Quantum Splines

Dorje C. Brody,[1] Darryl D. Holm,[2] and David M. Meier[2]
[1]Mathematical Sciences, Brunel University, Uxbridge UB8 3PH, United Kingdom
[2]Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom
(Received 20 June 2012; published 4 September 2012)

A quantum spline is a smooth curve parametrized by time in the space of unitary transformations, whose associated orbit on the space of pure states traverses a designated set of quantum states at designated times, such that the trace norm of the time rate of change of the associated Hamiltonian is minimized. The solution to the quantum spline problem is obtained, and is applied in an example that illustrates quantum control of coherent states. An efficient numerical scheme for computing quantum splines is discussed and implemented in the examples.

DOI: 10.1103/PhysRevLett.109.100501 PACS numbers: 03.67.Ac, 02.30.Xx, 02.60.Ed, 42.50.Dv

Controlling the evolution of the unitary transformations that generate quantum dynamics is vital in quantum information processing. There is a substantial volume of literature devoted to the investigation of the many aspects of quantum control [1]. The objective of quantum control is the unitary transformation of one quantum state, pure or mixed, into another one, subject to certain criteria. For example, one may wish to transform a given quantum state $|\psi\rangle$ into another state $|\phi\rangle$ unitarily in the shortest possible time, with finite energy resource [2–4]. When only the initial and final states are involved, many time-independent Hamiltonians are available that achieve the unitary evolution $|\psi\rangle \to |\phi\rangle$, and we simply need to find one that is optimal. However, transforming a given quantum state $|\psi\rangle$ along a path that traverses through a sequence of designated quantum states $|\psi\rangle \to |\phi_1\rangle \to |\phi_2\rangle \to \cdots \to |\phi_m\rangle$ cannot be achieved by a time-independent Hamiltonian. To realize this chain of transformations in the shortest possible time, one chooses the optimal Hamiltonian H_j for each interval $|\phi_j\rangle \to |\phi_{j+1}\rangle$ [3,4], and switches the Hamiltonian from H_j to H_{j+1} when the state has reached $|\phi_{j+1}\rangle$. However, instantaneous switching of the Hamiltonian is in general not experimentally feasible.

In the present Letter, we consider the following quantum control problem: Let a set of quantum states $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_m\rangle$ and a set of times $t_1, t_2, \ldots, t_m$ be given. Starting from an initial state $|\psi_0\rangle$ at time $t_0 = 0$, find a time-dependent Hamiltonian $H(t)$ such that the evolution path $|\psi_t\rangle$ passes arbitrarily close to $|\phi_j\rangle$ at time $t = t_j$ for all $j = 1, \ldots, m$, and such that the change in the Hamiltonian, in a sense defined below, is minimized. The solution to this problem will generate a continuous curve in the space of quantum states that interpolates through the designated states, just as a spline curve interpolates through a given set of data points. We thus refer to this solution as a quantum spline.

There is a difference between a classical spline curve and a quantum spline. In the classical context the solution curve passes through a given set of points, whereas in the quantum context, a curve on the space of pure states in itself has no operational meaning. Thus, instead of finding a curve in the space of pure states where the designated states lie, we must find a time-dependent curve in the space of Hamiltonians that in turn will generate the curve in the unitary transformation group needed to produce an optimal trajectory. In other words, we shall seek a curve in the associated Lie algebra, which of course is equivalent to the space of Hamiltonians, up to multiplication by $i = \sqrt{-1}$.

Our approach involves variational calculus in the Lie algebra of skew-Hermitian matrices, with constraints that take values in the unitary group. In addition, since our optimality condition for quantum splines involves the time derivative of $iH(t)$, we shall make use of the techniques developed recently for the higher-order calculus of variations on Lie groups and their algebras [5,6]. By extending these results we are able to (a) derive the Euler-Lagrange equations (5) and (9) below that solve quantum spline problems, and (b) devise an efficient discretization scheme to numerically implement the solution. An example of such a solution for a two-level quantum system is sketched in Fig. 1. As an application, we illustrate how the results transform a quantum state along a path that lies entirely on the coherent-state subspace.

