \circ U_T$  be the corresponding “dynamical” function on  $L^2(\mathbb{R}_+; \mathbb{R})$ . Then  $\text{grad}\tilde{g}(\mathcal{E}) = d_{\mathcal{E}}U_T^*(\text{grad}g(U_T(\mathcal{E})))$  and

$$\|\text{grad}\tilde{g}(\mathcal{E})\| \leq \|d_{\mathcal{E}}U_T^*\| \|\text{grad}g(U_T(\mathcal{E}))\| \quad (80a)$$

$$\leq \sqrt{T}\|\mu\| \|\text{grad}g(U_T(\mathcal{E}))\|. \quad (80b)$$

Since  $g$  is smooth,  $\|\text{grad}g\|$  is continuous over  $\text{U}(\mathcal{H})$ , so that since  $\text{U}(\mathcal{H})$  is compact,  $\|\text{grad}g\|$  is uniformly bounded. Therefore,  $\|\text{grad}\tilde{g}\|$  is uniformly bounded over  $L^2(\mathbb{R}_+; \mathbb{R})$ . For any dynamical quantum control landscape constructed in this way, in particular the landscapes considered in the present paper, the slope of the landscape (i.e. the speed of the gradient flow) is uniformly bounded by some constant.

Taken together, these results show that, even though the control space  $\mathbb{K}$  is unbounded and infinite-dimensional and one might naively expect anything to happen, the landscapes under consideration are well-behaved, exhibiting gradient flows which do not get trapped (at least away from singular points) and which do not speed out of control.

## VI. SUMMARY

This work presented an expanded analysis of landscapes  $J_F$  and  $J_P$ , which are based on the Euclidean distances between unitary operators in  $\text{U}(\mathcal{H})$  and  $\text{PU}(\mathcal{H})$ , respectively. The expansion appears in several ways. First, additional freedom has been allowed in the landscape functions themselves, by admitting  $A$  matrices that are rank-deficient. Landscapes based on these rank-deficient  $A$  matrices measure the distance between unitary operators by their action on a subspace of the full state space. This can be the desired objective for designing a quantum information processor, for example, where only this subspace of the state space is to be used for the quantum

register. This additional freedom in defining the landscape is consistent with the principal finding of earlier work on landscapes of this form: they have no suboptimal minima (i.e., “traps”) that could impede a deterministic optimal control algorithm (such as gradient descent) from reaching the global minimum.

In addition to broadening the families of landscapes for consideration, we have provided more detail on the structure of the critical sets and the behavior of the landscape functions at these critical sets. The critical sets were shown to generally be disjoint unions of critical submanifolds and we have described the structure of these submanifolds, as products of Grassmann manifolds and unitary groups. Furthermore, we have shown that these critical submanifolds are generally nondegenerate in the Morse-Bott sense, so that the kinematic landscapes are generally Morse-Bott functions.

These results were related back to the corresponding dynamical landscapes through the control-to-propagator map  $U_T$ , implicitly defined by the Schrödinger equation, that takes a control function as input and returns the final time unitary evolution operator. This map was shown to be infinitely Fréchet differentiable, leading to the conclusion that, away from the singular points of  $U_T$ , the level sets and critical sets of the dynamical landscapes are  $C^\infty$  smooth, finite codimension submanifolds of the infinite-dimensional control space  $\mathbb{K} = L^2(\mathbb{R}_+; \mathbb{R})$ . Also, the number of positive and negative Hessian eigenvalues (and therefore the characterization as a minimum, maximum, or saddle) was shown to be identical for a kinematic critical point and a regular point of  $U_T$  that maps to it. This behavior implies that no traps exist in the dynamical landscape among the set of regular points of  $U_T$ . Furthermore, Morse-Bott nondegeneracy of the critical set is also preserved away from singular points of  $U_T$ , which can be important for certain numerical landscape exploration methods such as second order D-MORPH [14].

Finally, two additional landscapes were introduced that are based on the intrinsic or geodesic distance between operators in  $U(\mathcal{H})$  and  $PU(\mathcal{H})$ , respectively, rather than Euclidean distance. These kinematic landscapes have the desirable property of having no critical points except for the global minimum at the target. These landscapes may allow for more efficient performance of optimal control algorithms over  $J_F$  and  $J_P$ , since the latter have many saddle points where the gradient is zero.

## APPENDIX A

### DIFFERENTIABILITY OF $U(T, 0)$ WITH RESPECT TO THE CONTROL

Let  $\mathcal{M}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$  denote the space of Hermitian operators endowed with the real Hilbert-Schmidt inner product  $\langle A, B \rangle_{\text{HS}} = \Re \text{Tr}(A^\dagger B)$ , and let  $\mathbb{H}(\mathcal{H}) = L^2(\mathbb{R}_+; \mathcal{M}(\mathcal{H})) = \mathcal{M}(\mathcal{H}) \otimes L^2(\mathbb{R}_+; \mathbb{R})$  denote the space of all square-integrable time-dependent Hamiltonians on  $\mathcal{H}$  with inner product

$$\langle H_1, H_2 \rangle_{L^2} = \int_0^\infty \Re \text{Tr}[H_1(t)H_2(t)] dt. \quad (81)$$

