Exploring families of quantum controls for generating unitary transformations

2008 J. Phys. A: Math. Theor. 41 205305

(http://iopscience.iop.org/1751-8121/41/20/205305)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 205305 (21pp) doi:10.1088/1751-8113/41/20/205305

Exploring families of quantum controls for generating unitary transformations

Jason Dominy[1] **and Herschel Rabitz**[1,2]

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

E-mail: jdominy@princeton.edu and hrabitz@princeton.edu

Received 16 February 2008, in final form 4 April 2008
Published 2 May 2008
Online at stacks.iop.org/JPhysA/41/205305

Abstract
A numerical method is presented for exploring the intertwined roles of the control-field structure and the final time T in determining the unitary evolution operator $U(T, 0)$ for finite-level quantum control systems. The algorithm can (a) identify controls achieving a target unitary operator W at time T up to machine precision and (b) identify a continuous family of controls producing the same operator W over a continuous interval of final times. The high degree of precision is obtained, in part, by exploiting the geometry of the unitary group. In particular, geodesics of the unitary group are followed, both for tracking to a target transformation and for error management.

PACS numbers: 32.80.Qk, 02.30.Zz, 03.67−a

1. Introduction

A fundamental concern in the control of quantum systems is the degree of exhibited control. This paper considers finite N-level systems with Hamiltonians of the form

$$H(t) = H_0 - \mu\mathcal{E}(t), \tag{1}$$

typically represented in the basis of the eigenstates of the field-free Hamiltonian H_0, so that H_0 is real diagonal, and the dipole μ is Hermitian. The theory of controllability of right-invariant affine systems on Lie groups, developed in [1], and applied to finite-level quantum systems in [2], allows for assessing the controllability of such quantum systems by analyzing only the dimension of the Lie algebra $\mathcal{A}$ generated by the skew-Hermitian matrices $\mathrm{i}H_0$ and $\mathrm{i}\mu$ under matrix commutation. These fundamental controllability results state that if $\mathcal{A}$ is the entire space of $N \times N$ skew-Hermitian matrices $u(N)$, then for any $W \in U(N)$ (the $N \times N$ unitary group), there exists a control $\mathcal{E}$ and a final time T such that the unitary propagator $U(t, 0)$ evolves to W at T under this control (where $U(t, t_0)$ denotes the unitary propagator from time t_0 to time t). In addition, [1] reveals that, if the system is controllable and the Hamiltonian traceless (so that the propagator is restricted to the special unitary group $SU(N)$), then there exists a $\hat{T} > 0$ such that for any $T \geqslant \hat{T}$, and any $W \in SU(N)$, there exists a control such that $U(T, 0) = W$. The role of T in quantum control has also been explored in more recent investigations [3–10] into time-optimal control, where the smallest time T is sought at which a given target unitary operator W may be attained.

This paper introduces a numerical method, denoted unitary D-MORPH, that explores the relationship between the control field, the final time T and the resulting unitary propagator $U(T, 0)$. Unitary D-MORPH is a variation of the original D-MORPH (diffeomorphic modulation under observable-response-preserving homotopy) method, presented in [11–13], that considered the relationship between the control field and the transition probability $P_{i\to f}$. Unitary D-MORPH numerically inverts the control-field $\mapsto$ unitary-propagator map to find a path through control-field space that corresponds to given tracks for T and $U(T, 0)$ through $\mathbb{R}_+$ and $U(N)$, respectively. We find that unitary D-MORPH is able to identify controls that produce a target unitary transformation at a sufficiently large time T with remarkably high accuracy ($<10^{-13}$ for 2-qubit systems). In addition, the algorithm can identify a continuous family of controls producing the same operator W over a continuous interval of final times. Beyond these two particular applications examined in the present work, the unitary D-MORPH formulation may be configured to address other questions in quantum control. For example, it can be configured to explore a given level set (the set of all controls that produce some fixed unitary transformation W at some fixed time T) by sending continuous trajectories along it in analogy with previous work [13].

The remainder of the paper is organized as follows. Section 2 breaks down the nonlinear inverse problem into an iterative sequence of underdetermined linear inverse problems. Section 3 describes some aspects of the geometry of level sets within the space of control fields and places unitary D-MORPH within this geometric context. Section 4 derives the space of all solutions to the linear inverse problem in the case where the Hamiltonian is of the form (1), and H_0 and μ satisfy the Lie algebra rank condition for controllability. Section 5 describes the application of this algorithm to the control problems outlined above. Section 6 discusses possible ways of taking advantage of the underdetermined nature (and hence multiplicity of solutions) of the linear inverse problem to optimize properties of the solution. Some details of the numerical implementation of this method are considered in section 7, and section 8 presents numerical results related to these applications. Finally, section 9 summarizes the work and suggests further extensions and applications of the unitary D-MORPH algorithm.

2. General formulation

Let $\mathbb{K}$ denote the space of admissible control functions, taken to be the linear subspace of locally bounded functions within $L^2(\mathbb{R}_+; \mathbb{R})$. Let $\mathbb{M}(N)$ be the space of $N \times N$ Hermitian matrices, $\mathbb{H}$ the subspace of locally bounded maps within $L^2(\mathbb{R}_+; \mathbb{M}(N))$, and $\hat{\mathcal{H}} : \mathbb{K} \to \mathbb{H}$ the differentiable map that assigns a time-dependent Hamiltonian to each control function. We will consider $\mathbb{C}^{N\times N}$ to be a *real* Hilbert space with the real Hilbert–Schmidt inner product $\langle A, B\rangle = \mathrm{Re}\,\mathrm{Tr}(A^\dagger B)$. This inner product then induces an inner product on $\mathbb{M}(N)$ as well as a Riemannian metric on $U(N)$ that will be the conventions adopted herein.

Denote by $V : \mathbb{K} \times \mathbb{R}_+ \to U(N)$ the time-dependent input-state map of the quantum control system; i.e. $V(\mathcal{E}, t) = U(t, 0)[\mathcal{E}]$ is the unitary time-evolution operator for the Schrödinger equation with Hamiltonian $\hat{\mathcal{H}}(\mathcal{E})$, evaluated at time t. For simplicity, we will also use the notation $V_t : \mathbb{K} \to U(N)$ and $V_{\mathcal{E}} : \mathbb{R}_+ \to U(N)$ for the map V with fixed t and $\mathcal{E}$, respectively. Similarly, we will use $\mathcal{H}_t(\mathcal{E}) = \mathcal{H}(t, \mathcal{E}) = \hat{\mathcal{H}}(\mathcal{E})(t)$ to denote the Hamiltonian for control $\mathcal{E}$, evaluated at time t.

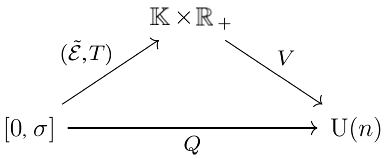

Figure 1. Commutative diagram relating V, the tracks Q and T, and the solution $\tilde{\mathcal{E}}$.

Suppose for some $\sigma > 0$ that $Q : [0, \sigma] \to U(N)$ and $T : [0, \sigma] \to \mathbb{R}_+$ are defined differentiable tracks (possibly constant maps) for $U(T, 0)$ and T, respectively. Then we may seek a path $\tilde{\mathcal{E}} : [0, \sigma] \to \mathbb{K}$ such that

$$V(\tilde{\mathcal{E}}(s), T(s)) = Q(s) \tag{2}$$

for all $s \in [0, \sigma]$, a relationship summarized in the commutative diagram in figure 1. This will be referred to as the *global nonlinear inverse problem*. Unitary D-MORPH seeks to solve this problem through a sequential linearization scheme. As shown in appendix A, V_T is a C^∞ map on $\mathbb{K}$ for each fixed $T > 0$. Differentiating (2) gives

$$\mathrm{d}_{\tilde{\mathcal{E}}(s)} V_{T(s)} \left(\frac{\mathrm{d}\tilde{\mathcal{E}}}{\mathrm{d}s} \right) + \frac{\mathrm{d}V_{\tilde{\mathcal{E}}(s)}}{\mathrm{d}t}(T(s)) \frac{\mathrm{d}T}{\mathrm{d}s} = \frac{\mathrm{d}Q}{\mathrm{d}s}, \tag{3}$$

where $\mathrm{d}_{\mathcal{E}} V_t : \mathbb{K} \to \mathrm{T}_{V_t(\mathcal{E})} U(N)$ is the differential of V_t at $\mathcal{E} \in \mathbb{K}$. Here and elsewhere in this paper, $\mathrm{T}_x X$ denotes the tangent space to the manifold X at the point $x \in X$. So, if we have some $\mathcal{E}_0 \in \mathbb{K}$ such that $V(\mathcal{E}_0, T(0)) = Q(0)$, then the solution $\tilde{\mathcal{E}}$ to (2) is the integral of the initial value problem $\frac{\mathrm{d}\tilde{\mathcal{E}}}{\mathrm{d}s} = \delta\tilde{\mathcal{E}}(s)$ where, for each $s \in [0, \sigma]$, $\delta\tilde{\mathcal{E}}(s)$ is *any* solution of the underdetermined linear inverse problem

$$\mathrm{d}_{\tilde{\mathcal{E}}(s)} V_{T(s)}(\delta\tilde{\mathcal{E}}(s)) = Q'(s) - T'(s) \frac{\mathrm{d}V_{\tilde{\mathcal{E}}(s)}}{\mathrm{d}t}(T(s)), \tag{4}$$

referred to henceforth as the *local linear inverse problem*.

To solve the local linear inverse problem, we need to determine expressions for the derivatives $\mathrm{d}V_{\mathcal{E}}/\mathrm{d}t$ and $\mathrm{d}_{\mathcal{E}} V_T$. The expression for $\mathrm{d}V_{\mathcal{E}}/\mathrm{d}t$ may be obtained directly from the Schrödinger equation

$$\frac{\mathrm{d}V_{\mathcal{E}}}{\mathrm{d}t}(T) = \left.\frac{\mathrm{d}U(t, 0)}{\mathrm{d}t}\right|_{t=T} = -\frac{\mathrm{i}}{\hbar}\mathcal{H}(T, \mathcal{E})U(T, 0) = -\frac{\mathrm{i}}{\hbar}\mathcal{H}(T, \mathcal{E})V_{\mathcal{E}}(T). \tag{5}$$

With some analysis, it may be shown (see appendix A) that V_T is Fréchet differentiable with respect to the control field and that the differential is given by

$$\mathrm{d}_{\mathcal{E}} V_T(\delta\mathcal{E}) = -\frac{\mathrm{i}}{\hbar} V_T(\mathcal{E}) \int_0^T V_t^\dagger(\mathcal{E})\, \mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E})(t) V_t(\mathcal{E})\, \mathrm{d}t. \tag{6}$$

Note that since $Q(s)$ defines a differentiable curve through $U(N)$, this curve may be specified by a Schrödinger-like initial value problem in s

$$\frac{\mathrm{d}Q}{\mathrm{d}s}(s) = \mathrm{i}Q(s)B(s), \tag{7}$$

where $B : [0, \sigma] \to \mathbb{M}(N)$ fills the role of the 'Hamiltonian', and $Q(0) = V(\mathcal{E}(\cdot, 0), T(0))$ is given; specifying $B(\cdot)$ in turn yields a track $Q(\cdot)$. With equations (5)–(7), we may now rewrite the linear inverse problem of (4) as

$$-\int_0^T V_t^\dagger(\tilde{\mathcal{E}}(s))\, \mathrm{d}_{\tilde{\mathcal{E}}(s)}\hat{\mathcal{H}} \left(\frac{\mathrm{d}\tilde{\mathcal{E}}}{\mathrm{d}s} \right)(t) V_t(\tilde{\mathcal{E}}(s))\, \mathrm{d}t = \frac{\mathrm{d}T}{\mathrm{d}s}(s) V_T^\dagger(\tilde{\mathcal{E}}(s))\mathcal{H}(T, \tilde{\mathcal{E}}(s))V_T(\tilde{\mathcal{E}}(s)) + \hbar B(s). \tag{8}$$

Note that these equations may not have a solution for $\mathrm{d}T/\mathrm{d}s \neq 0$ or $B(s) \neq 0$. One way to formally guarantee a solution is to assume that $\mathrm{d}_{\mathcal{E}} V_T$ is surjective for all $\mathcal{E} \in \mathbb{K}$ and all $T \in \mathbb{R}_+$. This is equivalent to the claim in several quantum control landscape papers (e.g. [14]) that the input-state map V_T is a submersion (i.e. has no critical points). Such an assumption is not necessary, however, to make use of the unitary D-MORPH algorithm, as will be discussed in the following section. Note also that, while this framework allows us to observe the response to a steadily decreasing T, these regularity requirements on V_T almost certainly prevent it from approaching the optimal time for a given target W. As a consequence, D-MORPH is likely not an effective tool for finding time-optimal controls.