The optimal curve $H(t)$ that solves the quantum spline problem is the minimizer of a "cost functional" (action) consisting of two terms: The first term measures the overall change in the Hamiltonian during the evolution. For this purpose we shall consider the trace norm; i.e., for a pair of trace-free skew-Hermitian matrices A and B we define their inner product by

$$\langle A, B\rangle = -2\,\mathrm{tr}(AB), \tag{1}$$

where the factor -2 is purely conventional. Thus, if H is a time-dependent Hamiltonian and $\dot{H}$ its time derivative, the instantaneous penalty arising from changing the Hamiltonian is given by $\frac{1}{2}\langle i\dot{H}, i\dot{H}\rangle = \mathrm{tr}(\dot{H}^2)$. The second term penalizes the "mismatch" between the state $|\psi_{t_j}\rangle$ at

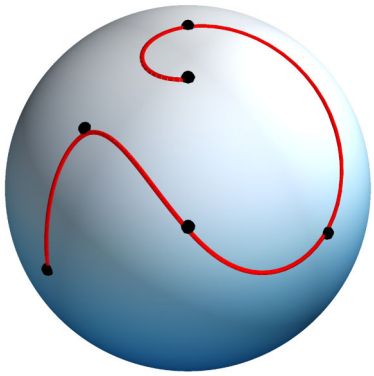


FIG. 1 (color online). A quantum spline for a two-level system. The lower-left initial state and the targets are represented by black dots. The variational formulation of the problem requires us to minimize a functional that measures both the cost related to the change of the Hamiltonian, and the amount of mismatch between the trajectory and the target points.

time t_j and the target state $|\phi_j\rangle$. For this purpose we shall use the standard geodesic distance

$$D(\psi, \phi) = 2\arccos\sqrt{\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}} \tag{2}$$

for a pair of states $|\psi\rangle$ and $|\phi\rangle$. Writing $U(t)$ for the parametric family of unitary operators generated by $H(t)$ so that $|\psi_{t_j}\rangle = U(t_j)|\psi_0\rangle$, the mismatch penalty is chosen to be $\frac{1}{2}D^2(U(t_j)\psi_0, \phi_j)/\sigma^2$, where the tolerance $\sigma > 0$ is a tunable parameter so that the penalty is high when σ is small, and the factor of a half is purely conventional.

The action, of course, must be minimized subject to the constraint that the dynamical evolution of the state is unitary. That is, U must satisfy the Schrödinger equation $\dot{U} = -iHU$, in units $\hbar = 1$. Therefore, given an initial state $|\psi_0\rangle$ at time $t_0 = 0$, a set of target states $|\phi_1\rangle, \ldots, |\phi_m\rangle$ at times $t_1, \ldots, t_m$, and an initial Hamiltonian $H(0) = H_0$, we wish to find the minimizer of

$$\mathcal{J} = \int_{t_0}^{t_m}\Big(\frac{1}{2}\langle i\dot{H}, i\dot{H}\rangle + \langle M, \dot{U}U^{-1} + iH\rangle\Big)dt + \frac{1}{2\sigma^2}\sum_{j=1}^{m} D^2(U(t_j)\psi_0, \phi_j), \tag{3}$$

where the minimization is over curves $U(t) \in SU(n+1)$ and $iH(t), M(t) \in \mathfrak{su}(n+1)$. Additionally, we require smoothness of these curves on open intervals (t_j, t_{j+1}) for $j = 0, \ldots, m-1$; $U(0) = \mathbb{1}$, and the continuity of $U(t)$ and $H(t)$ is assumed everywhere. The curve $M(t)$ acts as a Lagrange multiplier enforcing the kinematic constraint.

Before we proceed to vary the action $\mathcal{J}$ let us comment on the choice of the initial Hamiltonian H_0. We let H_0 be such that the trajectory $e^{-iH_0t}|\psi_0\rangle$ corresponds to the geodesic curve on the space of pure states joining $|\psi_0\rangle$ and $|\phi_1\rangle$; the construction of such a Hamiltonian can be found in [4]. Intuitively, since the first target time t_1 is fixed, this choice generates the most direct traverse $|\psi_0\rangle \to |\phi_1\rangle$, hence requiring least change in the Hamiltonian at initial times $t \ll t_1$.