Let  $Z_T : \mathbb{H}(\mathcal{H}) \rightarrow U(\mathcal{H})$  be the map, defined implicitly through the Schrödinger equation, that takes a time-dependent Hamiltonian  $H(\cdot) \in \mathbb{H}(\mathcal{H})$  and produces the corresponding unitary time-evolution operator at time  $T$ :  $U(T, 0) \in U(\mathcal{H})$ . This map is well-defined over the entire domain because of the absolute convergence of the Dyson series over  $\mathbb{H}(\mathcal{H})$ :

$$\begin{aligned} Z_T(H) &= \mathbb{1} + \left(-\frac{i}{\hbar}\right) \int_0^T dt_1 H(t_1) \\ &\quad + \left(-\frac{i}{\hbar}\right)^2 \int_0^T dt_1 H(t_1) \int_0^{t_1} dt_2 H(t_2) + \dots \end{aligned} \quad (82)$$

In this appendix, we will prove that  $Z_T$  is infinitely Fréchet differentiable over  $\mathbb{H}(\mathcal{H})$ . A corollary is that the map  $U_T : \mathbb{K} \rightarrow U(\mathcal{H})$  defined in the body of the paper is infinitely Fréchet differentiable over all of  $\mathbb{K} = L^2(\mathbb{R}_+; \mathbb{R})$ .

**Lemma 1.** *If  $f : [a, b] \rightarrow \mathbb{R}$  is integrable, then*

$$\begin{aligned} \int_a^b dt_1 f(t_1) \int_a^{t_1} dt_2 f(t_2) \cdots \int_a^{t_{n-1}} dt_n f(t_n) \\ = \frac{1}{n!} \left( \int_a^b f(t) dt \right)^n. \end{aligned} \quad (83)$$

*Proof:* Note first that (83) holds trivially for  $n = 1$ . Suppose that it holds for  $n = m$ . Then

$$\begin{aligned} \int_a^b dt_1 f(t_1) \int_a^{t_1} dt_2 f(t_2) \cdots \int_a^{t_m} dt_{m+1} f(t_{m+1}) \\ = \frac{1}{m!} \int_a^b f(t_1) \left( \int_a^{t_1} f(t) dt \right)^m dt_1 \end{aligned} \quad (84a)$$

$$= \frac{1}{m!} \int_a^b \frac{d}{dt_1} \left[ \frac{1}{m+1} \left( \int_a^{t_1} f(t) dt \right)^{m+1} \right] dt_1 \quad (84b)$$

$$= \frac{1}{(m+1)!} \left( \int_a^b f(t) dt \right)^{m+1} \quad (84c)$$

and the lemma follows for arbitrary  $n \in \mathbb{N}$  by induction. ■

**Definition 1.** *For integrable operator-valued functions  $A_i : [0, T] \rightarrow \mathcal{B}(\mathcal{H})$  and for integrable real-valued functions  $a_i : [0, T] \rightarrow \mathbb{R}$ , we will use the following short-hand notation for the Dyson-esque terms*

$$\begin{aligned} \Upsilon_T[A_1, A_2, \dots, A_n] \\ := \int_0^T dt_1 A_1(t_1) \int_0^{t_1} dt_2 A_2(t_2) \cdots \int_0^{t_{n-1}} dt_n A_n(t_n) \end{aligned} \quad (85a)$$

$$\begin{aligned} \Upsilon_T[a_1, a_2, \dots, a_n] \\ := \int_0^T dt_1 a_1(t_1) \int_0^{t_1} dt_2 a_2(t_2) \cdots \int_0^{t_{n-1}} dt_n a_n(t_n). \end{aligned} \quad (85b)$$

**Lemma 2.** *If  $g_1, \dots, g_m \in L^2(\mathbb{R}_+; \mathbb{R}_+)$  are non-negative square-integrable functions on  $[0, \infty)$ , then*

$$\Upsilon_T[g_1, g_2, \dots, g_m] \leq T^{m/2} \|g_1\|_{L^2} \cdots \|g_m\|_{L^2} \quad (86)$$

*Proof:* Since the  $g_i$ 's are non-negative functions, we get the inequalities

$$\Upsilon_T[g_1, g_2, \dots, g_m] \leq \int_0^T dt_1 g_1(t_1) \int_0^T dt_2 g_2(t_2) \cdots \int_0^T dt_m g_m(t_m) \quad (87a)$$

$$\leq T^{m/2} \prod_{i=1}^m \left( \int_0^T dt_i g_i^2(t_i) \right)^{\frac{1}{2}} \quad (87b)$$

$$\leq T^{m/2} \|g_1\|_{L^2} \cdots \|g_m\|_{L^2} \quad (87c)$$

by extension of the integrals out to the interval  $[0, T]$ , followed by application of the Cauchy-Schwarz inequality, and finally extension out to  $[0, \infty)$ . ■