3. Level sets: the geometry of D-MORPH

For a fixed $W \in U(N)$ and a fixed $T > 0$, we denote by $\Phi_{W,T}$ the level set (or fiber) of W at T,

$$\Phi_{W,T} = V_T^{-1}(W) = \{\mathcal{E} \in \mathbb{K} : V_T(\mathcal{E}) = W\}. \tag{9}$$

For a given control $\mathcal{E} \in \mathbb{K}$, the level set containing $\mathcal{E}$ may also be denoted by

$$\Phi_{\mathcal{E},T} = V_T^{-1}(V_T(\mathcal{E})) = \{\hat{\mathcal{E}} \in \mathbb{K} : V_T(\hat{\mathcal{E}}) = V_T(\mathcal{E})\}. \tag{10}$$

When the tracks $T(s)$ and $Q(s)$ are constant maps, the right-hand side of equation (8) is zero, and the control function 'curve' $\mathcal{E}(\cdot, s) := \tilde{\mathcal{E}}(s)$ is confined to the level set $\Phi_{\mathcal{E}(\cdot,0),T}$. These level sets exist in an infinite-dimensional space, and are not easy to visualize. However, the local submersion theorem [15] reveals that in a neighborhood of a regular point $\mathcal{E}$ of V_T (the only regions where D-MORPH may be applied), the level set containing $\mathcal{E}$, $\Phi_{\mathcal{E},T}$, is a codimension N^2 submanifold of $\mathbb{K}$. As the tracks $Q(s)$ and $T(s)$ vary smoothly with s, the corresponding level set $\Phi_{Q(s),T(s)}$ will also vary smoothly away from critical points. From this perspective, it is helpful to picture D-MORPH geometrically, as an algorithm defining a path through $\mathbb{K}$ that stays on the correct level set as that level set evolves with the parameter s.

From this picture the critical points evidently pose a hazard while implementing D-MORPH. As the algorithm evolves over s, the appearance of critical points in $\Phi_{Q(s),T(s)}$ can herald topological changes in the level set. These changes may include disconnected components being pinched off, merged, disappearing entirely, or emerging. Fortunately however, such behavior seems rare in practice, and in some cases may be avoided in a subsequent run by changing aspects of the track being followed. In addition, although such events will effectively terminate a D-MORPH run, they may help to reveal important control landscape topological information.

4. The algorithm in the dipole approximation

For the analysis that follows, let $\{\Omega_j\}_{j=1,\ldots,N^2}$ be any basis for $\mathbb{M}(N)$ (the space of $N \times N$ Hermitian matrices, endowed with the Hilbert–Schmidt inner product). For the numerical examples presented herein, we use the particular orthogonal basis of matrices of the form $|q\rangle\langle q|$, $(1/2)(|j\rangle\langle k| + |k\rangle\langle j|)$ or $(\mathrm{i}/2)(|j\rangle\langle k| - |k\rangle\langle j|)$ for $1 \leqslant q \leqslant N$ and $1 \leqslant j < k \leqslant N$. In addition, let $\boldsymbol{\nu} : \mathbb{M}(N) \to \mathbb{R}^{N^2}$ and $\nu_j : \mathbb{M}(N) \to \mathbb{R}$ be maps $\nu_j(A) := (\boldsymbol{\nu}(A))_j = \mathrm{Tr}(A\Omega_j) = \langle A, \Omega_j\rangle$ that 'vectorize' a Hermitian matrix A with respect to the chosen basis $\{\Omega_j\}$.

Let $\hat{\mathcal{H}}(\mathcal{E}) = H_0 - \mu\mathcal{E}$ for some fixed H_0 and μ. Then $\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}) = -\mu\delta\mathcal{E}$, so that the left-hand side of (8) becomes

$$-\int_0^T U^\dagger(t,0)\,\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}\,(\delta\mathcal{E})\,(t)U(t,0)\,\mathrm{d}t = \int_0^T U^\dagger(t,0)\mu U(t,0)\delta\mathcal{E}(t) = \int_0^T \mu(t)\delta\mathcal{E}(t)\,\mathrm{d}t, \tag{11}$$

where $\mu(t) = U^\dagger(t,0)\mu U(t,0)$. Since the algorithm consists of solving a sequence of linear problems—each at a fixed *s*—we have adopted the notational convention of leaving the *s*-dependence implicit.

If we define $\boldsymbol{\alpha}_{\mathcal{E},T} : \mathbb{R}\times\mathbb{M}(N)\to\mathbb{R}^{N^2}$ by

$$\boldsymbol{\alpha}_{\mathcal{E},T}(\delta T, B) = \delta T\boldsymbol{\nu}\big(V_T^\dagger(\mathcal{E})\mathcal{H}(T,\mathcal{E})V_T(\mathcal{E})\big) + \hbar\boldsymbol{\nu}(B), \tag{12}$$

then (8) can be written

$$\int_0^T \boldsymbol{\nu}(\mu(t))\delta\mathcal{E}(t)\,\mathrm{d}t = \boldsymbol{\alpha}_{\mathcal{E},T}(\delta T, B), \tag{13}$$

which is equivalent to the system of equations

$$\langle\omega_{\mathcal{E},T,j}, \delta\mathcal{E}\rangle_{\mathbb{K}} = \alpha_{\mathcal{E},T,j}(\delta T, B) \qquad \forall j = 1,\ldots,N^2, \tag{14}$$

where $\omega_{\mathcal{E},T,j} : \mathbb{R}_+\to\mathbb{R}$ is given by

$$\omega_{\mathcal{E},T,j}(t) = \begin{cases}\nu_j(\mu(t)) = \mathrm{Tr}(U^\dagger(t,0)\mu U(t,0)\Omega_j), & t\in[0,T],\\ 0, & \text{else,}\end{cases} \tag{15}$$

or equivalently $\omega_{\mathcal{E},T,j} = \hbar(\mathrm{d}_{\mathcal{E}}V_T)^*(\mathrm{i}V(\mathcal{E},T)\Omega_j)$, where $(\mathrm{d}_{\mathcal{E}}V_T)^*$ denotes the adjoint operator of the differential, and $\{\mathrm{i}V(\mathcal{E},T)\Omega_j\}$ is a basis for $\mathrm{T}_{V(\mathcal{E},T)}U(N)$.

Since this is a linear inverse problem, it may be solved by finding the set of all solutions to the homogeneous problem and one particular solution to the inhomogeneous problem. The space of solutions $\delta\mathcal{E}$ to the homogeneous problem

$$\langle\omega_{\mathcal{E},T,j}, \delta\mathcal{E}\rangle_{\mathbb{K}} = 0 \qquad \forall\, j = 1,\ldots,N^2, \tag{16}$$

is the linear subspace of $\mathbb{K}$ which is the orthogonal complement of the span of $\{\omega_{\mathcal{E},T,j}\}$. Assuming that the functions $\{\omega_{\mathcal{E},T,j}\}$ are linearly independent, this space of solutions will have codimension N^2 within the infinite-dimensional space $\mathbb{K}$. We can define the Gram (or overlap) matrix

$$(\mathcal{S}_{\mathcal{E},T})_{i,j} = \langle\omega_{\mathcal{E},T,i}, \omega_{\mathcal{E},T,j}\rangle_{\mathbb{K}} = \int_0^\infty \omega_{\mathcal{E},T,i}(t)\omega_{\mathcal{E},T,j}(t)\,\mathrm{d}t = \int_0^T \omega_{\mathcal{E},T,i}(t)\omega_{\mathcal{E},T,j}(t)\,\mathrm{d}t \tag{17}$$

and the map $\mathbf{p}_{\mathcal{E},T} : \mathbb{K}\to\mathbb{R}^{N^2}$ by

$$(\mathbf{p}_{\mathcal{E},T}(f))_i = \langle\omega_{\mathcal{E},T,i}, f\rangle_{\mathbb{K}} = \int_0^\infty \omega_{\mathcal{E},T,i}(t)f(t)\,\mathrm{d}t = \int_0^T \omega_{\mathcal{E},T,i}(t)f(t)\,\mathrm{d}t. \tag{18}$$

It may be observed that $\boldsymbol{\omega}_{\mathcal{E},T}, \mathbf{p}_{\mathcal{E},T}$ and $\mathcal{S}_{\mathcal{E},T}$ are representations of $(\mathrm{d}_{\mathcal{E}}V_T)^*, \mathrm{d}_{\mathcal{E}}V_T$ and $(\mathrm{d}_{\mathcal{E}}V_T)\circ(\mathrm{d}_{\mathcal{E}}V_T)^*$, respectively, all relative to the basis $\{\mathrm{i}V(\mathcal{E},T)\Omega_j\}$ of $\mathrm{T}_{V(\mathcal{E},T)}U(N)$. The space of solutions to the homogeneous problem can then be described by projecting $\mathbb{K}$ into the desired subspace

$$N_{\mathcal{E},T} = \big\{\delta\mathcal{E} = f - \boldsymbol{\omega}_{\mathcal{E},T}^T\,\mathcal{S}_{\mathcal{E},T}^{-1}\mathbf{p}_{\mathcal{E},T}(f) : f\in\mathbb{K}\big\}. \tag{19}$$

The inhomogeneous problem admits the particular solution $\delta\mathcal{E} = \boldsymbol{\omega}_{\mathcal{E},T}^T\mathcal{S}_{\mathcal{E},T}^{-1}\boldsymbol{\alpha}_{\mathcal{E},T}(\delta T, B)$. Hence, the full set of solutions to the inhomogeneous problem of (8) is given by

$$M_{\mathcal{E},T}(\delta T, B) = \big\{\delta\mathcal{E} = f + \boldsymbol{\omega}_{\mathcal{E},T}^T\,\mathcal{S}_{\mathcal{E},T}^{-1}(\boldsymbol{\alpha}_{\mathcal{E},T}(\delta T, B) - \mathbf{p}_{\mathcal{E},T}(f)) : f\in\mathbb{K}\big\}. \tag{20}$$

Returning finally to the global nonlinear inverse problem, assume that $Q : [0,\sigma]\to U(N)$ and $T : [0,\sigma]\to\mathbb{R}_+$ are defined differentiable tracks (possibly constant maps) for $U(T,0)$

and T, respectively. Assume further that $\mathcal{H}(\cdot, \mathcal{E}(\cdot, 0)) := H_0 - \mu\mathcal{E}(\cdot, 0)$ is a Hamiltonian that integrates to $Q(0)$ at time $T(0)$ (i.e. $V(\mathcal{E}(\cdot, 0), T(0)) = Q(0)$). We may then solve for $\mathcal{E}(\cdot, s)$ by integrating

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\cdot, s) = \delta\tilde{\mathcal{E}}(s) \in M_{\mathcal{E}(\cdot,s),T(s)}\left(\frac{\mathrm{d}T}{\mathrm{d}s}, B(s)\right), \tag{21}$$

where, for each $s \in [0, \sigma]$, $\delta\tilde{\mathcal{E}}(s)$ may be chosen to be *any* tangent vector within the codimension N^2 (hence infinite dimensional) space of possibilities: $M_{\mathcal{E}(\cdot,s),T(s)}\,(\mathrm{d}T/\mathrm{d}s, B(s))$, making (21) a set-valued differential equation or *differential inclusion*. The multiplicity of possible solutions implied by this equation will be exploited in section 6.

4.1. *Operation on* SU(N)

If $\mathrm{Tr}(\mu) = \mathrm{Tr}(H_0) = 0$, then the image of V is contained entirely within SU(N), and the D-MORPH formulation presented above will fail because $\mathcal{S}_{H,T}$ is singular everywhere. However the formulation may be modified slightly in this circumstance to allow it to function properly. We need to simply replace the chosen basis for $\mathbb{M}(N)$, $\{\Omega_j\}$, with a basis for the traceless Hermitian matrices $\{\hat{\Omega}_j\}$. For example, one could use the matrices of the form $|q\rangle\langle q| - (1/N)\mathbb{I}$, $(1/2)(|j\rangle\langle k| + |k\rangle\langle j|)$ or $(\mathrm{i}/2)(|j\rangle\langle k| - |k\rangle\langle j|)$ for $1 \leqslant q \leqslant N-1$ and $1 \leqslant j < k \leqslant N$. This leaves a set of $N^2 - 1$ elements $\{\hat{\Omega}_j\}$, forming a basis for the space of traceless Hermitian $N \times N$ matrices. We may then define $\hat{\omega}_{\mathcal{E},T,j} = \hbar(\mathrm{d}_{\mathcal{E}}V_T)^*(\mathrm{i}V(\mathcal{E}, T)\hat{\Omega}_j)$, as well as the corresponding $\hat{\mathcal{S}}_{\mathcal{E},T}$ and $\hat{\mathbf{p}}_{\mathcal{E},T}$ to get the full set of local solutions

$$\hat{M}_{\mathcal{E},T}(\delta T, B) = \left\{\delta\mathcal{E} = f + \hat{\omega}^T_{\mathcal{E},T}\,\hat{\mathcal{S}}^{-1}_{\mathcal{E},T}(\alpha_{\mathcal{E},T}(\delta T, B) - \hat{\mathbf{p}}_{\mathcal{E},T}(f)) : f \in \mathbb{K}\right\} \tag{22}$$

for B traceless and Hermitian.

5. Theoretical applications

In this section, two distinct means of configuring the tracks $T(s)$ and $B(s)$ are considered, each leading to a different application of unitary D-MORPH to problems in quantum control. Concrete numerical simulations for each of these configurations will be discussed in section 8.