The Euler-Lagrange equations governing stationary points of (3) are obtained by taking the variation of $\mathcal{J}$ and requiring $\delta\mathcal{J} = 0$. Writing $A = (\delta U)U^{-1}$ we have

$$\begin{aligned}\delta\mathcal{J} &= \int_{t_0}^{t_m}(\langle i\dot{H}, i\delta\dot{H}\rangle + \langle M, \dot{A} - [\dot{U}U^{-1}, A] + i\delta H\rangle + \langle \delta M, \dot{U}U^{-1} + iH\rangle)dt + \frac{1}{2\sigma^2}\sum_{j=1}^{m}\delta D^2(\psi_{t_j}, \phi_j)\\ &= \int_{t_0}^{t_m}(\langle M - i\ddot{H}, i\delta H\rangle + \langle -\dot{M} + [\dot{U}U^{-1}, M], A\rangle + \langle \delta M, \dot{U}U^{-1} + iH\rangle)dt + \frac{1}{2\sigma^2}\sum_{j=1}^{m}\delta D^2(\psi_{t_j}, \phi_j)\\ &\quad + \sum_{j=1}^{m-1}[\langle \Delta M(t_j), A(t_j)\rangle + \langle i\Delta\dot{H}(t_j), i\delta H(t_j)\rangle] + \langle M(t_m), A(t_m)\rangle + \langle i\dot{H}(t_m), i\delta H(t_m)\rangle,\end{aligned} \tag{4}$$

where in the second step we have integrated by parts, and used the notations $\Delta M(t_j) = M(t_j^-) - M(t_j^+)$ and $\Delta\dot{H}(t_j) = \dot{H}(t_j^-) - \dot{H}(t_j^+)$, with $M(t_i^+) = \lim_{t\downarrow t_i} M(t)$ and $M(t_i^-) = \lim_{t\uparrow t_i} M(t)$, and similarly for $\dot{H}(t_j^\pm)$. It follows from (4) that on the open intervals (t_j, t_{j+1}), $j = 0, \ldots, m-1$, the following equations hold:

$$i\ddot{H} - M = 0, \qquad \dot{M} + [M, \dot{U}U^{-1}] = 0, \qquad \dot{U}U^{-1} + iH = 0. \tag{5}$$

Additionally, at the nodes $t = t_j$, we require matching conditions. To work them out, let us calculate the variation $\delta D^2 = 2D\delta D$ appearing in (4). From the definition (2) and the relation

$$\begin{aligned}&\frac{\langle\psi|e^{-\varepsilon A}|\phi\rangle\langle\phi|e^{\varepsilon A}|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}\\ &\quad\approx \frac{\langle\psi|(1-\varepsilon A)|\phi\rangle\langle\phi|(1+\varepsilon A)|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}\\ &\quad= \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} + \frac{2\Re[\langle\psi|\phi\rangle\langle\phi|A|\psi\rangle]}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle}\varepsilon + \mathcal{O}(\varepsilon^2),\end{aligned} \tag{6}$$

which holds for any $A = -A^\dagger$, we find, bearing in mind that if $D = 2\arccos(\sqrt{x})$ then $dD/dx = -2/\sin(D)$,

$$\delta D = \frac{d}{d\varepsilon} D(e^{\varepsilon A}\psi, \phi)\Big|_{\varepsilon=0} = \frac{-4\Re[\langle\psi|\phi\rangle\langle\phi|A|\psi\rangle]}{\sin(D)\langle\phi|\phi\rangle\langle\psi|\psi\rangle}. \tag{7}$$

From (7), and writing $D_j = D(\psi_{t_j}, \phi_j)$, we deduce that $\delta D_j^2 = 2D_j\langle F_j, A(t_j)\rangle$, where

$$F_j = \frac{\langle\psi_{t_j}|\phi_j\rangle|\psi_{t_j}\rangle\langle\phi_j| - \langle\phi_j|\psi_{t_j}\rangle|\phi_j\rangle\langle\psi_{t_j}|}{\sin(D_j)\langle\phi_j|\phi_j\rangle\langle\psi_{t_j}|\psi_{t_j}\rangle}. \tag{8}$$

The relevant matching conditions at the nodes are therefore given by:

$$\dot{H}(t_j^+) - \dot{H}(t_j^-) = 0, \quad M(t_j^+) - M(t_j^-) = D_j F_j/\sigma^2, \tag{9}$$