**Lemma 3.** *If  $f, g_1, \dots, g_m \in L^2(\mathbb{R}_+; \mathbb{R}_+)$  are non-negative square-integrable functions on  $[0, \infty)$ , then*

$$\begin{aligned} & \Upsilon_T \underbrace{[f, \dots, f]}_{\beta_0 \text{ terms}} \underbrace{[g_1, f, \dots, f]}_{\beta_1 \text{ terms}} \underbrace{[g_2, f, \dots, f]}_{\beta_2 \text{ terms}} \cdots \underbrace{[g_m, f, \dots, f]}_{\beta_m \text{ terms}} \\ & \leq \frac{T^{\frac{1}{2} \sum \beta_i}}{\prod_{i=0}^m \beta_i!} \|f\|_{L^2}^{\sum \beta_i} \Upsilon_T[g_1, g_2, \dots, g_m] \end{aligned} \quad (88)$$

*Proof:* Let  $\sigma_i = \beta_0 + \dots + \beta_i + i + 1$  for  $i = 0, \dots, m$ . Then, using Fubini's theorem, we may rearrange the order of integration as follows:

$$\begin{aligned} & \Upsilon_T \underbrace{[f, \dots, f]}_{\beta_0 \text{ terms}} \underbrace{[g_1, f, \dots, f]}_{\beta_1 \text{ terms}} \underbrace{[g_2, f, \dots, f]}_{\beta_2 \text{ terms}} \cdots \underbrace{[g_m, f, \dots, f]}_{\beta_m \text{ terms}} \\ & = \int_0^T dt_{\sigma_0} g_1(t_{\sigma_0}) \int_0^{t_{\sigma_0}} dt_{\sigma_1} g_2(t_{\sigma_1}) \cdots \\ & \quad \int_0^{t_{\sigma_{m-2}}} dt_{\sigma_{m-1}} g_m(t_{\sigma_{m-1}}) \int_{t_{\sigma_0}}^T dt_1 f(t_1) \int_{t_{\sigma_0}}^{t_1} dt_2 f(t_2) \\ & \quad \cdots \int_{t_{\sigma_0}}^{t_{\sigma_0-2}} dt_{\sigma_0-1} f(t_{\sigma_0-1}) \int_{t_{\sigma_1}}^{t_{\sigma_0}} dt_{\sigma_0+1} f(t_{\sigma_0+1}) \\ & \quad \int_{t_{\sigma_1}}^{t_{\sigma_0+1}} dt_{\sigma_0+2} f(t_{\sigma_0+2}) \cdots \int_{t_{\sigma_1}}^{t_{\sigma_1-2}} dt_{\sigma_1-1} f(t_{\sigma_1-1}) \\ & \quad \cdots \int_{t_{\sigma_{m-1}}}^{t_{\sigma_{m-2}}} dt_{\sigma_{m-2}+1} f(t_{\sigma_{m-2}+1}) \int_{t_{\sigma_{m-1}}}^{t_{\sigma_{m-2}+1}} dt_{\sigma_{m-2}+2} f(t_{\sigma_{m-2}+2}) \\ & \quad \cdots \int_{t_{\sigma_{m-1}}}^{t_{\sigma_{m-1}-2}} dt_{\sigma_{m-1}-1} f(t_{\sigma_{m-1}-1}) \int_0^{t_{\sigma_{m-1}}} dt_{\sigma_{m-1}+1} f(t_{\sigma_{m-1}+1}) \\ & \quad \cdots \int_0^{t_{\sigma_{m-2}}} dt_{\sigma_{m-1}} f(t_{\sigma_{m-1}}) \quad (89a) \\ & = \frac{1}{\prod_{i=0}^m \beta_i!} \int_0^T dt_{\sigma_0} g_1(t_{\sigma_0}) \int_0^{t_{\sigma_0}} dt_{\sigma_1} g_2(t_{\sigma_1}) \cdots \\ & \quad \cdots \int_0^{t_{\sigma_{m-2}}} dt_{\sigma_{m-1}} g_m(t_{\sigma_{m-1}}) \left( \int_{t_{\sigma_0}}^T f(t) dt \right)^{\beta_0} \times \\ & \quad \times \left( \int_{t_{\sigma_1}}^{t_{\sigma_0}} f(t) dt \right)^{\beta_1} \cdots \left( \int_0^{t_{\sigma_{m-1}}} f(t) dt \right)^{\beta_m} \end{aligned} \quad (89b)$$

where the last step follows from Lemma 1. Then, since  $f$  is

a non-negative function, we get the inequality

$$\begin{aligned} & \Upsilon_T \underbrace{[f, \dots, f]}_{\beta_0 \text{ terms}} \underbrace{[g_1, f, \dots, f]}_{\beta_1 \text{ terms}} \underbrace{[g_2, f, \dots, f]}_{\beta_2 \text{ terms}} \cdots \underbrace{[g_m, f, \dots, f]}_{\beta_m \text{ terms}} \\ & \leq \frac{1}{\prod_{i=0}^m \beta_i!} \left( \int_0^T f(t) dt \right)^{\sum_{i=0}^m \beta_i} \times \\ & \quad \times \int_0^T dt_1 g_1(t_1) \int_0^{t_1} dt_2 g_2(t_2) \cdots \int_0^{t_{m-1}} dt_m g_m(t_m) \end{aligned} \quad (90a)$$

$$\leq \frac{T^{\frac{1}{2} \sum \beta_i}}{\prod_{i=0}^m \beta_i!} \|f\|_{L^2}^{\sum \beta_i} \Upsilon_T[g_1, g_2, \dots, g_m] \quad (90b)$$

by first extending the  $f$  integrals to the interval  $[0, T]$ . and then invoking the Cauchy-Schwarz inequality. ■