5.1. *The reachable time set*

For any given system (H_0, μ) and final time $T > 0$, let $A(\mathbb{I}, T) = \mathrm{Range}(V_T) \subset U(N)$ denote the set of all $N \times N$ unitary evolution operators that may be attained at time T. Then, for any $W \in U(N)$, let $R(W) = \{T \in \mathbb{R}_+ : W \in A(\mathbb{I}, T)\}$ denote the set of all final times at which W may be attained. If the system is controllable, then for any $W \in U(N)$, $R(W)$ is non-empty. If $T \in R(W)$ and there exists a control $\mathcal{E}$ such that $V_T(\mathcal{E}) = W$ and $\mathcal{E}$ is a regular point of V_T, then according to the D-MORPH formalism, T is an interior point of $R(W)$. Starting from a control field and time T that produce W, if $\mathcal{S}_{\mathcal{E},T}$ is invertible at this point, then other nearby times and control fields may be found using unitary D-MORPH that also produce W. This is done by taking $B = 0$ and $\delta T = \mathrm{d}T/\mathrm{d}s \neq 0$ in (20).

5.2. *Tracking to a given* $W \in U(N)$ *along a geodesic*

Suppose we have an initial electric field $\mathcal{E}(t, 0)$ and define $U_0 \in U(N)$ by $U_0 := V_T(\tilde{\mathcal{E}}(0))$ (i.e., U_0 is $U(T, 0)$ for the Hamiltonian $H(t) = H_0 - \mu\mathcal{E}(t, 0)$). Let $A = -(\mathrm{i}/\sigma)\log\left(U_0^\dagger W\right)$ where $\sigma > 0$ and the matrix logarithm is defined as follows. If the Hamiltonian has nonzero trace and the dynamics range over $U(N)$, this matrix logarithm is defined by choosing $\log\left(U_0^\dagger W\right)$ from the principal branch, i.e. all eigenvalues of $\log\left(U_0^\dagger W\right)$ must lie in $(-\mathrm{i}\pi, \mathrm{i}\pi]$. However, if the Hamiltonian is traceless, then A must also be made traceless. The trace of $\log\left(U_0^\dagger W\right)$ is always equal to $2\mathrm{i}\pi k$ for some integer k between $-N/2$ and $N/2$. If k is positive, we subtract $2\mathrm{i}\pi$ from each of the k largest eigenvalues of $\log\left(U_0^\dagger W\right)$. If k is negative, we add $2\mathrm{i}\pi$ to each of the k smallest eigenvalues. Then it may be shown [16] that $Q : [0, \sigma] \to U(N)$ given by $Q(s) = U_0 \exp(\mathrm{i}As)$ describes a minimal geodesic from U_0 to W. For $B(s)$ defined by $\mathrm{i}Q(s)B(s) := (\mathrm{d}Q/\mathrm{d}s) = \mathrm{i}Q(s)A$, we find that $B(s) = A$ describes this path. With this choice of $B(s)$, $U(T, 0; s)$ hits W exactly at $s = \sigma$ (and will continue past it if the simulation is not stopped), rather than approaching W asymptotically as can happen with other optimization schemes. This method clearly produces the shortest, straightest path from U_0 to W. However, because of the complexity in the dynamical evolution, the corresponding path $\mathcal{E}(\cdot, s)$ may be highly convoluted.

6. Making choices among the multiplicity of solutions

Because of the strongly underdetermined nature of the posed nonlinear inverse problem, the local linear inverse problem is likewise underdetermined, leading to a continuum of local solutions as seen in (20). As a consequence, the original nonlinear inverse problem also has a continuum of possible solutions. In this section we consider ways to take advantage of the infinite dimension of the space of local solutions to choose members that exhibit certain favorable qualities, leading toward an optimal value of some prescribed functional. Since all of the local solutions identified in (20) solve the problem exactly, the additional qualities may be gained without compromise.

Suppose that $\mathcal{J} : \mathbb{K} \to \mathbb{R}$ is some differentiable functional expressing a property of the solution control function that is advantageous to either maximize or minimize. In the homogeneous case where the tracks $Q(s)$ and $T(s)$ are constant maps, the control-field curve $\mathcal{E}(\cdot, s)$ lies entirely within the level set $\Phi_{\mathcal{E}(\cdot,0),T}$. Let $\mathcal{J}_0 : \Phi_{\mathcal{E}(\cdot,0),T} \to \mathbb{R}$ be the restriction of $\mathcal{J}$ to this level set, and since critical points are undesirable within unitary D-MORPH, we assume that $\mathcal{E}(\cdot, s)$ is a regular point of V_T for every $s \in [0, \sigma]$, so that $\Phi_{\mathcal{E}(\cdot,0),T}$ is locally a codimension N^2 submanifold near each $\mathcal{E}(\cdot, s)$. Recall that the gradient of $\mathcal{J}$ at $\mathcal{E}(\cdot, s)$ is the unique element $\nabla\mathcal{J}(\mathcal{E}(\cdot, s))$ of $\mathrm{T}_{\mathcal{E}(\cdot,s)}\mathbb{K}$ such that $\langle\nabla\mathcal{J}(\mathcal{E}(\cdot, s)), \delta\mathcal{E}(\cdot)\rangle = \mathrm{d}_{\mathcal{E}(\cdot,s)}\mathcal{J}(\delta\mathcal{E})$ for all $\delta\mathcal{E} \in \mathrm{T}_{\mathcal{E}(\cdot,s)}\mathbb{K}$. Likewise, the gradient of $\mathcal{J}_0$ at $\mathcal{E}(\cdot, s)$ is the unique element $\nabla\mathcal{J}_0(\mathcal{E}(\cdot, s))$ of $\mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T}$ such that $\langle\nabla\mathcal{J}_0(\mathcal{E}(\cdot, s)), \delta\mathcal{E}\rangle = \mathrm{d}_{\mathcal{E}(\cdot,s)}\mathcal{J}_0(\delta\mathcal{E})$ for all $\delta\mathcal{E} \in \mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T}$. Now, if $\delta\mathcal{E} \in \mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T} \subset \mathrm{T}_{\mathcal{E}(\cdot,s)}\mathbb{K}$ is tangent to $\Phi_{\mathcal{E}(\cdot,0),T}$ at $\mathcal{E}(\cdot, s)$, then $\mathrm{d}_{\mathcal{E}(\cdot,s)}\mathcal{J}_0(\delta\mathcal{E}) = \mathrm{d}_{\mathcal{E}(\cdot,s)}\mathcal{J}(\delta\mathcal{E})$. Hence, $\langle\nabla\mathcal{J}_0(\mathcal{E}(\cdot, s)), \delta\mathcal{E}\rangle = \langle\nabla\mathcal{J}_0(\mathcal{E}(\cdot, s)), \delta\mathcal{E}\rangle$ for all $\delta\mathcal{E} \in \mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T}$. Since $\nabla\mathcal{J}_0 \in \mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T}$, this implies that $\nabla\mathcal{J}_0$ is the projection of $\nabla\mathcal{J}$ into $\mathrm{T}_{\mathcal{E}(\cdot,s)}\Phi_{\mathcal{E}(\cdot,0),T}$. With the machinery already developed, this can be written

$$\nabla\mathcal{J}_0(\mathcal{E}) = \nabla\mathcal{J}(\mathcal{E}) - \boldsymbol{\omega}_{\mathcal{E},T}^T \mathcal{S}_{\mathcal{E},T}^{-1} \mathbf{p}_{\mathcal{E},T}(\nabla\mathcal{J}(\mathcal{E})) \tag{23}$$

which, in this homogeneous case, is a valid solution to the local inverse problem. Within the space of all possible solutions, $\nabla\mathcal{J}_0(\mathcal{E})$ is the one that points in the direction of steepest ascent of $\mathcal{J}$. So if we use $\nabla\mathcal{J}$ as the free function in D-MORPH, $\mathcal{E}(\cdot, s)$ will follow the gradient flow of $\mathcal{J}_0$ on $\Phi_{\mathcal{E}(\cdot,0),T}$.

In the inhomogeneous case (i.e., where one or both of $B(s)$ and $\mathrm{d}T/\mathrm{d}s(s)$ are nonzero) the situation is less straightforward. Let $g = \boldsymbol{\omega}_{\mathcal{E},T}^T \mathcal{S}_{\mathcal{E},T}^{-1} \boldsymbol{\alpha}_{\mathcal{E},T}(\delta T, B)$, so that for any $\lambda \in \mathbb{R}$, $\delta\mathcal{E} = \lambda\nabla\mathcal{J}_0(\mathcal{E}) + g$ is a valid solution to the local inverse problem. If $\mathrm{d}_{\mathcal{E}}\mathcal{J}(g) \geqslant 0$, then choosing $\nabla\mathcal{J}(\mathcal{E})$ as the free function will cause the value of $\mathcal{J}$ to never decrease. If $\mathrm{d}_{\mathcal{E}}\mathcal{J}(g) < 0$ and $\|\nabla\mathcal{J}_0(\mathcal{E})\| = 0$, then every $\delta\mathcal{E} \in M_{\mathcal{E},T}(\delta T, B)$ is a direction that decreases

$\mathcal{J}$. On the other hand, if $\mathrm{d}_{\mathcal{E}}\mathcal{J}(g) < 0$ and $\|\nabla\mathcal{J}_0(\mathcal{E})\| > 0$, then choosing $\lambda > 0$ large enough will result in a $\delta\mathcal{E}$ that increases $\mathcal{J}$. However, this may produce a very large $\delta\mathcal{E}$ (i.e., one with large norm), which is likely to be undesirable from a numerical standpoint. So, considering both analytic and numerical viewpoints, the use of $\nabla\mathcal{J}(\mathcal{E})$ (without any local scaling term λ) as the free function appears to be the most practical choice for maximizing a functional $\mathcal{J}$ over the course of the D-MORPH run.

One of the simplest and most useful applications of the concept above is to minimize the fluence

$$\mathcal{J}_{F,T}(\mathcal{E}) = \int_0^T \mathcal{E}^2(t)\mathcal{W}(t)\,\mathrm{d}t \tag{24}$$

where $\mathcal{W} : \mathbb{R}_+ \to \mathbb{R}_+$ is some arbitrary weight function. Then the gradient is just the functional derivative

$$\frac{\delta\mathcal{J}_{F,T}}{\delta\mathcal{E}(t)} = \begin{cases} 2\mathcal{E}(t)\mathcal{W}(t) & t \in [0, T] \\ 0 & \text{else.} \end{cases} \tag{25}$$

Minimizing the fluence with the free function $f(t) = -\delta\mathcal{J}_{F,T}/\delta\mathcal{E}(t)$ helps to 'regularize' the control field. Also, the fluence minimized fields tend to be much easier to interpret physically as nonessential intensity features in the control field are removed.

7. Numerical considerations

7.1. Error management through tracking

Recall that $Q : \mathbb{R}_+ \to U(N)$ represents the desired unitary track specified by $B(\cdot)$. Suppose that at some s_k, $U(T, 0; s_k) \neq Q(s_k)$. Then we could introduce a correction in the track that would attempt to reduce this error at s_{k+1}. Since $B(s_k)$ already is the approximate track to get from $Q(s_k)$ to $Q(s_{k+1})$, the correction $C(s_k)$ should only attempt to track from $U(T, 0; s_k)$ back to $Q(s_k)$. Then $B(s_k) + C(s_k)$ will be an approximate track to get from $U(T, 0; s_k)$ to $Q(s_{k+1})$. In particular, if $B \equiv 0$ (so that Q is fixed), then $C(s_k)$ will define a track back to Q.

Let $G : [0, s_{k+1} - s_k] \to U(N)$ be the minimal geodesic from $U(T, 0; s_k)$ (the actual computed solution at s_k) to $Q(s_k)$ (the desired solution at s_k) in an interval of length $\Delta s_k = s_{k+1} - s_k$. Then $G(r) = U(T, 0; s_k)\exp(\mathrm{i}rC(s_k))$, where $C(s_k) = -(\mathrm{i}/\Delta s_k)\log(U^\dagger(T, 0; s_k)Q(s_k))$ and the matrix logarithm should be defined as in section 5.2. Letting $\gamma = \nu(C(s_k))$, we can incorporate this correction as

$$\delta\mathcal{E}(t) = f(t) + \omega_{\mathcal{E},T}^T \mathcal{S}_{\mathcal{E},T}^{-1}(\alpha_{\mathcal{E},T}(\delta T, B) + \gamma - \mathbf{p}_{\mathcal{E},T}(f)). \tag{26}$$

7.2. Combined tracking and error correction

Better accuracy may be achieved if we eliminate this dual approach to tracking and error correction, and simply define a track $B(s_k)$ to get from $U(T, 0; s_k)$ (the actual computed solution at s_k) to $Q(s_{k+1})$ (the desired solution at s_{k+1}) in one s-step. Then in analogy with the error correction derivation, we define

$$B(s_k) = -\frac{\mathrm{i}}{s_{k+1} - s_k}\log(U^\dagger(T, 0; s_k)Q(s_{k+1})), \tag{27}$$

and

$$\delta\mathcal{E}(t) = f(t) + \omega_{\mathcal{E},T}^T \mathcal{S}_{\mathcal{E},T}^{-1}(\alpha_{\mathcal{E},T}(\delta T, B(s_k)) - \mathbf{p}_{\mathcal{E},T}(f)). \tag{28}$$

Due to lack of commutativity in $U(N)$, the $\alpha_{\mathcal{E},T}(\delta T, B(s_k))$ term from (28) is close, but not identical, to $\alpha_{\mathcal{E},T}(\delta T, B) + \gamma$ from (26).