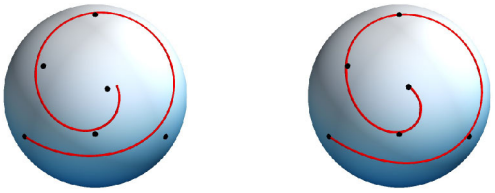
whereas we require $\dot{H}(t_m) = 0$ and $M(t_m) + D_m F_m/\sigma^2 = 0$ at the terminal point. Quantum spline problems are therefore solved by finding a solution to Eqs. (5) and (9) that satisfies, in addition, the terminal conditions at t_m. On open time intervals (t_i, t_{i+1}) equation (5) yields

$$\ddot{H} + i[H, \ddot{H}] = 0. \tag{10}$$

This is the right-reduced equation for the so-called Riemannian cubics on $SU(n+1)$ with respect to the bi-invariant Riemannian metric induced by the inner product (1). That is, $U(t)$ is a Riemannian cubic on the open time intervals (t_i, t_{i+1}). Here, by a Riemannian cubic we mean a solution to a certain fourth-order equation for a curve on a Riemannian manifold (see Ref. [7] for further details). The node conditions (9) imply that $U(t)$ is a Riemannian cubic spline, a twice continuously differentiable curve that is composed of a series of cubics.

We remark on the important structure of the Lagrange multiplier $M(t)$ implied by the equations of motion that makes it sufficient to consider a subspace of $\mathfrak{su}(n+1)$ when searching for the optimal initial value $M(0)$. Let us denote by $\mathcal{P}_\psi$ the totality of tracefree skew-Hermitian generators of unitary motions that leave the state $|\psi\rangle$ invariant, and $\mathcal{P}_\psi^\perp$ its complement with respect to the inner product (1). Then, we have the following Lemma: $M(t) \in \mathcal{P}_{\psi_t}^\perp$ [this holds because for all j, $D_j F_j \in \mathcal{P}_{\psi_{t_j}}^\perp$, and from (5), $M(t)_{t\in(t_j,t_{j+1})} = \mathrm{Ad}_{U(t)U(t_{j+1})^{-1}} M(t_{j+1}^-)$]. This result is significant, because the search for the optimal $M(0)$ can be restricted to the $2n$-dimensional subspace $\mathcal{P}_{\psi_0}^\perp$ of the $n(n+2)$-dimensional Lie algebra $\mathfrak{su}(n+1)$.

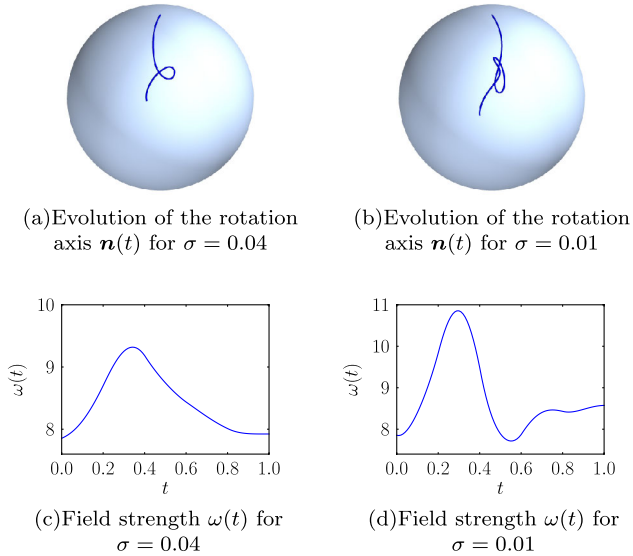
Before we indicate the process for the implementation of the optimization scheme, let us show some results first. Consider a two-level system ($n = 1$). We can think of this system as a spin-$\frac{1}{2}$ particle immersed in a magnetic field. If $\boldsymbol{n}(t)$ is the unit direction of the field at time t, the Hamiltonian of the system can be written in the form



(a)$\sigma = 0.04$ (b)$\sigma = 0.01$

FIG. 2 (color online). Orbits on the state space generated by the solution to the quantum spline problem. The black dots indicate the initial (lower left) and the target points. The optimal trajectories are shown for two different values of the tolerance parameter: $\sigma = 0.04$ and $\sigma = 0.01$. Lower values of the tolerance parameter translate, through the cost functional $\mathcal{J}$, into a stronger penalty on the mismatch.