**Definition 2.** *For  $m = 0, 1, 2, \dots$ , let  $\mathcal{B}^m(\text{T}\mathbb{H}(\mathcal{H}); \mathcal{B}(\mathcal{H}))$  denote the space of bounded  $m$ -multilinear operators from  $(\text{T}\mathbb{H}(\mathcal{H}))^m = \text{T}\mathbb{H}(\mathcal{H}) \times \text{T}\mathbb{H}(\mathcal{H}) \times \dots \times \text{T}\mathbb{H}(\mathcal{H})$  to  $\mathcal{B}(\mathcal{H})$ , with the norm*

$$\|A\| = \sup_{\{\|\delta H_j\| \neq 0\}} \frac{\|A(\delta H_1, \dots, \delta H_m)\|}{\|\delta H_1\| \cdots \|\delta H_m\|} \quad (91)$$

for each  $A \in \mathcal{B}^m(\text{T}\mathbb{H}(\mathcal{H}); \mathcal{B}(\mathcal{H}))$ . Then let  $\varphi_{T,m} : \mathbb{H} \rightarrow \mathcal{B}^m(\text{T}\mathbb{H}(\mathcal{H}); \mathcal{B}(\mathcal{H}))$  be defined by

$$\begin{aligned} & \varphi_{T,m}(H)(\delta H_1, \delta H_2, \dots, \delta H_m) \\ & := \sum_{n=m}^{\infty} \sum_{a_0 + \dots + a_m = n-m} \sum_{\pi \in S_m} \left( -\frac{i}{\hbar} \right)^n \times \\ & \quad \times \Upsilon_T \underbrace{[H, \dots, H]}_{a_0 \text{ terms}} \underbrace{[\delta H_{\pi(1)}, H, \dots, H]}_{a_1 \text{ terms}} \underbrace{[\delta H_{\pi(2)}, H, \dots]}_{a_2 \text{ terms}} \cdots \underbrace{[\delta H_{\pi(m)}, H, \dots, H]}_{a_m \text{ terms}} \end{aligned} \quad (92)$$

where  $S_m$  denotes the symmetric group on  $m$  elements (i.e., the group of permutations of  $m$  elements). For  $m, q = 0, 1, 2, \dots$ , let  $\Psi_{T,m,q} : \mathbb{H} \oplus \text{T}\mathbb{H} \rightarrow \mathcal{B}^m(\text{T}\mathbb{H}(\mathcal{H}); \mathcal{B}(\mathcal{H}))$  be defined by

$$\begin{aligned} & \Psi_{T,m,q}(H, \delta H)(\delta H_1, \delta H_2, \dots, \delta H_m) \\ & := \sum_{n=m+q}^{\infty} \sum_{\substack{a_0 + \dots + a_m = n-q \\ = n-m-q}} \sum_{b_0 + \dots + b_m = q} \sum_{\pi \in S_m} \left( -\frac{i}{\hbar} \right)^n \times \\ & \quad \times \Upsilon_T \underbrace{[H, \dots, H]}_{a_0 \text{ terms}} \underbrace{[A_1, H, \dots, H]}_{a_1 \text{ terms}} \underbrace{[A_2, H, \dots]}_{a_2 \text{ terms}} \cdots \underbrace{[H, A_{m+q}, H, \dots, H]}_{a_{m+q} \text{ terms}} \end{aligned} \quad (93)$$

where

$$\begin{aligned} & \{A_1, A_2, \dots, A_{m+q}\} \\ & = \left\{ \underbrace{\delta H, \dots, \delta H}_{b_0 \text{ terms}}, \underbrace{\delta H_{\pi(1)}, \delta H, \dots, \delta H}_{b_1 \text{ terms}}, \delta H_{\pi(2)}, \delta H, \dots \right. \\ & \quad \left. \dots, \delta H, \delta H_{\pi(m)}, \underbrace{\delta H, \dots, \delta H}_{b_m \text{ terms}} \right\} \end{aligned} \quad (94)$$

**Lemma 4.**  *$\varphi_{T,m}$  and  $\Psi_{T,m,q}$  are well-defined since their defining sums converge absolutely, and for each  $H \in \mathbb{H}$  and*

$\delta H \in \mathbb{T}\mathbb{H}$ ,  $\varphi_{T,m}(H)$  and  $\Psi_{T,m,q}(H)(\delta H)$  are bounded  $m$ -multilinear operators.