7.3. Additional correction steps

In situations where very high accuracy is desired, the on-the-fly error correction described in the previous two sections may be insufficient. Orders of magnitude of additional accuracy may be obtained by introducing dedicated error correction steps between the standard 's' steps. Suppose that at some s_k, the algorithm has evolved the Hamiltonian to $H(\cdot, s_k)$ which produces $U(T, 0; s_k)$, but the desired unitary solution is $Q(s_k)$. Then defining again $C = -\mathrm{i}\log(U^\dagger(T, 0; s_k)Q(s_k))$, and $\gamma = \boldsymbol{\nu}(C)$, we can perturb the control field by

$$\delta\mathcal{E} = \boldsymbol{\omega}_{\mathcal{E},T}^T \mathcal{S}_{\mathcal{E},T}^{-1}\gamma. \tag{29}$$

It has been found empirically that ten such steps are sufficient to reduce the Hilbert–Schmidt norm $\|U(T, 0; s_k) - Q(s_k)\|$ from $\sim 10^{-4}$ to $\sim 10^{-14}$ for 2-qubit systems (i.e., down to machine precision). The high quality of unitary D-MORPH performance may be attributed to the fact that the geometry of the unitary group has been built into the algorithm, especially through the use of geodesic curves such as are employed in these correction steps.

7.4. Remeshing time

When the track $T(s)$ is non-constant, it typically becomes necessary to redefine the time vector and the control-field vector after each s-step. During these events, both the values of the t_k's as well as the number of such time points may change. Extra time points may be added to keep the spacing $\Delta t = t_{k+1} - t_k$ from growing too much when T increases, or to provide additional resolution as the control field $\mathcal{E}(t)$ grows either in intensity or in high-frequency structure so as to maintain control over the error in the integration of the Hamiltonian. Once the new time vector has been specified, the control field may be recreated using spline interpolation (and extrapolation when T is increasing).

7.5. Time mesh scaling

In addition to the 2-qubit simulations presented in the following section, a small number of 3- and 4-qubit simulations were performed. It was found that larger systems require much larger time meshes, and therefore much more memory and computational time. Numerically, we can only work with controls in $\mathbb{K}$ that are well approximated by piecewise constant functions with a modest number of 'pieces'. For small systems, this set of controls appears sufficient for controllability. However, because of the greater numbers of internal degrees of freedom in larger systems, this set of 'numerically nice' controls quickly becomes insufficient for complete unitary controllability as the system size increases. In particular, the greater numbers of natural frequencies in these larger systems necessitate finer time meshes to resolve higher frequencies and larger T to resolve smaller differences between the frequencies. Consequently, we believe this time mesh scaling behavior is due to the nature of the control landscape itself, and is largely independent of the algorithm employed in exploration.

8. Numerical examples

8.1. Variable final time example

8.1.1. CNOT gate. In this example, we consider the four-level (2-qubit) system [17]

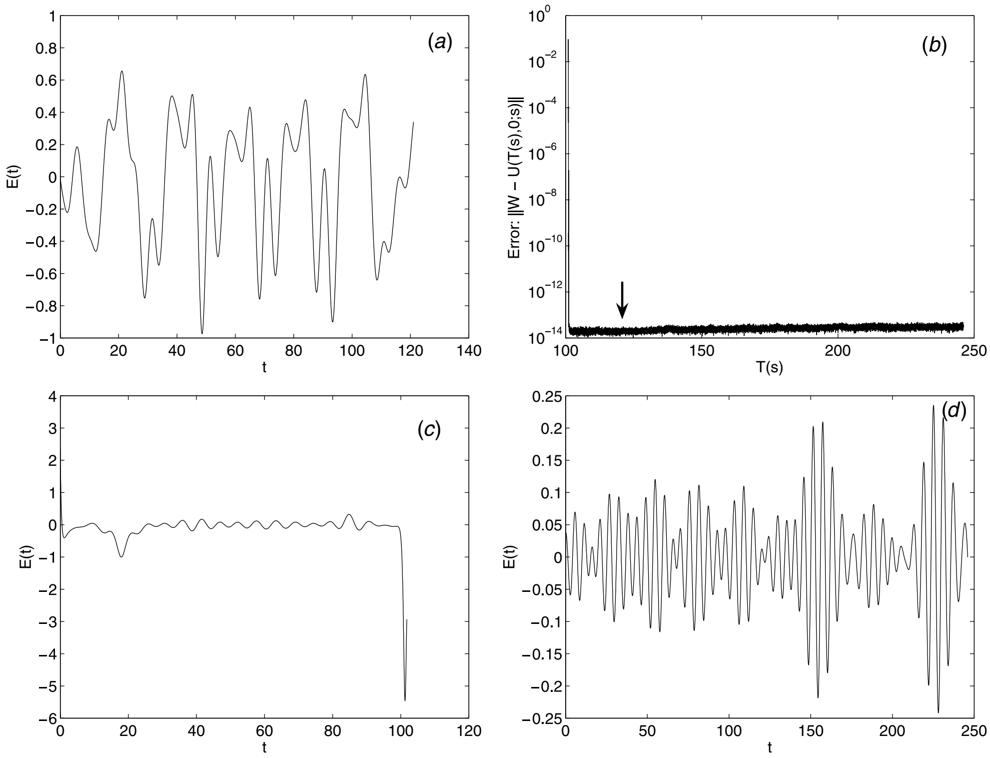

Figure 2. Results of continuously varying T for a 2-qubit system, while holding $U(T, 0; s) =$ CNOT to high precision. Panel (a) shows the initial field $\mathcal{E}_0$ which yields CNOT at $T = 121.05$. Panel (b) shows the error as T is decreased as well as increased from 121.05 (indicated with an arrow). Panel (c) is the control field at $T(s) = 101.7706$, just before the error began increasing. Panel (d) is the control field for $T(s) = 245.55$.

$$H_0 = \frac{1}{2}\begin{bmatrix} \omega_1 + \omega_2 - \frac{1}{2}\gamma & 0 & 0 & 0 \\ 0 & \omega_1 - \omega_2 + \frac{1}{2}\gamma & -\gamma & 0 \\ 0 & -\gamma & \omega_2 - \omega_1 + \frac{1}{2}\gamma & 0 \\ 0 & 0 & 0 & -\omega_1 - \omega_2 - \frac{1}{2}\gamma \end{bmatrix}$$

$$\mu = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \tag{30}$$

where $\omega_1 = 1$ and $\omega_2 = \pi - 2.05 \approx 1.0926$ are the transition angular frequencies for the two qubits, and $\gamma = 0.1$ is the Heisenberg exchange interaction coupling constant between them. We wish to address the problem of generating the controlled NOT (CNOT) gate

$$W := \exp\left(\mathrm{i}\frac{5\pi}{4}\right)\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{31}$$

while varying the final time T. We start with the field $\tilde{\mathcal{E}}(0)$ in figure 2(a) from [17], but further optimized using the error correction schemes of unitary D-MORPH. This starting field has

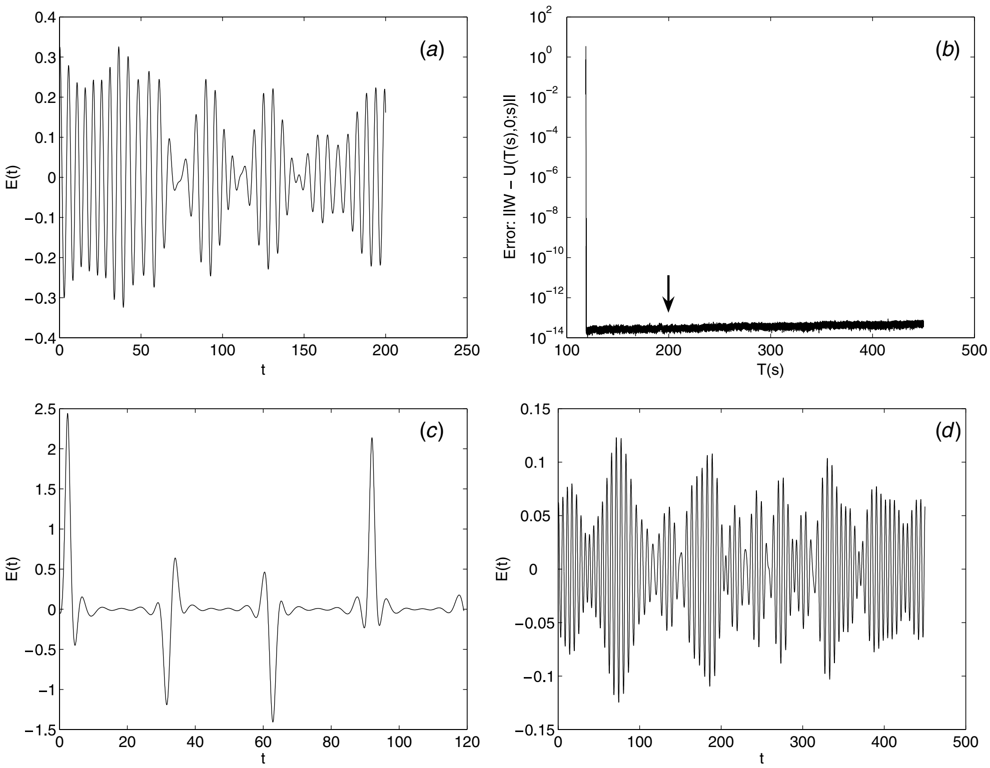

Figure 3. Results of continuously varying T for a 2-qubit system, while holding $U(T, 0; s) = \mathbb{I}$ to high precision. Panel (a) shows the initial field $\mathcal{E}_0$ which yields the identity transformation at $T = 200.05$. Panel (b) shows the error as T is decreased as well as increased from 200.05 (indicated with an arrow). Panel (c) is the control field found by unitary D-MORPH that yields the identity transformation at $T(s) = 119.05$, just before the error began increasing. Panel (d) is the control field for $T(s) = 450.05$.

an initial terminal time of $T(0) = 121.05$, and we defined an exponential track for T by specifying the derivative $\mathrm{d}T/\mathrm{d}s(s) = -0.5\exp(-0.5s/T(0))$, and a constant track for Q by specifying $B(s) \equiv 0$ for all s. We then propagated s with step size $\Delta s = 0.01$, the fluence minimizing free function, and ten additional error correction steps. Figure 2(b) indicates that for decreasing T, the error (measured as the Hilbert–Schmidt norm $\|U(T, 0; s) - W\|$) was kept below 5×10^{-14} until T reached about 101, at which point the performance stopped. Figure 2(c) represents the field at $T = 101.7706$ when the error was still 1.5943×10^{-14}. The most likely explanation for the catastrophic behavior near $T = 101$ is that the connected component of the level set that D-MORPH was following, vanished, leaving D-MORPH far away from any other field that can yield the CNOT gate (if, indeed, any such field exists for this T).

The algorithm was then restarted in the same configuration, but with a linear track for T, given by $\mathrm{d}T/\mathrm{d}s = 0.5$, in order to explore larger values of T. Figure 2(b) also shows the error observed in this case of $T > 121.05$. Again, the algorithm was able to keep the error below 5×10^{-14} for all T up to 245.55, at which point the algorithm was manually terminated. The field at $T = 245.55$ is shown in figure 2(d) to be quite regular in structure and of lower amplitude than in figure 2(a).

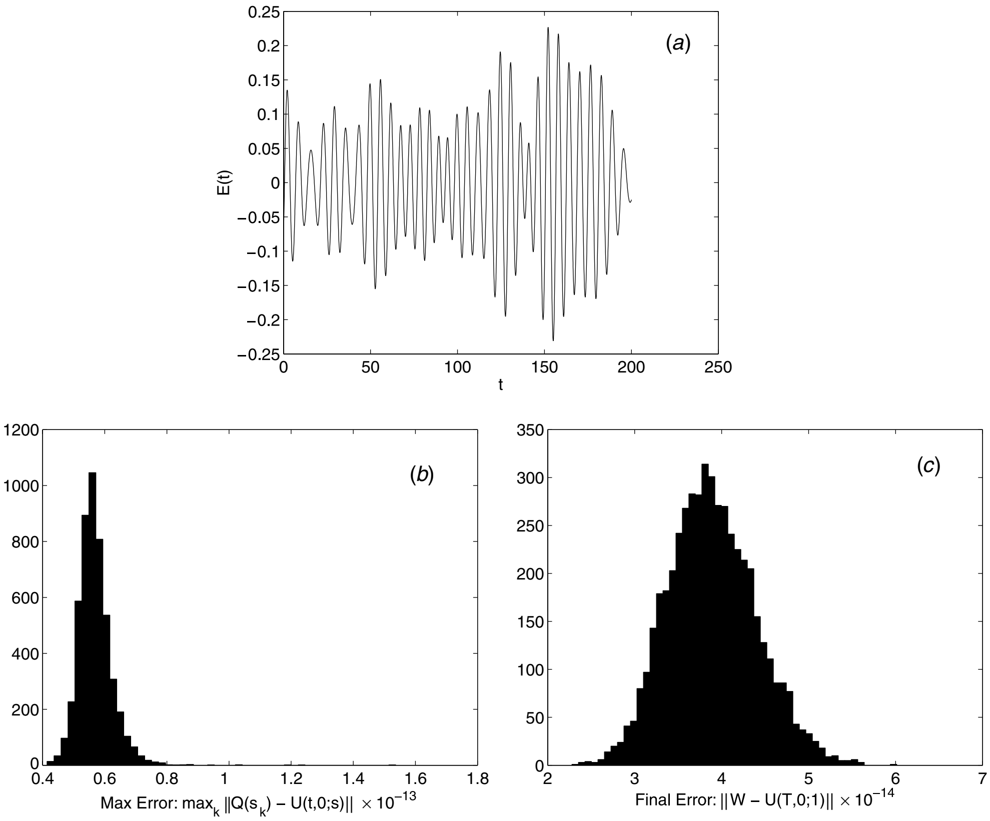

Figure 4. Results of a run of 5000 individual random geodesic tracks on the 2-qubit system with $T = 200.05$. Panel (*a*) shows the starting control field for the tracks. Panel (*b*) shows a histogram of the maximum error encountered in the 100-step tracks. Panel (*c*) shows the error histogram at the conclusion of each of the 5000 tracks.