$H(t) = \omega(t)\boldsymbol{\sigma}\cdot\boldsymbol{n}(t)$, where $\omega(t)$ is the field strength. In this case the state space is just the Bloch sphere S^2. We have implemented the optimization for a set of target states on S^2, an initial state $|\psi_0\rangle$, and a set of times. Using the resulting Hamiltonian we have generated the dynamics of the state, as illustrated in Fig. 1. In Fig. 2 we have sketched the effect of choosing different tolerance levels. When the value of σ is reduced, the resulting orbit $|\psi_t\rangle$ traverses closer to the vicinities of the target states $\{|\phi_j\rangle\}$. From (3), one sees that this may be realized at the expense of varying the Hamiltonian $H(t)$ more rapidly. This effect can be visualized in the case of a two-level system, since $H(t)$ is characterized by the the scalar field strength $\omega(t)$ and the unit vector $\boldsymbol{n}(t) \in \mathbb{R}^3$. In Fig. 3 we have plotted the



(a)Evolution of the rotation axis $\boldsymbol{n}(t)$ for $\sigma = 0.04$ (b)Evolution of the rotation axis $\boldsymbol{n}(t)$ for $\sigma = 0.01$

(c)Field strength $\omega(t)$ for $\sigma = 0.04$ (d)Field strength $\omega(t)$ for $\sigma = 0.01$

FIG. 3 (color online). The quantum spline $H(t)$. Hamiltonians that generate the dynamical trajectories in Fig. 2. The top row shows the orbits of the endpoint of the rotation axis $\boldsymbol{n}(t)$. The bottom row shows the field strength $\omega(t)$. These images illustrate the fact that as the value of σ is decreased, the amount of change in the optimal Hamiltonian $H(t)$ increases.

endpoint of the unit vector $\boldsymbol{n}(t)$ on a sphere, and the values of $\omega(t)$, for different choices of σ. These plots show that both $\boldsymbol{n}(t)$ and $\omega(t)$ vary more rapidly at smaller tolerance level (i.e., smaller σ).

Another example we consider here is a controlled motion of a quantum state on the coherent-state subspace of the state space. Consider $SU(n+1)$ coherent states [8,9] in arbitrary dimensions. These coherent states can be generated by taking symmetric tensor products of "single-particle"states. In the context of quantum information theory, these states correspond to totally disentangled states inside the symmetric subspace of the Hilbert space of the combined system. Each coherent state thus corresponds to the image of a map, known as the Veronese embedding [10,11], of a pure state. Therefore, given a set of points on a coherent-state space we identify them with states on a single-particle Hilbert space, solve the quantum spline problem as indicated above, and map the result back to the larger Hilbert space. In particular, the coherent quantum spline is generated by the symmetric tensor product Hamiltonian $\otimes_s H(t)$. This elementary procedure works because (a) the Veronese embedding commutes with the action of $SU(n+1)$, and (b) the natural metrics on the spaces of coherent states are scalar multiples of the metric (2) used here [10].

Next we discuss a numerical approach for finding a local minimum of the cost functional (3). The search can be restricted to solutions of (5) and (9), which are encoded by their initial conditions $M(0)$ and $\dot{H}(0)$. We can therefore regard (3) as a function of these initial conditions, and perform a descent algorithm on that function. The terminal conditions at t_m can then be used to test whether we have arrived at a local minimum.

For a numerical implementation we can discretize the equations of motion (5) and (9), and find the approximate gradient of $\mathcal{J}$; alternatively, we can introduce an approximation $\mathcal{J}_d$ of $\mathcal{J}$ defined on a discrete path space, and take its variation, which yields a set of discrete equations of motion. Here we follow the latter method, which permits the use of adjoint equations [6] for an efficient calculation of the exact gradient of $\mathcal{J}_d$. This method is highly effective in dealing with higher-dimensional ($n > 1$) systems. Moreover, in this method discrete critical curves of $\mathcal{J}_d$ satisfy a version of the terminal conditions at $t = t_m$ *exactly*, and this leads to a precise method for testing convergence. In addition, such curves fulfil the conditions for the above-stated Lemma on their discrete time domain, which can be exploited by restricting the search for the optimal initial value of M to $\mathcal{P}^{\perp}_{\psi_0}$.

