

# OPTIMAL CONTROL ALGORITHMS FOR QUANTUM SYSTEMS AND THEIR LIMITATIONS

**Sophie G. Schirmer**  
D.A.M.T.P.  
University of Cambridge  
U.K.  
sgs29@cam.ac.uk

**Pierre Becq de Fouquieres**  
D.A.M.T.P.  
University of Cambridge  
U.K.  
plbd2@cam.ac.uk

**Xiaoting Wang**  
D.A.M.T.P.  
University of Cambridge  
U.K.  
xw233@cam.ac.uk

## Abstract

Some of the limitations of popular optimal control algorithms such as Lyapunov control, the Krotov method and other iterative algorithms based on functional gradients are considered. Lyapunov control, while elegant and generally effective for ideal systems, struggles for realistic systems. Iterative techniques such as the Krotov method tend to be more effective, but the inclusion of an energy penalty term implies that the resulting solutions are not critical points of the objective functional alone, and do not maximize it. These problems can be avoided by using functional gradients without penalty terms, but even these techniques fail for some problems.

## 1 Introduction

Optimal control has been shown to be a powerful approach to solve optimal pulse design and Hamiltonian engineering problems for quantum systems [1]. Typically optimal control involves several steps: (1) formulation of the control problem as an optimization problem by specifying a functional to be maximized, (2) parametrization of the control fields to reduce the dimensionality of the (generally infinite dimensional) state space, and (3) numerical algorithms to solve the resulting finite-dimensional optimization problem. In this paper we focus on the final task, i.e., optimal control algorithms, although the first two tasks are clearly non-trivial.

---

\*SGS acknowledges support from an EPSRC Advanced Research Fellowship, the QIP IRC and Hitachi. PLBF acknowledges support from an EPSRC CASE studentship in collaboration with Hitachi. XW is supported by a studentship from the Cambridge Overseas Trust and an Elizabeth Cherry Major Scholarship from Hughes Hall, Cambridge. Some of this work was inspired by discussions with other programme participants while SGS was visiting the Institute for Mathematics and its Applications (IMA) at the University of Minnesota.

<sup>†</sup>The authors are affiliated with the Dept of Applied Maths and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK. Correspondence should be addressed to sgs29@cam.ac.uk.

Although optimization is a rather well-developed field, the optimization problems that arise in quantum control are typically not simple convex optimization problems. For this reason global search techniques, e.g., based on genetic or evolutionary strategies were initially advocated to solve the optimal control problems arising in areas such as quantum chemistry or quantum information processing. However, global search strategies are computationally expensive, and it was soon realized that the control landscape, although not convex, may after all be very simple. E.g., in [2] it was observed that the objective function  $J = \text{Tr}[A\rho(t_F)]$ , where  $A = |\Psi_f\rangle\langle\Psi_f|$  is a projector onto a pure state and  $\rho(t_F) = U(t_F)|\Psi_0\rangle\langle\Psi_0|U(t_F)^\dagger$ , has only two critical values  $J = 0$  and  $J = 1$  corresponding to the global minimum and maximum respectively, if the optimization is performed over the entire unitary group  $U(N)$ . It can therefore be argued that for controllable systems [3] and state transfer problems with objective functionals of the form above, all peaks (maxima) of  $J$  have the same height, and thus a simple local search using a gradient-based method should suffice to find a globally optimal solution.

Various gradient-based algorithms have been proposed to solve the resulting optimal control problems. Among the most popular are Lyapunov control [4; 5; 6; 7; 8], iterative algorithms based on generalizations of the Krotov method [12], and gradient-ascent methods such as GRAPE [13]. Although most of these strategies can be shown to exhibit desirable properties such as monotonic convergence, and have been successfully applied to a wide variety of problems to find control pulses that perform far better than simple, intuitive control schemes, there are open questions and gaps in existing arguments that should be addressed. In the following we will highlight some of these issues.