8.1.2. Identity gate. This example uses the system described in (30), this time attempting to generate the identity matrix over a continuous interval of times *T*. The starting field for this exploration was found by taking the field at $T = 200.05$ from the CNOT example in section 8.1.1 (this field is shown in figure 4(*a*)), and tracking it along a geodesic with fixed *T* to get from CNOT to the identity. This procedure gave the starting field $\mathcal{E}_0$ shown in figure 3(*a*). From this point, a decreasing linear track for *T* was defined, with $\mathrm{d}T/\mathrm{d}s = -0.5$. We then propagated *s* with step size $\Delta s = 0.01$, the fluence minimizing free function, and ten additional error correction steps. For $T < 200.05$, figure 3(*b*) indicates that the error (measured as the Hilbert–Schmidt norm $\|U(T, 0; s) - \mathbb{I}\|$) was again maintained below 5×10^{-14} until *T* reached about 119, at which point the performance could not be maintained as in the example of figure 2. Figure 3(*c*) represents the field at $T = 119.05$ when the error was still 2.203×10^{-14}.

The algorithm was also restarted in the same configuration, but with an increasing linear track for $T > 200.05$, given by $\mathrm{d}T/\mathrm{d}s = 0.5$. Figure 3(*b*) shows the error observed in this case. Again, the algorithm was able to keep the error below 5×10^{-14} for all *T* up to 450.05, at which point the algorithm was manually terminated. The field at $T = 450.05$, shown in figure 3(*d*), is quite regular with largely amplitude modulation.

8.2. Exact-time controllability example

In this example, we take the fixed system in (30) and the fixed starting field $\mathcal{E}_0$ shown in figure 4(*a*), which was found at $T = 200.05$ during the course of the CNOT example described in section 8.1.1. Then 5000 random (Haar-distributed [18, 19]) special unitary matrices W were generated from SU(4). For each of these matrices W, a geodesic track was defined from the CNOT matrix generated by $\mathcal{E}_0$ to W as described in section 5.2. For each W, this geodesic track was discretized into 100 steps and followed, while allowing ten additional error correction steps between each s-step and using the fluence minimizing free function. The histograms in figures 4(*b*) and (*c*) document the maximum error seen during the course of each track, as well as the final error seen at the conclusion of the tracking procedure. It is evident that very high accuracy is maintained both during the evolution and at the conclusion of each track.

8.3. A remark on robustness

The numerical examples presented above employed the gradient of the fluence as the free function. As a result, the scheme was not specifically directed to seek out solutions that are robust to small perturbations (e.g., noise) in the control field. Indeed, post-facto analysis of some of these solutions reveals that the quality of the results is very sensitive to even 1% on-resonance Gaussian noise, bringing the error up from $\sim 10^{-14}$ to ~ 0.05. This observation is supported by the upper bound $\|\mathrm{d}_{\mathcal{E}} V_T(\delta\mathcal{E})\| \leqslant (\|\mu\|/\hbar)\|\delta\mathcal{E}\|_1$. It may be possible to avoid this upper bound by changing the statement of the optimization problem. In unitary D-MORPH, this could be accomplished by using the gradient of some robustness metric, rather than the fluence, as the free function f in (20).

9. Conclusions

A new algorithm is presented for exploring the relationship between the control field $\mathcal{E}$, the final time T, and the resulting unitary propagator $U(T, 0)[\mathcal{E}]$ in quantum control. By developing the algorithm around the geometry of the unitary and special unitary groups, a remarkable degree of accuracy may be achieved in reaching target unitary transformations and in holding that accuracy while the target time T is varied. This accuracy was demonstrated in two classes of numerical examples. The first type explored the continuous changes necessary in the control field to achieve a certain desired unitary transformation as the final time T was varied. And the second class of example found families of controls for a fixed T as the target unitary transformation was varied. The unitary D-MORPH algorithm enjoys a great deal of flexibility, only some of which was exploited in the present paper. The basic framework can easily be adapted to explore a variety of other questions in quantum control.

Acknowledgments

JD acknowledges support from the Program in Plasma Science and Technology at Princeton. This work was supported by the US Department of Energy.

Appendix. Differentiability with respect to the control

Standard theorems exist that prove the differentiability of the solution of a differential equation with respect to parameters (or controls), however they traditionally assume that the vector field defining the flow varies smoothly with time. Since the space of admissible controls in this paper ($\mathbb{K}$) includes non-smooth functions, these theorems do not directly apply. For completeness, this appendix proves that the unitary solution to Schrödinger's equation $V_T(\mathcal{E})$ is infinitely Fréchet differentiable with respect to the control, and gives an expression for the first derivative. This result shows that the notion of smooth control landscapes over $\mathbb{K}$ is well defined.

It should be noted that, since $H \in \mathbb{H}$ may be discontinuous, a classical solution of the differential form of the Schrödinger equation may not exist. However, a solution in the sense of Carathéodory [20] exists and is unique for every $H \in \mathbb{H}$, and this solution exists for all $t \in [0, \infty)$ [1]. In other words, for each $H \in \mathbb{H}$ (or each $\mathcal{E} \in \mathbb{K}$), there exists a unique absolutely continuous unitary path $U(t, 0)$ for all $t \in [0, \infty)$, with $U(0, 0) = \mathbb{I}$, and such that almost everywhere on $[0, \infty)$, $\mathrm{d}U/\mathrm{d}t$ exists and is equal to $-(\mathrm{i}/\hbar)H(t)U(t, 0)$. $V : \mathbb{K} \times \mathbb{R}_+ \to U(N)$ is therefore well defined on the entire domain in terms of these Carathéodory solutions.

For the theorem that follows, we will need this fundamental inequality:

Lemma 1 (Gronwall's inequality)**.** *Let φ, ψ, χ be real-valued measurable functions defined on an interval $[a, b]$ such that χ and $\chi\psi$ are integrable, and $\forall t \in [a, b]$*

$$\varphi(t) \leqslant \psi(t) + \int_a^t \chi(s)\varphi(s)\,\mathrm{d}s. \tag{A.1}$$

Then for all $t \in [a, b]$,

$$\varphi(t) \leqslant \psi(t) + \int_a^t \chi(s)\psi(s)\exp\left(\int_s^t \chi(u)\,\mathrm{d}u\right)\mathrm{d}s. \tag{A.2}$$

Proof. See [20, p 37]. □

Corollary 1. *In the previous lemma, if ψ is a non-negative, non-decreasing function and χ is non-negative, then*

$$\varphi(t) \leqslant \psi(t)\exp\left(\int_a^t \chi(s)\,\mathrm{d}s\right). \tag{A.3}$$

Proof. First, note that $\psi(s)$ can be pulled out of the integral in (A.2), leaving

$$\varphi(t) \leqslant \psi(t)\left(1 + \int_a^t \chi(s)\exp\left(\int_s^t \chi(u)\,\mathrm{d}u\right)\mathrm{d}s\right). \tag{A.4}$$

Using integration by parts, one can show that

$$\int_a^t \chi(s)\left(\int_s^t \chi(u)\,\mathrm{d}u\right)^k \mathrm{d}s = \frac{1}{k+1}\left(\int_a^t \chi(s)\,\mathrm{d}s\right)^{k+1} \tag{A.5}$$

from which the result follows by the power series representation of exp. □

The arguments to be presented make much use of the topologies of various Hilbert and Banach spaces, so we briefly summarize these topologies here. Although these spaces include complex elements, they are all taken to be vector spaces over $\mathbb{R}$, with real inner products in the case of the Hilbert spaces. As was stated earlier, we take the real Hilbert–Schmidt inner product $\langle A, B\rangle = \mathrm{Re}\,\mathrm{Tr}(A^\dagger B)$ for the topology on $\mathbb{C}^{N\times N}$ and use this to induce the inner product on $\mathbb{M}(N)$ and the Riemannian metric on $U(N)$. $L^2(\mathbb{R}_+; \mathbb{R})$ is given the inner product $\langle f, g\rangle_2 = \int_0^\infty f(t)g(t)\,\mathrm{d}t$. $L^2(\mathbb{R}_+; \mathbb{C}^{N\times N})$ is endowed with the inner product $\langle A, B\rangle_2 = \int_0^\infty \mathrm{Re}\,\mathrm{Tr}(A^\dagger(t)B(t))\,\mathrm{d}t$. $L^\infty(\mathbb{R}_+; \mathbb{R})$ is given the norm $\|f\|_\infty = \mathrm{ess\,sup}_{t\in[0,\infty)}|f(t)|$, i.e. $\|f\|_\infty$ is the smallest number such that $|f(t)| \leqslant \|f\|_\infty$ almost everywhere on $[0, \infty)$. We give $L^\infty(\mathbb{R}_+;\mathbb{C}^{N\times N})$ the norm $\|A\|_\infty = \text{ess sup}_{t\in[0,\infty)}\|A(t)\|$, where $\|A(t)\|$ is the norm defined by the real Hilbert–Schmidt inner product on $\mathbb{C}^{N\times N}$. Finally, we also make use of induced operator norms. If X and Y are two Banach spaces and A is a linear operator from X to Y, then we define $\|A\| := \sup_{\|x\|>0}(\|Ax\|/\|x\|)$.

Definition 1. *Let $\mathbb{T} = \mathcal{C}(\mathbb{R}_+;\mathbb{C}^{N\times N})$ be the space of continuous trajectories through $\mathbb{C}^{N\times N}$, where $\mathbb{T}$ is given the compact-open topology (i.e., the topology of uniform convergence on compact sets).*

Theorem 1. *Suppose that $\hat{\mathcal{H}} : \mathbb{K} \to \mathbb{H}$ is differentiable, and is such that $\hat{V} : \mathbb{K} \to \mathbb{T}$, defined implicitly through the Schrödinger equation $[\hat{V}(\mathcal{E})(t) = V_t(\mathcal{E})]$, is continuous. Then, for each fixed $T > 0$, the map $V_T : \mathbb{K} \to \mathbb{C}^{N\times N}$ is differentiable, with Fréchet derivative*

$$\mathrm{d}_{\mathcal{E}}V_T(\delta\mathcal{E}) = -\frac{\mathrm{i}}{\hbar}V_T(\mathcal{E})\int_0^T V_t^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E})\,\mathrm{d}t. \tag{A.6}$$

Proof. This proof follows closely the proof of lemma 4.1.9 in [15]. Let

$$\Psi(t,\mathcal{E},\delta\mathcal{E}) := -\frac{\mathrm{i}}{\hbar}V_t(\mathcal{E})\int_0^t V_\tau^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(\tau)V_\tau(\mathcal{E})\,\mathrm{d}\tau. \tag{A.7}$$

We wish to show that for any $T > 0$, and all $\mathcal{E} \in \mathbb{K}$,

$$\lim_{\|\delta E\|_2\to 0}\frac{\|V_T(\mathcal{E}+\delta\mathcal{E}) - V_T(\mathcal{E}) - \Psi(T,\mathcal{E},\delta\mathcal{E})\|}{\|\delta\mathcal{E}\|_2} = 0. \tag{A.8}$$

To that end, observe first that $V_T(\mathcal{E})$ satisfies the integral form of Schrödinger's equation: $V_T(\mathcal{E}) = \mathbb{I} - (\mathrm{i}/\hbar)\int_0^T H(t,\mathcal{E})V_t(\mathcal{E})\,\mathrm{d}t$. In addition, since $V_t(\mathcal{E})$ is absolutely continuous with respect to t and $V_t^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E})$ is locally bounded and measurable and therefore integrable, $\Psi(t,\mathcal{E},\delta\mathcal{E})$ is also absolutely continuous in t and therefore differentiable with respect to t almost everywhere. In fact,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Psi(t,\mathcal{E},\delta\mathcal{E}) &= -\frac{\mathrm{i}}{\hbar}\left(\frac{\mathrm{d}}{\mathrm{d}t}V_t(\mathcal{E})\right)\int_0^t V_\tau^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(\tau)V_\tau(\mathcal{E})\,\mathrm{d}\tau\\ &\quad - \frac{\mathrm{i}}{\hbar}V_t(\mathcal{E})V_t^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E}) \quad \text{a.e.}\\ &= \left(-\frac{\mathrm{i}}{\hbar}\right)^2\mathcal{H}_t(\mathcal{E})V_t(\mathcal{E})\int_0^t V_\tau^\dagger(\mathcal{E})(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(\tau)V_\tau(\mathcal{E})\,\mathrm{d}\tau\\ &\quad - \frac{\mathrm{i}}{\hbar}(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E}) \quad \text{a.e.}\\ &= -\frac{\mathrm{i}}{\hbar}\mathcal{H}_t(\mathcal{E})\Psi(t,\mathcal{E},\delta\mathcal{E}) - \frac{\mathrm{i}}{\hbar}(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E}) \quad \text{a.e.}\end{aligned} \tag{A.9}$$