*Proof:* Let  $f(t) = \|H(t)\|$ ,  $g_i(t) = \|A_i(t)\|$ ,  $h_j(t) = \|\delta H_j(t)\|$ , and  $h(t) = \|\delta H(t)\|$ . Then

$$\begin{aligned} & \|\Upsilon_T[\underbrace{H, \dots, H}_{a_0 \text{ terms}}, A_1, H, \dots, H, A_{m+q}, \underbrace{H, \dots, H}_{a_{m+q} \text{ terms}}]\| \\ & \leq \Upsilon_T[\underbrace{f, \dots, f}_{a_0 \text{ terms}}, \underbrace{g_1, \dots, g_1}_{a_1 \text{ terms}}, \underbrace{g_2, f, \dots, f}_{a_{m+q} \text{ terms}}, \underbrace{f, \dots, f}_{a_{m+q} \text{ terms}}] \end{aligned} \quad (95a)$$

$$\leq \frac{T^{\frac{n-m-q}{2}}}{\prod_{i=0}^{m+q} a_i!} \|f\|_{L^2}^{n-m-q} \Upsilon_T[g_1, g_2, \dots, g_{m+q}] \quad (95b)$$

$$\begin{aligned} & = \frac{T^{\frac{n-m-q}{2}}}{\prod_{i=0}^{m+q} a_i!} \|H\|_{L^2}^{n-m-q} \times \\ & \quad \times \Upsilon_T[\underbrace{h, \dots, h}_{b_0 \text{ terms}}, \underbrace{h_{\pi(1)}, \dots, h_{\pi(1)}}_{b_1 \text{ terms}}, \dots, \underbrace{h, h_{\pi(m)}, \dots, h_{\pi(m)}}_{b_m \text{ terms}}] \end{aligned} \quad (95c)$$

$$\leq \frac{T^{\frac{n-m}{2}} \|H\|_{L^2}^{n-m-q} \|h\|_{L^2}^q}{\prod_{i=0}^{m+q} a_i! \prod_{j=0}^m b_j!} \Upsilon_T[h_{\pi(1)}, h_{\pi(2)}, \dots, h_{\pi(m)}] \quad (95d)$$

$$\leq \frac{T^{\frac{n}{2}} \|H\|_{L^2}^{n-m-q} \|h\|_{L^2}^q}{\prod_{i=0}^{m+q} a_i! \prod_{j=0}^m b_j!} \|h_1\|_{L^2} \|h_2\|_{L^2} \cdots \|h_m\|_{L^2} \quad (95e)$$

$$\leq \frac{T^{\frac{n}{2}} \|H\|_{L^2}^{n-m-q} \|\delta H\|_{L^2}^q}{\prod_{i=0}^{m+q} a_i! \prod_{j=0}^m b_j!} \|\delta H_1\|_{L^2} \|\delta H_2\|_{L^2} \cdots \|\delta H_m\|_{L^2} \quad (95f)$$

by appealing to Lemmas 2 and 3. Recall, as a special case of the multinomial theorem [28], that

$$\sum_{c_1 + \dots + c_r = p} \frac{p!}{\prod_{i=1}^r c_i!} = r^p. \quad (96)$$

It follows that

$$\begin{aligned} & \|\Psi_{T,m,q}(H, \delta H)(\delta H_1, \delta H_2, \dots, \delta H_m)\| \\ & \leq \sum_{n=m+q}^{\infty} \sum_{\substack{a_0 + \dots + a_{m+q} = n-m-q \\ b_0 + \dots + b_m = q}} \sum_{\pi \in S_m} \frac{T^{\frac{n}{2}}}{\hbar^n \prod_{i=0}^{m+q} a_i! \prod_{j=0}^m b_j!} \times \\ & \quad \times \|H\|_{L^2}^{n-m-q} \|\delta H\|_{L^2}^q \prod_{j=1}^m \|\delta H_j\|_{L^2} \end{aligned} \quad (97a)$$

$$\begin{aligned} & = \sum_{n=m+q}^{\infty} \frac{m! T^{\frac{n}{2}} (m+q+1)^{n-m-q} (m+1)^q}{\hbar^n (n-m-q)! q!} \|H\|_{L^2}^{n-m-q} \times \\ & \quad \times \|\delta H\|_{L^2}^q \prod_{j=1}^m \|\delta H_j\|_{L^2} \end{aligned} \quad (97b)$$

$$\begin{aligned} & = \frac{m! T^{\frac{m+q}{2}} (m+1)^q}{\hbar^{m+q} q!} \exp\left(\frac{(m+q+1)\sqrt{T}}{\hbar} \|H\|_{L^2}\right) \times \\ & \quad \times \|\delta H\|_{L^2}^q \prod_{j=1}^m \|\delta H_j\|_{L^2} \end{aligned} \quad (97c)$$

so that the sum converges absolutely,  $\Psi_{T,m,q}(H, \delta H)$  is a bounded  $m$ -multilinear operator for each  $H \in \mathbb{H}$  and  $\delta H \in \mathbb{T}\mathbb{H}$ , and

$$\begin{aligned} & \|\Psi_{T,m,q}(H, \delta H)\| \\ & \leq \frac{m! T^{\frac{m+q}{2}} (m+1)^q}{\hbar^{m+q} q!} \exp\left(\frac{(m+q+1)\sqrt{T}}{\hbar} \|H\|_{L^2}\right) \|\delta H\|_{L^2}^q \\ & < \infty. \end{aligned} \quad (98)$$