## 2 Trajectory Tracking via Lyapunov Control

One formulation of an optimal control problem is in terms of steering the system to a target state  $\rho_d$

asymptotically, or tracking a target trajectory  $\rho_d(t)$ . In this case the objective is to find an admissible control  $\mathbf{f}(t)$  such that the trajectory  $\rho(t)$  of the controlled system asymptotically tracks the target trajectory, i.e.,  $\|\rho(t) - \rho_d(t)\| \rightarrow 0$  as  $t \rightarrow \infty$ . The states referred to here can be pure-state wavefunctions, mixed states or unitary operators. We shall assume that  $\rho(t)$  is a density operator, i.e., a positive unit-trace operator on the system's Hilbert space  $\mathcal{H}$ , representing a pure or mixed state of the system. Assuming a bilinear Hamiltonian control system, the simplest and most common type considered in the literature, the system and target states,  $\rho(t)$  and  $\rho_d(t)$ , respectively, satisfy the following dynamical laws

$$\dot{\rho}(t) = -i[H_0 + f(t)H_1, \rho(t)], \quad (1a)$$

$$\dot{\rho}_d(t) = -i[H_0, \rho_d(t)], \quad (1b)$$

where  $H_0$  and  $H_1$  are a system and control Hamiltonian, respectively, and  $f(t)$  is the control field. It is easy to verify that the so-called feedback control law

$$f(t) = \text{Tr}([-iH_1, \rho(t)]\rho_d(t)) \quad (2)$$

ensures that the Hilbert-Schmidt distance

$$\begin{aligned} V(t) = V(\rho(t), \rho_d(t)) &= \frac{1}{2} \|\rho(t) - \rho_d(t)\|^2 \\ &= \text{Tr}[\rho_d^2(t)] - \text{Tr}[\rho(t)\rho_d(t)] \end{aligned} \quad (3)$$

is monotonically decreasing as

$$\dot{V}(\rho(t), \rho_d(t)) = -f(t) \text{Tr}([-iH_1, \rho(t)]\rho_d(t)) \leq 0, \quad (4)$$

and thus  $V(\rho, \rho_d)$  decreases along any trajectory  $(\rho(t), \rho_d(t))$  of the system. If  $V(\rho, \rho_d) \rightarrow 0$  then  $\rho(t) \rightarrow \rho_d(t)$  and the control objective is realized. It is easy to see that the control is optimal in this case. In fact, it can be shown that  $f(t) = \langle \rho_*(t) | \partial J_*(\rho) / \partial \rho \rangle$ , where  $J_*(\rho)$  satisfies the Hamilton-Jacobi equations.

However, we cannot conclude  $V(\rho, \rho_d) \rightarrow 0$  in general. Rather, we can only employ LaSalle's Invariance Principle to show that the system converges to a maximal invariant set  $E = \{\dot{V} \equiv 0\}$ , which unfortunately, is generally large. Careful analysis shows that we can distinguish two cases [8]:

**Ideal Systems:** Lyapunov control is generally effective, i.e., system state  $\rho(t)$  converges to  $\rho_d(t)$  as  $t \rightarrow \infty$  for almost all initial states  $\rho(0)$ , except for a measure-zero set of bad target states  $\rho_d$ .

**Non-ideal Systems:** Lyapunov control is generally not effective, i.e., target state is a centre on a centre manifold, and almost all trajectories converge to centre manifold but not  $\rho_d$ .

Here, the dynamical system with  $H = H_0 + f(t)H_1$  is considered ideal if  $H_0$  is strongly regular and  $H_1$  is fully connected. Although ideal Hamiltonians are "generic" in an abstract sense, most physical systems not ideal even if they are controllable, except for  $n = 2$ . Thus, while the method is theoretically effective and optimal for most systems and target states, this is generally not the case for systems of physical interest.

### 3 Generalized Krotov Method

An alternative to optimal control techniques based on instantaneous minimization of a Lyapunov function are iterative techniques. A particularly popular approach in quantum chemistry [9; 10; 11], especially for optimization problems where the objective is to maximize the expectation value of a target observable  $A$  at a fixed target time  $t_F$ , are generalized Krotov methods. These are based on choosing a trial field  $\mathbf{f}^{(0)}(t)$  with bounded components

$$\|\mathbf{f}_m^{(0)}\|_\infty \equiv \max_{t_0 \leq t \leq t_f} |f_m^{(0)}(t)| < \infty \quad (5)$$

for  $m = 1, \dots, M$ , and then repeatedly solving the *initial value problem* for the variational trial function  $\rho_v^{(n)}(t)$