Therefore $\Psi(T,\mathcal{E},\delta\mathcal{E})$ satisfies the integral equation

$$\Psi(T,\mathcal{E},\delta\mathcal{E}) = -\frac{\mathrm{i}}{\hbar}\int_0^T[\mathcal{H}_t(\mathcal{E})\Psi(t,\mathcal{E},\delta\mathcal{E}) + (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E})]\,\mathrm{d}t. \tag{A.10}$$

Then

$$\begin{aligned}V_T(\mathcal{E}+\delta\mathcal{E}) - V_T(\mathcal{E}) - \Psi(T,\mathcal{E},\delta\mathcal{E}) &= -\frac{\mathrm{i}}{\hbar}\int_0^T \mathcal{H}_t(\mathcal{E}+\delta\mathcal{E})V_t(\mathcal{E}+\delta\mathcal{E}) - \mathcal{H}_t(\mathcal{E})V_t(\mathcal{E})\,\mathrm{d}t\\ &\quad + \frac{\mathrm{i}}{\hbar}\int_0^T \mathcal{H}_t(\mathcal{E})\Psi(t,\mathcal{E},\delta\mathcal{E}) + (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)V_t(\mathcal{E})\,\mathrm{d}t\end{aligned} \tag{A.11}$$

$$= -\frac{\mathrm{i}}{\hbar}\int_0^T [\mathcal{H}_t(\mathcal{E}+\delta\mathcal{E}) - \mathcal{H}_t(\mathcal{E}) - (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)]V_t(\mathcal{E}+\delta\mathcal{E})\,\mathrm{d}t$$
$$- \frac{\mathrm{i}}{\hbar}\int_0^T (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)[V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})]\,\mathrm{d}t$$
$$- \frac{\mathrm{i}}{\hbar}\int_0^T \mathcal{H}_t(\mathcal{E})[V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E}) - \Psi(t,\mathcal{E},\delta\mathcal{E})]\,\mathrm{d}t. \qquad (A.12)$$

From this, we get the inequality

$$\|V_T(\mathcal{E}+\delta\mathcal{E}) - V_T(\mathcal{E}) - \Psi(T,\mathcal{E},\delta\mathcal{E})\|$$
$$\leqslant \frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E}+\delta\mathcal{E}) - \mathcal{H}_t(\mathcal{E}) - (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|\|V_t(\mathcal{E}+\delta\mathcal{E})\|\,\mathrm{d}t$$
$$+ \frac{1}{\hbar}\int_0^T \|(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|\|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|\,\mathrm{d}t$$
$$+ \frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E})\|\|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E}) - \Psi(t,\mathcal{E},\delta\mathcal{E})\|\,\mathrm{d}t \qquad (A.13)$$
$$= A(T,\mathcal{E},\delta\mathcal{E})$$
$$+ \frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E})\|\|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E}) - \Psi(t,\mathcal{E},\delta\mathcal{E})\|\,\mathrm{d}t, \qquad (A.14)$$

where $A(T,\mathcal{E},\delta\mathcal{E})$ encompasses the first two terms of (A.13). Now, since by assumption, for each $\mathcal{E} \in \mathbb{K}$, $\|\mathcal{H}_t(\mathcal{E})\|$ is locally integrable, we can apply Gronwall's inequality (Cor. 1) to (A.14) for each fixed $\mathcal{E}, \delta\mathcal{E} \in \mathbb{K}$ to get

$$\|V_T(\mathcal{E}+\delta\mathcal{E}) - V_T(\mathcal{E}) - \Psi(T,\mathcal{E},\delta\mathcal{E})\| \leqslant A(T,\mathcal{E},\delta\mathcal{E})\exp\left(\frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E})\|\,\mathrm{d}t\right), \qquad (A.15)$$

where $\exp\left(\frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E})\|\mathrm{d}t\right)$ is a finite number independent of $\delta\mathcal{E}$.

It remains to show that $A(T,\mathcal{E},\delta\mathcal{E})/\|\delta\mathcal{E}\|_2 \to 0$ as $\|\delta\mathcal{E}\|_2 \to 0$. Consider the first term of $A(T,\mathcal{E},\delta\mathcal{E})$. Since $V_t(\mathcal{E}) \in U(N)$ for all $t \geqslant 0$ and $\mathcal{E} \in \mathbb{K}$, $\|V_t(\mathcal{E}+\delta\mathcal{E})\| = \sqrt{N}$. Then invoking the Cauchy–Schwarz inequality, we can show that

$$\frac{1}{\hbar}\int_0^T \|\mathcal{H}_t(\mathcal{E}+\delta\mathcal{E}) - \mathcal{H}_t(\mathcal{E}) - (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|\|V_t(\mathcal{E}+\delta\mathcal{E})\|\,\mathrm{d}t$$
$$\leqslant \frac{\sqrt{NT}}{\hbar}\left(\int_0^T \|\mathcal{H}_t(\mathcal{E}+\delta\mathcal{E}) - \mathcal{H}_t(\mathcal{E}) - (\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$
$$\leqslant \frac{\sqrt{NT}}{\hbar}\|\hat{\mathcal{H}}(\mathcal{E}+\delta\mathcal{E}) - \hat{\mathcal{H}}(\mathcal{E}) - \mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E})\|. \qquad (A.16)$$

Now, for the second term of $A(T,\mathcal{E},\delta\mathcal{E})$, note that since $\hat{\mathcal{H}}$ is differentiable, $\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}$ is a bounded linear operator so that for each $\mathcal{E} \in \mathbb{K}$, $\|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}\| < \infty$ and for every $\delta\mathcal{E} \in \mathbb{K}$, $\|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E})\| \leqslant \|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}\|\|\delta\mathcal{E}\|_2$. So by the Cauchy–Schwarz inequality we get

$$\frac{1}{\hbar}\int_0^T \|(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|\|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|\,\mathrm{d}t$$
$$\leqslant \frac{1}{\hbar}\left(\int_0^T \|(\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}))(t)\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}\left(\int_0^T \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$
$$\leqslant \frac{\|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E})\|}{\hbar}\left(\int_0^T \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$

$$\leqslant \frac{\|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}\|\|\delta\mathcal{E}\|_2}{\hbar}\left(\int_0^T \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}. \tag{A.17}$$

Therefore, we get

$$\frac{A(T,\mathcal{E},\delta\mathcal{E})}{\|\delta\mathcal{E}\|_2} \leqslant \frac{\sqrt{NT}}{\hbar}\frac{\|\hat{\mathcal{H}}(\mathcal{E}+\delta\mathcal{E}) - \hat{\mathcal{H}}(\mathcal{E}) - \mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E})\|}{\|\delta\mathcal{E}\|_2} + \frac{\|\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}\|}{\hbar}\left(\int_0^T \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}. \tag{A.18}$$

The first of these terms vanishes as $\|\delta\mathcal{E}\|_2 \to 0$ because of the differentiability of $\hat{\mathcal{H}}$. And, because of the assumed continuity of the map $\hat{V} : \mathbb{K} \to \mathbb{T}$ and the compact-open topology placed on $\mathbb{T}$, as $\|\delta\mathcal{E}\|_2 \to 0$, $\sup_{t\in[0,T]} \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\| \to 0$, so that $\left(\int_0^T \|V_t(\mathcal{E}+\delta\mathcal{E}) - V_t(\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}} \to 0$. Therefore the second term also vanishes as $\|\delta\mathcal{E}\|_2 \to 0$, so

$$\lim_{\|\delta\mathcal{E}\|_2\to 0} \frac{A(T,\mathcal{E},\delta\mathcal{E})}{\|\delta\mathcal{E}\|_2} = 0 \tag{A.19}$$

and the proof is complete. □

Corollary 2. *If $\hat{\mathcal{H}}$ is given by $\hat{\mathcal{H}}(\mathcal{E}) = H_0 - \mu\mathcal{E}$ for some fixed $N \times N$ Hermitian H_0 and μ, then V_T is Fréchet differentiable and*

$$\mathrm{d}_{\mathcal{E}} V_T(\delta\mathcal{E}) = \frac{\mathrm{i}}{\hbar} V_T(\mathcal{E}) \int_0^T V_t^{\dagger}(\mathcal{E})\mu V_t(\mathcal{E})\delta\mathcal{E}(t)\,\mathrm{d}t. \tag{A.20}$$

Proof. It was proved in [1, 21] that the map $\hat{V} : \mathbb{K} \to \mathbb{T}$ is continuous for this form of the Hamiltonian. In addition $\hat{\mathcal{H}}$ is differentiable with $\mathrm{d}_{\mathcal{E}}\hat{\mathcal{H}}(\delta\mathcal{E}) = -\mu\delta\mathcal{E}$. So theorem 1 applies and gives the stated result. □

Henceforth in this appendix, we will assume for simplicity that $\hat{\mathcal{H}}(\mathcal{E}) = H_0 - \mu\mathcal{E}$.

Definition 2. *If X and Y are normed linear spaces, let $\mathcal{B}^r(X;Y)$ be the space of bounded r-multilinear operators from $X^r = X \times X \times \ldots \times X$ to Y, with the norm*

$$\|A\| = \sup_{\{\|x_j\|\neq 0\}} \frac{\|A(x_1,\ldots,x_r)\|}{\|x_1\|\cdots\|x_r\|} \tag{A.21}$$

for each $A \in \mathcal{B}^r(X;Y)$. Let $\mathcal{B}^0(X;Y) = Y$. The arguments in this appendix will make particular use of the spaces $\mathcal{B}^r(\mathbb{K}; L^2([0,T];\mathbb{C}^{N\times N}))$ and $\mathcal{B}^r(\mathbb{K}; L^\infty([0,T];\mathbb{C}^{N\times N}))$. To distinguish between these two norm topologies where it is not clear from the context, the norm on $\mathcal{B}^r(\mathbb{K}; L^2([0,T];\mathbb{C}^{N\times N}))$ will be denoted $\|\cdot\|_2$ and that on $\mathcal{B}^r(\mathbb{K}; L^\infty([0,T];\mathbb{C}^{N\times N}))$ will be denoted $\|\cdot\|_\infty$.

Definition 3. *For any permutation $\pi \in S_r$ (the symmetric group on r elements), let $\mathcal{P}_\pi : \mathcal{B}^r(X;Y) \to \mathcal{B}^r(X;Y)$ be defined by $\mathcal{P}_\pi(A)(x_1,\ldots,x_r) := A(x_{\pi(1)},\ldots,x_{\pi(r)})$ for any $A \in \mathcal{B}^r(X;Y)$. For any $\hat{\pi} \in \bigoplus_{r=0}^{\infty} S_r$, let $\hat{\mathcal{P}}_{\hat{\pi}} : \bigoplus_{r=0}^{\infty}\mathcal{B}^r(X;Y) \to \bigoplus_{r=0}^{\infty}\mathcal{B}^r(X;Y)$ be defined by $(\hat{\mathcal{P}}_{\hat{\pi}}(\hat{A}))_r := P_{\pi_r}(A_r) \in \mathcal{B}^r(X;Y)$, where π_r and A_r are the rth 'coordinates' of $\hat{\pi}$ and $\hat{A}$.*

Lemma 2. *$\mathcal{P}_\pi$ is a bounded linear operator on $\mathcal{B}^r(X;Y)$ with operator norm $\|\mathcal{P}_\pi\| = 1$, for any normed linear spaces X and Y.*

Proof. $\|\mathcal{P}_\pi(A)\| = \sup_{\{\|x_j\|\neq 0\}} \frac{\|\mathcal{P}_\pi(A)(x_1,\ldots,x_r)\|}{\|x_1\|\cdots\|x_r\|} = \sup_{\{\|x_{\pi(j)}\|\neq 0\}} \frac{\|A(x_{\pi(1)},\ldots,x_{\pi(r)})\|}{\|x_{\pi(1)}\|\cdots\|x_{\pi(r)}\|} = \|A\|$. □