Since  $\varphi_{T,m}(H) = \Psi_{T,m,0}(H, \delta H)$ , this conclusion also holds for  $\varphi_{T,m}(H)$ . ■

**Theorem 3.**  $Z_T$  is infinitely Fréchet differentiable, i.e.  $C^\infty$ , everywhere on  $\mathbb{H}$ .

*Proof:* We begin by establishing that  $\varphi_{T,m}$  is Fréchet differentiable for each  $m = 0, 1, 2, \dots$ , with derivative  $\varphi_{T,m+1}$ . Observe that

$$\begin{aligned} & \varphi_{T,m+1}(H)(\delta H_1, \dots, \delta H_m, \delta H) \\ & = \Psi_{T,m,1}(H, \delta H)(\delta H_1, \dots, \delta H_m) \end{aligned} \quad (99)$$

and that since the defining sum for  $\varphi_{T,m}(H + \delta H)$  converges absolutely, it may be rearranged as

$$\begin{aligned} & \varphi_{T,m}(H + \delta H)(\delta H_1, \dots, \delta H_m) \\ & = \sum_{q=0}^{\infty} \Psi_{T,m,q}(H, \delta H)(\delta H_1, \dots, \delta H_m). \end{aligned} \quad (100)$$

Then by appealing to the bound of  $\Psi_{T,m,q}$  in (97c), we get

$$\begin{aligned} & \|\varphi_{T,m}(H + \delta H) - \varphi_{T,m}(H) - \varphi_{T,m+1}(H)(\cdot, \dots, \cdot, \delta H)\| \\ & = \sup_{\{\|\delta H_j\|=1\}} \|\varphi_{T,m}(H + \delta H)(\delta H_1, \dots, \delta H_m) \\ & \quad - \varphi_{T,m}(H)(\delta H_1, \dots, \delta H_m) \\ & \quad - \varphi_{T,m+1}(H)(\delta H_1, \dots, \delta H_m, \delta H)\| \end{aligned} \quad (101a)$$

$$= \sup_{\{\|\delta H_j\|=1\}} \left\| \sum_{q=2}^{\infty} \Psi_{T,m,q}(H, \delta H)(\delta H_1, \dots, \delta H_m) \right\| \quad (101b)$$

$$\leq \sup_{\{\|\delta H_j\|=1\}} \sum_{q=2}^{\infty} \|\Psi_{T,m,q}(H, \delta H)(\delta H_1, \dots, \delta H_m)\| \quad (101c)$$

$$\begin{aligned} & \leq \sum_{q=2}^{\infty} \frac{m! T^{\frac{m+q}{2}} (m+1)^q}{\hbar^{m+q} q!} \times \\ & \quad \times \exp\left(\frac{(m+q+1)\sqrt{T}}{\hbar} \|H\|_{L^2}\right) \|\delta H\|_{L^2}^q \end{aligned} \quad (101d)$$

$$\begin{aligned} & = \frac{m! T^{\frac{m+2}{2}} (m+1)^2}{\hbar^{m+2}} \exp\left(\frac{(m+3)\sqrt{T}}{\hbar} \|H\|_{L^2}\right) \|\delta H\|_{L^2}^2 \\ & \quad \times \exp\left[\frac{(m+1)\sqrt{T}}{\hbar} \|\delta H\|_{L^2} \exp\left(\frac{\sqrt{T}}{\hbar} \|H\|_{L^2}\right)\right]. \end{aligned} \quad (101e)$$

Hence,

$$\begin{aligned} & \lim_{\|\delta H\| \rightarrow 0} \frac{1}{\|\delta H\|} \|\varphi_{T,m}(H + \delta H) - \varphi_{T,m}(H) \\ & \quad - \varphi_{T,m+1}(H)(\cdot, \dots, \cdot, \delta H)\| = 0 \end{aligned} \quad (102)$$

and therefore  $\varphi_{T,m}$  is Fréchet differentiable with derivative  $\varphi_{T,m+1}$ . Since  $Z_T = \varphi_{T,0}$ , this implies that  $Z_T$  is infinitely Fréchet differentiable, and that the  $m$ 'th derivative of  $Z_T$  is  $\varphi_{T,m}$ . ■

**Lemma 5.** Let  $\hat{\mathcal{H}} : \mathbb{K} \rightarrow \mathbb{H}(\mathcal{H})$  be defined by  $\hat{\mathcal{H}}(\mathcal{E})(t) = H_0 - \mu\mathcal{E}(t)$  for some fixed Hermitian operators  $H_0$  and  $\mu$  in  $\mathcal{M}(\mathcal{H})$ . Then  $\hat{\mathcal{H}}$  is infinitely Fréchet differentiable, i.e.  $C^\infty$ .