$$\frac{\partial}{\partial t} \rho_v^{(n)}(t) = -\frac{i}{\hbar} \mathfrak{L}[\mathbf{f}^{(n)}(t), \rho_v^{(n)}(t)], \quad \rho_v^{(n)}(t_0) = \rho_0, \quad (6)$$

followed by the *final value problem* for the variational trial function  $A_v^{(n)}(t)$

$$\frac{\partial}{\partial t} A_v^{(n)}(t) = -\frac{i}{\hbar} \mathfrak{L}[\tilde{\mathbf{f}}^n(t), A_v^{(n)}(t)], \quad A_v^{(n)}(t_F) = A, \quad (7)$$

while updating the control fields after each step according to the rule

$$f_m^{(n)} = (1 - \alpha) \tilde{f}_m^{(n-1)} - \frac{i\alpha}{\lambda_m} \langle \langle A_v^{(n-1)} | \mathfrak{L}_m | \rho_v^{(n)} \rangle \rangle \quad (8a)$$

$$\tilde{f}_m^{(n)} = (1 - \beta) f_m^{(n)} - \frac{i\beta}{\lambda_m} \langle \langle A_v^{(n)} | \mathfrak{L}_m | \rho_v^{(n)} \rangle \rangle \quad (8b)$$

One can show that for  $\alpha, \beta \in [0, 2)$  this algorithm converges to a solution of the Euler-Lagrange equations

$$0 = \frac{\delta \mathcal{J}_{\text{tot}}}{\delta A_v} = \frac{\partial}{\partial t} \rho_v(t) + \frac{i}{\hbar} \mathfrak{L}[\mathbf{f}(t), \rho_v(t)], \quad (9a)$$

$$0 = \frac{\delta \mathcal{J}_{\text{tot}}}{\delta \rho_v} = \frac{\partial}{\partial t} A_v(t) + \frac{i}{\hbar} \mathfrak{L}[\mathbf{f}(t), A_v(t)], \quad (9b)$$

$$0 = \frac{\delta \mathcal{J}_{\text{tot}}}{\delta f_m} = \frac{\lambda_m}{\hbar} f_m(t) + \frac{i}{\hbar} \langle \langle A_v(t) | \mathfrak{L}_m | \rho_v(t) \rangle \rangle. \quad (9c)$$

for the target functional

$$\mathcal{J} = \mathcal{A} - \mathcal{C} = \text{Tr}[\hat{A}\hat{\rho}(t_F)] - \sum_{m=1}^M \frac{\lambda_m}{2\hbar} \|f_m\|_2^2. \quad (10)$$

Thus, the resulting solutions  $(\rho(t), A(t), \mathbf{f}(t))$  are optimal in the sense that they are critical points of the target functional  $\mathcal{J}$ . However, due to the added penalty term  $\mathcal{C}$ , the critical points of  $\mathcal{J}$  are generally not critical points of  $\mathcal{A}$ , and one can verify that the solutions generally do not satisfy  $[\rho(t_F), A(t_F)] = 0$ . Since the critical points  $\rho(t)$  of  $\mathcal{A}$  must commute with the observable at the target time, we must conclude that the solutions found by the Krotov method are generally not optimal solutions for the unconstrained optimization problem, i.e., they do not maximize  $\mathcal{A}$ .

#### 4 Gradient-Ascent Algorithms without Penalty Terms

The previous section shows that the introduction of the penalty term results in solutions that fail to maximize the objective function. A logical step therefore is to eliminate the penalty term and attempt to maximize the objective functional directly. For a simple (pure) state transfer problem of steering the system from the initial state  $|\Psi_0\rangle$  to a final state  $|\Psi_d\rangle$  at time  $t_F$  by applying a suitable control  $f(t)$ , the objective function can be simplified to

$$\mathcal{A} = |\langle \Psi_d | U_{f(t)}(t_F) | \Psi_0 \rangle|^2. \quad (11)$$

where  $U_{f(t)}(t)$  must satisfy the Schrodinger equation

$$\frac{d}{dt} U_{f(t)}(t) = -i[H_0 - f(t)H_1]U_{f(t)}(t) \quad (12)$$

for  $t \in [0, t_F]$ , and  $U_{f(t)}(0) = I$ .