Definition 4. *Fix $T > 0$. Let $\mathcal{F}$ be any space of integrable matrix-valued functions $F : [0, T] \to \mathbb{C}^{N\times N}$, that is closed under indefinite integration, i.e. for any $F \in X$, the function $\hat{F}$ defined by $\hat{F}(t) = \int_0^t F(t)dt$ also lies in $\mathcal{F}$. Let $\mathcal{I}_T : \mathcal{B}^r(\mathbb{K}; \mathcal{F}) \to \mathcal{B}^r(\mathbb{K}; \mathbb{C}^{N\times N})$ be the definite integral*

$$\mathcal{I}_T(M)(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r) = \int_0^T M(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t)\,\mathrm{d}t \tag{A.22}$$

for any $M \in \mathcal{B}^r(\mathbb{K}; \mathcal{F})$ and let $\hat{\mathcal{I}}_T : \mathcal{B}^r(\mathbb{K}; \mathcal{F}) \to \mathcal{B}^r(\mathbb{K}; \mathcal{F})$ be the indefinite integral

$$\hat{\mathcal{I}}_T(M)(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t) = \int_0^t M(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(\tau)\,\mathrm{d}\tau \qquad \forall t \in [0, T]. \tag{A.23}$$

Lemma 3. *$\mathcal{I}_T$ and $\hat{\mathcal{I}}_T$ are bounded linear operators in the cases where the space $\mathcal{F}$ is either $L^2([0, T]; \mathbb{C}^{N\times N})$ or $L^\infty([0, T]; \mathbb{C}^{N\times N})$, hence $\mathcal{I}_T$ and $\hat{\mathcal{I}}_T$ are continuous and Fréchet differentiable and their derivatives are $\mathcal{I}_T$ and $\hat{\mathcal{I}}_T$, respectively. In addition, $\mathcal{I}_T$ may be thought of as a linear operator on $\mathcal{B}^r(\mathbb{K}, \mathcal{F})$ where the output is constant with respect to time, i.e. $\mathcal{I}_T(M)(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t) = \int_0^T M(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(\tau)\,\mathrm{d}\tau$ for all $t \in [0, T]$. In this context, and for either choice of $\mathcal{F}$, $\mathcal{I}_T$ remains a bounded linear operator.*

Definition 5. *Denote by $\Gamma_{T,r}$ the vector space (over $\mathbb{R}$) of Fréchet differentiable maps $B : \mathbb{K} \to \mathcal{B}^r(\mathbb{K}; L^2([0, T]; \mathbb{C}^{N\times N}))$ such that $B(\mathbb{K}) \subset \mathcal{B}^r(\mathbb{K}; L^\infty([0, T]; \mathbb{C}^{N\times N}))$, and B is continuous when considered as a map $\mathbb{K} \to \mathcal{B}^r(\mathbb{K}; L^\infty([0, T]; \mathbb{C}^{N\times N}))$. Let $\Gamma_T = \bigoplus_{r=0}^\infty \Gamma_{T,r}$ be the direct sum of these spaces. On Γ_T, define multiplication of elements $B \in \Gamma_{T,r}$ and $C \in \Gamma_{T,q}$ by*

$$(B\cdot C)(\mathcal{E})(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_{r+q})(t) = [B(\mathcal{E})(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t)][C(\mathcal{E})(\delta\mathcal{E}_{r+1}, \ldots, \delta\mathcal{E}_{r+q})(t)] \tag{A.24}$$

for all $\mathcal{E} \in \mathbb{K}$, $\{\delta\mathcal{E}_j\} \subset \mathbb{K}$, and $t \in [0, T]$. The Fréchet derivative of an element of Γ_T will be defined 'grade-wise'. In other words, $B \in \Gamma_T$ may be thought of as a sequence $B = \{B_i\}$ with $B_i \in \Gamma_{T,i}$, and $\mathrm{d}B$ is defined to be $\mathrm{d}B = \{\mathrm{d}B_i\}$.

Lemma 4. *With respect to this multiplication, $\Gamma_{T,r}\cdot\Gamma_{T,q} \subset \Gamma_{T,r+q}$ and $\mathrm{d}_{\mathcal{E}}(B\cdot C)(\delta\mathcal{E}) = (\mathrm{d}_{\mathcal{E}}B(\delta\mathcal{E}))\cdot C + B\cdot(\mathrm{d}_{\mathcal{E}}C(\delta\mathcal{E}))$, i.e. $\mathrm{d}(B\cdot C) = (\mathrm{d}B)\cdot C + \mathcal{P}_\pi \circ [B\cdot(\mathrm{d}C)]$ for some $\pi \in S_{r+q+1}$. Hence, Γ_T is a graded associative algebra of maps on which the usual product rule applies. In addition $\hat{\mathcal{I}}_T \circ \Gamma_T \subset \Gamma_T$ and $\mathcal{I}_T \circ \Gamma_T \subset \Gamma_T$ (where $\mathcal{I}_T$ is regarded as a map to $\mathcal{B}^r(\mathbb{K}; L^2([0, T]; \mathbb{C}^{N\times N}))$ whose output is constant with respect to time), i.e. Γ_T is closed under both definite and indefinite time integration. Furthermore, each $\Gamma_{T,r}$ is closed with respect to the permutation operations $\mathcal{P}_\pi$ for all $\pi \in S_r$ and Γ_T is closed with respect to the permutations $\hat{\mathcal{P}}_{\hat{\pi}}$ for all $\hat{\pi} \in \bigoplus_{r=0}^\infty S_r$, i.e. $\mathcal{P}_\pi \circ \Gamma_{T,r} = \Gamma_{T,r}$ and $\hat{\mathcal{P}}_{\hat{\pi}} \circ \Gamma_T = \Gamma_T$.*

Proof.

$$\|(B\cdot C)(\mathcal{E})\|_\infty = \sup_{\|\delta\mathcal{E}_j\|\neq 0}\ \operatorname*{ess\,sup}_{t\in[0,T]} \frac{\|(B(\mathcal{E})(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t))(C(\mathcal{E})(\delta\mathcal{E}_{r+1}, \delta\mathcal{E}_{r+q})(t))\|}{\|\delta\mathcal{E}_1\|\cdots\|\delta\mathcal{E}_{r+q}\|} \tag{A.25}$$

$$\leqslant \sup_{\|\delta\mathcal{E}_j\|\neq 0}\ \operatorname*{ess\,sup}_{t\in[0,T]} \frac{\|B(\mathcal{E})(\delta\mathcal{E}_1, \ldots, \delta\mathcal{E}_r)(t)\|}{\|\delta\mathcal{E}_1\|\cdots\|\delta\mathcal{E}_r\|}\frac{\|C(\mathcal{E})(\delta\mathcal{E}_{r+1}, \delta\mathcal{E}_{r+q})(t)\|}{\|\delta\mathcal{E}_{r+1}\|\cdots\|\delta\mathcal{E}_{r+q}\|} \tag{A.26}$$

$$\leqslant \|B(\mathcal{E})\|_\infty\|C(\mathcal{E})\|_\infty < \infty, \tag{A.27}$$

so $(B\cdot C)(\mathbb{K})\subset\mathcal{B}^{r+q}(L^\infty([0,T];\mathbb{C}^{N\times N}))$. For continuity with respect to the L^∞ topology, it may be shown that

$$\begin{aligned}\|(B\cdot C)(\mathcal{E}+\delta\mathcal{E})-(B\cdot C)(\mathcal{E})\|_\infty &\leqslant \|B(\mathcal{E}+\delta\mathcal{E})-B(\mathcal{E})\|_\infty\|C(\mathcal{E}+\delta\mathcal{E})-C(\mathcal{E})\|_\infty\\ &\quad+\|B(\mathcal{E}+\delta\mathcal{E})-B(\mathcal{E})\|_\infty\|C(\mathcal{E})\|_\infty\\ &\quad+\|B(\mathcal{E})\|_\infty\|C(\mathcal{E}+\delta\mathcal{E})-C(\mathcal{E})\|_\infty \qquad (A.28)\\ &\to 0 \qquad \text{as} \quad \|\delta\mathcal{E}\|_2\to 0 \qquad (A.29)\end{aligned}$$

because by assumption, $\|B(\mathcal{E})\|_\infty$ and $\|C(\mathcal{E})\|_\infty$ are both bounded and B and C are both continuous with respect to the L^∞ topology on their codomains. Hence $B\cdot C$ is continous with respect to the L^∞ topology on its codomain.

For the differentiability of $B\cdot C$, observe that

$$\begin{aligned}&\frac{\|(B\cdot C)(\mathcal{E}+\delta\mathcal{E})-(B\cdot C)(\mathcal{E})-(\mathrm{d}_\mathcal{E}B(\delta\mathcal{E}))\cdot C(\mathcal{E})-B(\mathcal{E})\cdot(\mathrm{d}_\mathcal{E}C(\delta\mathcal{E}))\|_2}{\|\delta\mathcal{E}\|_2}\\ &\quad\leqslant\frac{1}{\|\delta\mathcal{E}\|_2}(\|B(\mathcal{E}+\delta\mathcal{E})-B(\mathcal{E})-\mathrm{d}_\mathcal{E}B(\delta\mathcal{E})\|_2\|C(\mathcal{E})\|_\infty\\ &\qquad+\|B(\mathcal{E})\|_\infty\|C(\mathcal{E}+\delta\mathcal{E})-C(\mathcal{E})-\mathrm{d}_\mathcal{E}C(\delta\mathcal{E})\|_2\\ &\qquad+\|B(\mathcal{E}+\delta\mathcal{E})-B(\mathcal{E})-\mathrm{d}_\mathcal{E}B(\delta\mathcal{E})\|_2\|C(\mathcal{E}+\delta\mathcal{E})-C(\mathcal{E})\|_\infty\\ &\qquad+\|\mathrm{d}_\mathcal{E}B(\delta\mathcal{E})\|_2\|C(\mathcal{E}+\delta\mathcal{E})-C(\mathcal{E})\|_\infty) \qquad (A.30)\\ &\quad\to 0 \qquad \text{as}\quad \|\delta\mathcal{E}\|_2\to 0 \qquad (A.31)\end{aligned}$$

since again by assumption, $\|B(\mathcal{E})\|_\infty$ and $\|C(\mathcal{E})\|_\infty$ are both finite and B and C are both continuous with respect to the L^∞ topology on their codomains, and differentiable with respect to the L^2 topology on their codomains. Hence $B\cdot C$ is differentiable with respect to the L^2 topology on its codomain, and we conclude that $B\cdot C\in\Gamma_T$.

Turning to closure with respect to time integration, let $B\in\Gamma_{T,r}$ for any r. Since both $\hat{\mathcal{I}}_T$ and $\mathcal{I}_T$ are bounded (hence continuous) on $\mathcal{B}^r(\mathbb{K};L^\infty([0,T];\mathbb{C}^{N\times N}))$ and $B(\mathbb{K})\subset\mathcal{B}^r(\mathbb{K};L^\infty([0,T];\mathbb{C}^{N\times N}))$, both $\hat{\mathcal{I}}_T\circ B$ and $\mathcal{I}_T\circ B$ map into $\mathcal{B}^r(\mathbb{K};L^\infty([0,T];\mathbb{C}^{N\times N}))$. Since B is continuous in the L^∞ topology, $\hat{\mathcal{I}}_T\circ B$ and $\mathcal{I}_T\circ B$ are continuous in the L^∞ topology. Finally, since $\hat{\mathcal{I}}_T$ and $\mathcal{I}_T$ are linear and bounded (hence differentiable) on $\mathcal{B}^r(\mathbb{K};L^2([0,T];\mathbb{C}^{N\times N}))$ and B is differentiable in the L^2 topology, $\hat{\mathcal{I}}_T\circ B$ and $\mathcal{I}_T\circ B$ are differentiable in the L^2 topology. Hence $\hat{\mathcal{I}}_T\circ B\in\Gamma_{T,r},\mathcal{I}_T\circ B\in\Gamma_{T,r}$ and $\Gamma_{T,r}$ is closed with respect to both definite and indefinite time integration. Both $\hat{\mathcal{I}}_T$ and $\mathcal{I}_T$ may be extended by linearity over all of Γ_T and Γ_T is closed with respect to both definite and indefinite time integration.