The implementation will make use of the Cayley map τ: $\mathfrak{su}(n+1) \to SU(n+1)$, which approximates the Lie exponential according to $X \mapsto (\mathbb{1} - X/2)^{-1}(\mathbb{1} + X/2)$. We will also need the left-trivialized differential d_l: $d_l\tau_X Y = (d/d\varepsilon)\tau(X + \varepsilon Y)\tau(X)^{-1}|_{\varepsilon=0}$, which is given by $(\mathbb{1} - X/2)^{-1} Y (\mathbb{1} + X/2)^{-1}$. We discretize the time interval $t_m - t_0$ into N steps such that $(t_m - t_0)/N = h$, and we let $t^\mu = t_0 + \mu h$ for $\mu = 0, \ldots, N$. For simplicity, we assume that the nodal times $\{t_j\}_{j=0,\ldots,m}$ coincide with some of the discrete time steps $t^{n_j} = t_0 + n_j h$, where $n_0 = 0$ and $n_m = N$. To obtain a discrete version of the cost functional, we approximate the time derivative $-i\dot{H}$ of the generator by the discrete variables $\{L_\mu\}$. The complete set of discrete variables is therefore $(U_\mu, iH_\mu, M_\mu, L_\mu)$, with $\mu = 0, \ldots, N$. Writing $\Delta_\mu = \delta_{\mu n_j} D_j F_j/\sigma^2$ and making use of the Euler method of [6], we obtain the following set of discrete equations of motion for $\mu = 0, \ldots, N-1$:

$$\begin{aligned} M_{\mu+1} &= (d_l\tau^{-1}_{ihH_{\mu+1}})(d_l\tau_{-ihH_{\mu+1}})(M_\mu + \Delta_\mu) \\ L_{\mu+1} &= L_\mu - h(d_l\tau_{ihH_{\mu+1}})M_{\mu+1} \\ U_{\mu+1} &= \tau(-ihH_{\mu+1})U_\mu, \\ H_{\mu+1} &= H_\mu + ihL_\mu. \end{aligned} \tag{11}$$

These equations can be integrated for given initial values M_0 and L_0 (recall that $U_0 = \mathbb{1}$ and H_0 are prescribed). The terminal conditions are $L_N = 0$ and $M_N + \Delta_N = 0$. The discrete cost functional $\mathcal{J}_d$ in terms of initial conditions (M_0, L_0) is

$$\mathcal{J}_d = \sum_{\mu=0}^{N-1} \frac{h}{2}\langle L_\mu, L_\mu\rangle + \frac{1}{2\sigma^2}\sum_{j=1}^{m} D^2(U_{n_j}\psi_0, \phi_j), \tag{12}$$

whereby the equations of motion (11) are implied.

A local minimum can be found by a gradient descent method, which requires the computation of the gradient of $\mathcal{J}_d$. The estimation of the gradient via finite-difference methods requires the repeated forward integration of the system of equations (11). The number of forward integrations increases with the number of dimensions of the Lie algebra (n^2 to leading order). Such estimation procedures thus quickly become unfeasible for higher-dimensional systems. This difficulty can be avoided by using the method of adjoint equations, which can be readily implemented for the discretization (11) and (12), presented here (see Supplemental Material [12] for details, and Ref. [13] for a numerical code). Then, the exact gradient is obtained at the cost of integrating twice (once forward, once backward) a system of equations of the same complexity as (11). This allows for an efficient treatment of the quantum spline problem when $n > 1$.

We thank C. Burnett and L. Noakes for fruitful and thoughtful discussions in the course of this work. D. D. H. and D. M. M. are also grateful for partial support by a European Research Council Advanced Grant.

[1] At the time of drafting this Letter there are 935 papers posted on the arXiv that contain the terms "quantum" and "control" in the title.

[2] D. C. Brody, J. Phys. A **36**, 5587 (2003).

[3] A Carlini, A. Hosoya, T. Koike, and Y. Okudaira, Phys. Rev. Lett. **96**, 060503 (2006).
[4] D. C. Brody and D. W. Hook, J. Phys. A **39**, L167 (2006); Corrigendum D. C. Brody and D. W. Hook, *ibid.* **40**, 10949 (2007).
[5] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, and F. X. Vialard, Commun. Math. Phys. **309**, 413 (2012).
[6] C. Burnett, D. D. Holm, and D. M. Meier, arXiv:1112.6037.
[7] L. Noakes, G. Heinzinger, and B. Paden, IMA J. Math. Control Inf. **6**, 465 (1989).
[8] A. M. Perelomov, Commun. Math. Phys. **26**, 222 (1972).
[9] R. Gilmore, Ann. Phys. (N.Y.) **74**, 391 (1972).
[10] D. C. Brody and E. M. Graefe, J. Phys. A **43**, 255205 (2010).
[11] D. C. Brody and L. P. Hughston, J. Geom. Phys. **38**, 19 (2001).
[12] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.109.100501 for a gradient computation using adjoint equations.
[13] D. C. Brody, D. D. Holm, and D. M. Meier, arXiv:1206.2675.