This state transfer problem was recently analyzed by [14], where it was attempted to show that the functional derivative kernel  $\frac{\delta \mathcal{A}}{\delta f(t)}$  can be identically zero for a given  $f(t)$  only when the associated objective function  $\mathcal{A}$  attains the extremal values of 0 or 1, provided that the system is controllable, i.e., that  $iH_0$  and  $iH_1$  generate the Lie algebra  $u(N)$  or  $su(N)$ , where  $N$  is the dimension of the underlying Hilbert space. We show that, although this appears to be true for most initial and final states, there are counter-examples. It is easy to verify that  $\frac{\delta \mathcal{A}}{\delta f}$  equals

$$2 \text{Im} \left( \langle \Psi_0 | U^\dagger(t) H_1 U(t) U^\dagger(t_F) | \Psi_d \rangle \langle \Psi_d | U(t_F) | \Psi_0 \rangle \right) \quad (13)$$

where we have dropped the subscript  $f(t)$  indicating the control dependence of  $U(t)$ .

From this we see one way for  $\delta \mathcal{A} / \delta f$  to vanish identically is if one of the factors on the RSH vanishes identically. The second factor vanishes identically if and only

if  $\langle \Psi_d | U(t_F) | \Psi_0 \rangle = 0$ , i.e., if the time-evolved initial state  $|\Psi(t_F)\rangle$  has zero overlap with the target state. It is easy to see that a control  $f(t)$  and corresponding trajectory  $U_{f(t)}(t)$  for which this happens is a global minimum of the objective functional  $\mathcal{A}$  and  $\mathcal{A}(f(t)) = 0$ . This is the most common case, but it is possible to construct examples such that  $\delta \mathcal{A} / \delta f \equiv 0$  but  $\mathcal{A}(f(t))$  takes values between 0 and 1.

One way to construct such examples is by making  $\langle \Psi_0 | U^\dagger(t) H_1 U(t) U^\dagger(t_F) | \Psi_d \rangle = 0$  for all  $t$ . That this is possible can be seen by choosing

$$|\Psi_0\rangle = \frac{1}{\sqrt{2}}(1, 0, 1)^T, \quad |\Psi_d\rangle = \frac{1}{\sqrt{2}}(1, 0, -1)^T$$

and

$$H_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

in a basis which diagonalises  $H_0$ . Then we have for  $f = 0$

$$\langle \Psi_0 | U^\dagger(t) H_1 U(t) U^\dagger(t_F) | \Psi_d \rangle = 0$$

for any  $t$  and  $t_F$ , since any product of the form

$$\begin{pmatrix} a \\ 0 \\ b \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} c \\ 0 \\ d \end{pmatrix} = 0$$

Setting  $H_0 = \text{diag}(\epsilon_1, \epsilon_2, \epsilon_3)$  and  $\omega_{13} = \epsilon_3 - \epsilon_1$  shows that  $\mathcal{A}(\Psi_0, \Psi_d, f) = \frac{1}{2}[1 - \cos(\omega_{13}t_F)]$  and thus depending on the target time  $t_F$  and energy gap  $\omega_{13}$ , any value in  $(0, 1)$  can be achieved. Furthermore, as  $H_1$  is connected, it suffices to choose  $H_0$  strongly regular, e.g.,  $H_0 = \text{diag}(2, 3, 5)$  to ensure controllability of the system.

The importance of the previous example, while it may seem contrived, is that there are choices of  $|\Psi_0\rangle$ ,  $|\Psi_d\rangle$  and  $f(t)$  such that  $\mathcal{A}(\Psi_0, \Psi_d, f)$  takes any value in  $(0, 1)$  even though  $\delta \mathcal{A} / \delta f$  vanishes identically, and this is the case even for fully controllable systems. More extensive analysis shows that in the generic case of

$$H_0 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad H_1 = \begin{pmatrix} c & d \\ d & c \end{pmatrix},$$

the choice of states

$$|\Psi_0\rangle = \frac{1}{\sqrt{2}|d|} \begin{pmatrix} |d| \\ d \end{pmatrix}, \quad |\Psi_d\rangle = \frac{1}{\sqrt{2}|d|} \begin{pmatrix} |d| \\ -d \end{pmatrix}$$

ensures that  $\delta\mathcal{A}/\delta f \equiv 0$  for  $f(t) = 0$ . Indeed, we can explicitly compute