Finally, since $\mathcal{P}_\pi$ is a bounded linear operator on both $\mathcal{B}^r(\mathbb{K};L^2([0,T];\mathbb{C}^{N\times N}))$ and $\mathcal{B}^r(\mathbb{K};L^\infty([0,T];\mathbb{C}^{N\times N}))$, the arguments used above for $\mathcal{I}_T$ and $\hat{\mathcal{I}}_T$ apply here as well to conclude that $\Gamma_{T,r}$ is closed with respect to permutations. It follows trivially that Γ_T is closed with respect to the permutations $\hat{\mathcal{P}}_{\hat\pi}$ for all $\hat\pi\in\bigoplus_{r=0}^\infty S_r$. □

Lemma 5. *Fix $T>0$ and let $\hat{V}_T:\mathbb{K}\to L^2([0,T];\mathbb{C}^{N\times N})$ be defined by $\hat{V}_T(\mathcal{E})(t)=V_t(\mathcal{E})$ for all $\mathcal{E}\in\mathbb{K}$ and all $t\in[0,T]$. Then $\hat{V}_T\in\Gamma_{T,0}$ and $\mathrm{d}_\mathcal{E}\hat{V}_T(\delta\mathcal{E})(t)=\mathrm{d}_\mathcal{E}V_t(\delta\mathcal{E})$.*

Proof. Since $\|\hat{V}_T(\mathcal{E})(t)\|=\sqrt{N}$ for all $\mathcal{E}\in\mathbb{K}$ and all $t\in[0,T]$, it is clear that $\hat{V}_T(\mathbb{K})\subset L^\infty([0,T];\mathbb{C}^{N\times N})$. In addition, it has been shown [1, 21] that $\hat{V}_T$ is continuous with respect to the L^∞ topology on its codomain. Finally let $\mathrm{d}_\mathcal{E}\hat{V}_T(\delta\mathcal{E})(t)=\mathrm{d}_\mathcal{E}V_t(\delta\mathcal{E})$. Then, using (A.15),

$$\frac{\|\hat{V}_T(\mathcal{E}+\delta\mathcal{E})-\hat{V}_T(\mathcal{E})-\mathrm{d}_{\mathcal{E}}\hat{V}_T(\delta\mathcal{E})\|_2}{\|\delta\mathcal{E}\|_2}=\frac{\left(\int_0^T\|V_t(\mathcal{E}+\delta\mathcal{E})-V_t(\mathcal{E})-\mathrm{d}_{\mathcal{E}}V_t(\delta\mathcal{E})\|^2\,\mathrm{d}t\right)^{\frac{1}{2}}}{\|\delta\mathcal{E}\|_2} \tag{A.32}$$

$$\leqslant T\frac{A(T,\mathcal{E},\delta\mathcal{E})}{\|\delta\mathcal{E}\|_2}\exp\left(\frac{1}{\hbar}\int_0^T\|\mathcal{H}_t(\mathcal{E})\|\,\mathrm{d}t\right) \tag{A.33}$$

$$\to 0 \qquad \text{as} \quad \|\delta\mathcal{E}\|_2\to 0 \tag{A.34}$$

as in (A.19). It is easily verified that $\mathrm{d}_{\mathcal{E}}\hat{V}_T$, defined in this way, is a bounded operator. Therefore $\hat{V}_T$ is differentiable with Fréchet derivative $\mathrm{d}_{\mathcal{E}}\hat{V}_T(\delta\mathcal{E})(t)=\mathrm{d}_{\mathcal{E}}V_t(\delta\mathcal{E})$, and $\hat{V}_T\in\Gamma_{T,0}$. □

Lemma 6. *Fix $T>0$ and $\delta\mathcal{E}\in\mathbb{K}$. Let $R:\mathbb{K}\to\mathcal{B}^1(\mathbb{K};L^2([0,T];\mathbb{C}^{N\times N}))$ be given by*

$$R(\mathcal{E})(\delta\mathcal{E}_1)(t):=\frac{\mathrm{i}}{\hbar}V_t^\dagger(\mathcal{E})\mu V_t(\mathcal{E})\delta\mathcal{E}_1(t) \qquad \forall\, t\in[0,T]. \tag{A.35}$$

Then $R\in\Gamma_{T,1}$ and $\mathrm{d}R=\mathcal{P}_\pi\circ(R\cdot(\tilde{\mathcal{I}}_T\circ R))-(\tilde{\mathcal{I}}_T\circ R)\cdot R$, where $\pi\in S_2$ is the transposition $(12)\mapsto(21)$. So the differential $\mathrm{d}R$ lies in $\Gamma_{T,2}$.

Proof. First, observe that $\Upsilon:\mathbb{K}\to\mathcal{B}^1(\mathbb{K};L^2([0,T];\mathbb{C}^{N\times N}))$ given by $\Upsilon(\mathcal{E})(\delta\mathcal{E}_1)=(\mathrm{i}/\hbar)\mu\delta\mathcal{E}$, is trivially a member of $\Gamma_{T,1}$. Also, for any $B\in\Gamma_T$, $B^\dagger$ is also a member of Γ_T, since the matrix adjoint operation (†) is linear (because $\mathbb{C}^{N\times N}$ is taken to be a vector space over $\mathbb{R}$) and bounded in both the L^2 and L^∞ topologies. In addition, $R=\hat{V}_T^\dagger\cdot\Upsilon\cdot\hat{V}_T$ is a product of elements of Γ_T and therefore itself lies in Γ_T. In particular, it lies in $\Gamma_{T,1}$.

Furthermore, by the product rule on Γ_T and the fact that $\mathrm{d}\Upsilon=0$,

$$\mathrm{d}R=\left(\mathrm{d}\hat{V}_T^\dagger\right)\cdot\Upsilon\cdot\hat{V}_T+\mathcal{P}_{(21)}\circ\left[\hat{V}_T^\dagger\cdot\Upsilon\cdot(\mathrm{d}\hat{V}_T)\right] \tag{A.36}$$

$$=\left[-(\hat{\mathcal{I}}_T\circ R)\cdot\hat{V}_T^\dagger\right]\cdot\Upsilon\cdot\hat{V}_T+\mathcal{P}_{(21)}\circ\left(\hat{V}_T^\dagger\cdot\Upsilon\cdot[\hat{V}_T\cdot(\hat{\mathcal{I}}_T\circ R)]\right) \tag{A.37}$$

$$=-(\hat{\mathcal{I}}_T\circ R)\cdot R+\mathcal{P}_{(21)}\circ[R\cdot(\hat{\mathcal{I}}_T\circ R)]. \tag{A.38}$$

□

Definition 6. *Let $\mathcal{A}_0$ be the subalgebra of Γ_T generated by R and closed under permutation. For $k=1,2,\ldots$, let $\mathcal{A}_k$ be the subalgebra of Γ_T generated by $\mathcal{A}_{k-1}\bigcup(\mathcal{I}_T\circ\mathcal{A}_{k-1})\bigcup(\hat{\mathcal{I}}_T\circ\mathcal{A}_{k-1})$ and closed under permutation. These $\{\mathcal{A}_k\}$ form a nested sequence of algebras. Let $\Lambda_T=\bigcup\mathcal{A}_k$.*

Lemma 7. *For each $B\in\mathcal{A}_k$ the Fréchet derivative $\mathrm{d}B$ lies in $\mathcal{A}_{k+1}$. Hence Λ_T is a subalgebra of Γ_T, closed with respect to $\mathcal{I}_T,\hat{\mathcal{I}}_T$, permutations $\hat{\mathcal{P}}_{\hat{\pi}}$, and differentiation by the control $\mathcal{E}$. As a consequence, each element of Λ_T is infinitely Fréchet differentiable.*

Proof. Note first that any B in $\mathcal{A}_0$ is a finite linear combination of permutations of finite powers of R. The Fréchet differential of B will then be a finite linear combination of terms like

$$\sum_{j=1}^J\mathcal{P}_{\pi_j}\circ(R^{j-1}\cdot\mathrm{d}R\cdot R^{J-j})=\sum_{j=1}^J\mathcal{P}_{\pi_j}\circ(R\cdots R\cdot\mathrm{d}R\cdot R\cdots R). \tag{A.39}$$

By lemma 6, $\mathrm{d}R=\mathcal{P}_\pi\circ(R\cdot(\tilde{\mathcal{I}}_T\circ R))-(\tilde{\mathcal{I}}_T\circ R)\cdot R$ which lies in $\mathcal{A}_1$. Since R also lies in $\mathcal{A}_1$, $\mathrm{d}B$ is a finite linear combination of permutations of finite products of terms in $\mathcal{A}_1$ and therefore itself lies in $\mathcal{A}_1$.

Now suppose for each $B\in\mathcal{A}_{k-1}$ that $\mathrm{d}B$ lies in $\mathcal{A}_k$. Let C be an arbitrary element of $\mathcal{A}_k$. As such, it is a finite linear combination of permutations of finite products of the generators in $\mathcal{A}_{k-1} \cup \mathcal{I}_T \circ \mathcal{A}_{k-1} \cup \hat{\mathcal{I}}_T \circ \mathcal{A}_{k-1}$. The Fréchet derivatives of any element in $\mathcal{A}_{k-1}$ lies in $\mathcal{A}_k$ by assumption, and the derivatives of any element in $\mathcal{I}_T \circ \mathcal{A}_{k-1}$ or $\hat{\mathcal{I}}_T \circ \mathcal{A}_{k-1}$ will lie in $\mathcal{I}_T \circ \mathcal{A}_k$ or $\hat{\mathcal{I}}_T \circ \mathcal{A}_k$, respectively. Hence the Fréchet derivative of any of these generators will lie in $\mathcal{A}_{k+1}$, as do all of the generators themselves. So the derivative of C will be a finite linear combination of finite products of elements of $\mathcal{A}_{k+1}$, and therefore the derivative itself lies in $\mathcal{A}_{k+1}$. The lemma follows by induction. □

Theorem 2. *If $\hat{\mathcal{H}}$ is given by $\hat{\mathcal{H}}(\mathcal{E}) = H_0 - \mu\mathcal{E}$ for some fixed $N \times N$ Hermitian H_0 and μ, then V_T is infinitely Fréchet differentiable (i.e. C^∞).*

Proof. The differential of V_T is given by $\mathrm{d}V_T = V_T \cdot (\mathcal{I}_T \circ R)$. Given any map $F = V_T \cdot (\mathcal{I}_T \circ B)$ for any $B \in \Lambda_T$, F is differentiable with Fréchet derivative $\mathrm{d}F = \hat{\mathcal{P}}_{\hat{\pi}} \circ [\mathrm{d}V_T \cdot (\mathcal{I}_T \circ B)] + V_T \cdot (\mathcal{I}_T \circ \mathrm{d}B)$ for some permutation $\hat{\mathcal{P}}_{\hat{\pi}}$. Permutation commutes with both integration and with V_T (since $V_T \in \Gamma_{T,0}$), hence the differential of F may be written $\mathrm{d}F = V_T \cdot (\mathcal{I}_T \circ C)$ where $C = \mathcal{P}_{\pi_1} \circ [(\mathcal{I}_T \circ R) \cdot B] + \mathrm{d}B$ is again an element of Λ_T. Therefore, each derivative of F is of the form $V_T \cdot (\mathcal{I}_T \circ C)$ for some $C \in \Lambda_T$, and so each such F is infinitely Fréchet differentiable. In particular, $\mathrm{d}V_T = V_T \cdot (\mathcal{I}_T \circ R)$ is of this form since $R \in \Lambda_T$. Hence, V_T is infinitely Fréchet differentiable. □

References

[1] Jurdjevic V and Sussmann H J 1972 Control systems on Lie groups *J. Diff. Eqns.* **12** 313–29
[2] Ramakrishna V, Salapaka M V, Dahleh M, Rabitz H and Pierce A 1995 Controllability of molecular systems *Phys. Rev.* A **51** 960–966
[3] Khaneja N, Brockett R and Glaser S J 2001 Time optimal control in spin systems *Phys. Rev.* A **63** 032308
[4] Vidal G, Hammerer K and Cirac J I 2002 Interaction cost of nonlocal gates *Phys. Rev. Lett.* **88** 237902
[5] Hammerer K, Vidal G and Cirac J I 2002 Characterization of nonlocal gates *Phys. Rev.* A **66** 062321
[6] Yuan H and Khaneja N 2005 Time optimal control of coupled qubits under nonstationary interactions *Phys. Rev.* A **72** 040301
[7] Dirr G, Helmke U, Hüper K, Kleinsteuber M and Liu Y 2006 Spin dynamics: a paradigm for time optimal control on compact Lie groups *J. Glob. Optim.* **35** 443–74
[8] Carlini A, Hosoya A, Koike T and Okudaira Y 2006 Time-optimal quantum evolution *Phys. Rev. Lett.* **96** 06050
[9] Carlini A, Hosoya A, Koike T and Okudaira Y 2007 Time-optimal unitary operations *Phys. Rev.* A **75** 042308
[10] Khaneja N, Heitmann B, Sporl A, Yuan H, Schulte-Herbruggen T and Glaser S J 2007 Shortest paths for efficient control of indirectly coupled qubits *Phys. Rev.* A **75** 012322
[11] Rothman A, Tak-San H and Rabitz H 2005 Observable-preserving control of quantum dynamics over a family of related systems *Phys. Rev.* A **72** 023416
[12] Rothman A, Tak-San H and Rabitz H 2005 Quantum observable homotopy tracking control *J. Chem. Phys.* **123** 134104
[13] Rothman A, Tak-San H and Rabitz H 2006 Exploring the level sets of quantum control landscapes *Phys. Rev.* A **73** 053401
[14] Rabitz H A, Hsieh M M and Rosenthal C M 2004 Quantum optimally controlled transition landscapes *Science* **303** 1998–2001
[15] Abraham R, Marsden J E and Ratiu T 1988 *Manifolds, Tensor Analysis, and Applications* 2nd edn (New York: Springer)
[16] do Carmo M P 1992 *Riemannian Geometry* (Boston: Birkhäuser)
[17] Grace M, Brif C, Rabitz H, Walmsley I A, Kosut R L and Lidar D A 2007 Optimal control of quantum gates and suppression of decoherence in a system of interacting two-level particles *J. Phys. B: At. Mol. Opt. Phys.* **40** S103–25
[18] Mezzadri F 2007 How to generate random matrices from the classical compact groups *Not. Am. Math. Soc.* **54** 592–604
[19] Edelman A and Rao N R 2005 Random matrix theory *Acta Numer.* **14** 1–65
[20] Coddington E A and Levinson N 1984 *Theory of Ordinary Differential Equations* (Krieger: Malabar, FL)
[21] Sussmann H J 1972 The 'bang-bang' problem for certain control systems in $gl(n, r)$ *SIAM J. Control* **10** 470–6