*Proof:* Let  $\zeta : \mathbb{K} \rightarrow \mathcal{B}(\mathbb{TK}; \mathbb{TH}(\mathcal{H}))$  be defined by  $\zeta(\mathcal{E})(\delta\mathcal{E})(t) = -\mu\delta\mathcal{E}(t)$ . For each  $\mathcal{E} \in \mathbb{K}$ ,  $\zeta(\mathcal{E})$  is linear, and  $\|\zeta(\mathcal{E})(\delta\mathcal{E})\|_{L^2} = \|\mu\|_{\text{HS}}\|\delta\mathcal{E}\|_{L^2}$ , so  $\zeta(\mathcal{E})$  is bounded. Now,

$$\lim_{\|\delta\mathcal{E}\| \rightarrow 0} \frac{\|\hat{\mathcal{H}}(\mathcal{E} + \delta\mathcal{E}) - \hat{\mathcal{H}}(\mathcal{E}) - \zeta(\mathcal{E})(\delta\mathcal{E})\|}{\|\delta\mathcal{E}\|} = 0 \quad (103)$$

so that  $\zeta$  is the Fréchet derivative of  $\hat{\mathcal{H}}$ . Since  $\zeta$  is constant (i.e.,  $\zeta(\mathcal{E})$  is the same linear operator regardless of which  $\mathcal{E} \in \mathbb{K}$  is input), the higher Fréchet derivatives also exist and are all equal to zero. ■

**Theorem 4.**  $U_T = Z_T \circ \hat{\mathcal{H}} : \mathbb{K} \rightarrow \mathbb{U}(\mathcal{H})$  is a composition of  $C^\infty$  maps and therefore is itself a  $C^\infty$  map.

## APPENDIX B GRADIENT OF $J_G$

Several steps in Section IV require differentiation of expressions involving the matrix logarithm. Since the expressions to be differentiated are all similar, this appendix will demonstrate the computation of the gradient of the kinematic landscape  $J_G$ , as the other variations follow along similar lines. To this end, we fix some target  $W \in \mathbb{U}(\mathcal{H})$  and recall

$$J_G(U) := \frac{1}{2} \|\log(U^\dagger W)\|_{\text{HS}}^2. \quad (104)$$

The differential of this function may then be written

$$\begin{aligned} d_U J_G(\delta U) &= \langle \log(U^\dagger W), d_{U^\dagger W} \log(-U^\dagger \delta U U^\dagger W) \rangle_{\text{HS}} \quad (105a) \\ &= \langle -U[(d_{U^\dagger W} \log)^*(\log(U^\dagger W))]W^\dagger U, \delta U \rangle, \quad (105b) \end{aligned}$$

where  $(d_{U^\dagger W} \log)^*$  is the adjoint (super-)operator. As a result, the gradient, which is the dual vector in  $\mathbb{T}_U \mathbb{U}(\mathcal{H})$  of the differential functional, is given by

$$\text{grad } J_G(U) = -U[(d_{U^\dagger W} \log)^*(\log(U^\dagger W))]W^\dagger U \quad (106)$$

It follows from the inverse relationship of the operator logarithm and exponential that  $\exp \circ \log = \text{id}$ , whence  $d_{\log Z} \exp \circ d_Z \log = \text{id}$  for any  $Z \in \mathbb{U}(\mathcal{H})$ , and therefore

$$(d_Z \log)^* \circ (d_{\log Z} \exp)^* = \text{id}. \quad (107)$$

Now, it is well-known [29], [30] that

$$d_{\log Z} \exp(X) = \int_0^1 e^{s \log(Z)} X e^{(1-s) \log(Z)} ds, \quad (108)$$

and therefore

$$(d_{\log Z} \exp)^*(X) = \int_0^1 e^{-s \log(Z)} X e^{-(1-s) \log(Z)} ds, \quad (109)$$

It follows that

$$\begin{aligned} (d_{\log U^\dagger W} \exp)^*(U^\dagger W \log(U^\dagger W)) &= \int_0^1 e^{-s \log(U^\dagger W)} U^\dagger W \log(U^\dagger W) e^{-(1-s) \log(U^\dagger W)} ds \\ &= \log(U^\dagger W), \end{aligned} \quad (110)$$

and consequently, using (107),

$$\begin{aligned} (d_{U^\dagger W} \log)^*(\log(U^\dagger W)) &= (d_{U^\dagger W} \log)^* \circ (d_{\log U^\dagger W} \exp)^*(U^\dagger W \log(U^\dagger W)) \\ &= U^\dagger W \log(U^\dagger W). \end{aligned} \quad (111)$$

We therefore can rewrite the gradient of  $J_G$  as

$$\begin{aligned} \text{grad } J_G(U) &= -U[(d_{U^\dagger W} \log)^*(\log(U^\dagger W))]W^\dagger U \\ &= -W \log(U^\dagger W)W^\dagger U \\ &= -U \log(U^\dagger W). \end{aligned} \quad (112)$$

where the last step follows from the fact that  $W^\dagger U$  commutes with  $\log(U^\dagger W)$ .

## REFERENCES

- [1] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal, "Quantum optimally controlled transition landscapes," *Science*, vol. 303, pp. 1998–2001, Mar. 26 2004. [Online]. Available: <http://www.sciencemag.org/content/303/5666/1998.abstract>
- [2] —, "Landscape for optimal control of quantum-mechanical unitary transformations," *Phys. Rev. A*, vol. 72, p. 052337, 2005. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.72.052337>
- [3] H. A. Rabitz, T.-S. Ho, M. M. Hsieh, R. Kosut, and M. Demiralp, "Topology of optimally controlled quantum mechanical transition probability landscapes," *Phys. Rev. A*, vol. 74, p. 012721, 2006. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.74.012721>
- [4] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal, "Optimal control landscapes for quantum observables," *J. Chem. Phys.*, vol. 124, p. 204107, 2006. [Online]. Available: <http://link.aip.org/link/?JCP/124/204107/1>
- [5] R. Wu, H. Rabitz, and M. Hsieh, "Characterization of the critical submanifolds in quantum ensemble control landscapes," *J. Phys. A: Math. Theor.*, vol. 41, p. 015006, 2008. [Online]. Available: <http://stacks.iop.org/1751-8121/41/i=1/a=015006>
- [6] R. Wu, A. Pechen, H. Rabitz, M. Hsieh, and B. Tsou, "Control landscapes for observable preparation with open quantum systems," *J. Math. Phys.*, vol. 49, p. 022108, 2008. [Online]. Available: <http://link.aip.org/link/?JMP/49/022108/1>
- [7] M. Hsieh, R. Wu, C. Rosenthal, and H. Rabitz, "Topological and statistical properties of quantum control transition landscapes," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 41, no. 7, p. 074020, 2008. [Online]. Available: <http://stacks.iop.org/0953-4075/41/i=7/a=074020>
- [8] M. Hsieh and H. Rabitz, "Optimal control landscape for the generation of unitary transformations," *Phys. Rev. A*, vol. 77, p. 042306, 2008. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.77.042306>
- [9] T.-S. Ho, J. Dominy, and H. Rabitz, "The landscape of unitary transformations in controlled quantum dynamics," *Phys. Rev. A*, vol. 79, p. 013422, 2009. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.79.013422>
- [10] R. Bott, "Nondegenerate critical manifolds," *Ann. of Math.*, vol. 60, no. 2, pp. 248–261, 1954. [Online]. Available: <http://www.jstor.org/stable/1969631>
- [11] M. F. Atiyah and R. Bott, "The Yang-Mills equations over Riemann surfaces," *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, vol. 308, no. 1505, pp. 523–615, 1983. [Online]. Available: <http://dx.doi.org/10.1098/rsta.1983.0017>
- [12] L. Nicolaescu, *An Invitation to Morse Theory*. New York: Springer, 2007.
- [13] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*. Springer, 1996.

- [14] V. Beltrani, J. Dominy, T.-S. Ho, and H. Rabitz, “Exploring the top and bottom of the quantum control landscape,” *J. Chem. Phys.*, vol. 134, no. 19, p. 194106, 2011. [Online]. Available: <http://link.aip.org/link/doi/10.1063/1.3589404>
- [15] R. Wu, R. Long, J. Dominy, T.-S. Ho, and H. Rabitz, “Singularities of quantum control landscapes,” *Phys. Rev. A*, vol. 86, no. 1, p. 013405, 2012. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.86.013405>
- [16] P. de Fouquieries and S. G. Schirmer, “Quantum control landscapes: A closer look,” 2010. [Online]. Available: <http://arxiv.org/abs/1004.3492>
- [17] A. N. Pechen and D. J. Tannor, “Are there traps in quantum control landscapes?” *Phys. Rev. Lett.*, vol. 106, p. 120402, Mar. 2011. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevLett.106.120402>
- [18] H. Rabitz, T.-S. Ho, R. Long, R. Wu, and C. Brif, “Comment on “are there traps in quantum control landscapes?”,” *Phys. Rev. Lett.*, vol. 108, p. 198901, May 2012. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevLett.108.198901>
- [19] J. P. Palao and R. Kosloff, “Optimal control theory for unitary transformations,” *Phys. Rev. A*, vol. 67, no. 6, p. 062308, Dec. 2003. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.68.062308>
- [20] M. P. do Carmo, *Riemannian Geometry*. Boston: Birkhäuser, 1992.
- [21] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*. New York: Springer, 1983.
- [22] A. W. Knap, *Lie Groups: Beyond an Introduction*, 2nd ed. Boston: Birkhäuser, 2004.
- [23] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*. Providence, RI: American Mathematical Society, 2001.
- [24] J. Milnor, *Morse Theory*, ser. Annals of Mathematics Studies. Princeton: Princeton University Press, 1973, vol. 51.
- [25] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, 2nd ed. New York: Springer, 1988.
- [26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.
- [27] J. Dominy and H. Rabitz, “Exploring families of controls for generating unitary transformations,” *J. Phys. A: Math. Theor.*, vol. 41, no. 20, p. 205305, May 2008. [Online]. Available: <http://stacks.iop.org/1751-8121/41/i=20/a=205305>
- [28] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.
- [29] R. Karplus and J. Schwinger, “A note on saturation in microwave spectroscopy,” *Phys. Rev.*, vol. 73, pp. 1020–1026, May 1948. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRev.73.1020>
- [30] R. Mathias, “Evaluating the Frechet derivative of the matrix exponential,” *Numer. Math.*, vol. 63, pp. 213–226, 1992. [Online]. Available: <http://dx.doi.org/10.1007/BF01385857>