$$\begin{aligned} & \langle \Psi_0 | U^\dagger(t) H_1 U(t) U^\dagger(t_F) | \Psi_d \rangle \langle \Psi_d | U(t_F) | \Psi_0 \rangle \\ &= \frac{|d|}{2} [\cos(\omega t) - \cos(\omega(t_F - t))] + \frac{c}{2} [1 - \cos(\omega t_F)] \end{aligned}$$

where  $\omega = b - a$ , which is clearly real, as well as  $\mathcal{A} = \frac{1}{2}[1 - \cos(\omega t_F)]$ , which shows that we can attain any value in  $(0, 1)$  for some  $t_F \in (0, \frac{\pi}{\omega})$ , and excluding the special cases when  $d = 0$  or  $a = b$  is enough to guarantee controllability.

More importantly, the previous construction provides counter-examples for systems of any dimension, not just  $N = 2$ , as for any two diagonal elements of  $H_1$  which are equal, we can restrict attention to the two-dimensional subspace they act on, construct  $|\Psi_0\rangle$  and  $|\Psi_d\rangle$  as above, and embed these states into  $\mathcal{H}$  to obtain examples for which  $\frac{\delta\mathcal{A}}{\delta f} = 0$  but  $\mathcal{A}$  assumes values in  $(0, 1)$ . The counter-examples thus constructed are such that the initial and target states both lie on a great circle in a 2D subspace of the Hilbert space. This is reminiscent of the special case of target states lying on a great circle in a 2D subspace of  $\mathcal{H}$  for Lyapunov control, for which the Lyapunov function could take any value between 0 and its maximum despite  $\dot{V}(t) \equiv 0$ . Although it seems that these counter-examples are quite special and form a very small set of optimal control problems, they do include some control problems of interest involving preparing/transforming CAT states or maximally entangled states. The construction above only provides examples for certain special non-stationary target states. But examples for which  $\delta\mathcal{A}/\delta f = 0$  but  $\mathcal{A} \neq 0, 1$  exist even for stationary target states. E.g. consider the three-level system

$$H_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 1 & \sqrt{2/3} & 0 \\ \sqrt{2/3} & 2 & \sqrt{1/3} \\ 0 & \sqrt{1/3} & 4 \end{pmatrix}$$

with initial state  $|\Psi_0\rangle = (1, 0, 0)^T$  and target state  $|\Psi_d\rangle = (0, 0, 1)^T$ . Controllability of this system again follows by connectedness of  $H_1$  and strong regularity of  $H_0$ , but setting, e.g.,  $t_F = \pi$  and  $f(t) = 1$ , we can explicitly calculate  $\mathcal{A} = \frac{8}{9}$ , while  $\frac{\delta\mathcal{A}}{\delta f}(t) = 2\text{Im}[\frac{4}{9}(3\cos(t) - 5)]$  indeed vanishes.

## 5 Conclusions

We have considered the limitations of several popular algorithms for solving quantum optimal control problems. In particular we have shown that none of the algorithms work for all control problems. Lyapunov control, though elegant, is generally the most restrictive in terms of the conditions imposed on the system Hamiltonian as well as the target states. Iterative numerical techniques such as the Krotov method are less likely to

fail completely but generally do not lead to solutions that maximize the objective functional due to the presence of an energy penalty term. We have also shown by constructing explicit counter-examples that even functional gradient techniques which have been claimed not to suffer from such problems, can fail for certain problems. In particular, it is possible for the functional gradients to vanish even when the objective functional is not maximized or minimized but assumes intermediate values, which was thought to be impossible.

## References

- S. G. Schirmer, P. J. Pemberton-Ross and X. Wang. In Proceedings of PhysCon 2007.
- H. Rabitz, M. Hsieh, C. M. Rosenthal, Science 303, 1998 (2004)
- V. Jurdjevic, J. Quinn, Controllability and stability. J. Differential Equations 28, 381-389 (1978)
- M. Mirrahimi and P. Rouchon, In Proceedings of the international symposium MTNS 2004.
- M. Mirrahimi and G. Turinici, Automatica 41, 1987-1994 (2005)
- C. Altafini, Feedback control of spin systems, Quantum Information Processing 6, 9 (2007)
- C. Altafini, IEEE Trans. Autom. Control 52, 1 (2007)
- X. Wang, S. Schirmer, <http://arXiv.org/abs/0801.0702>
- V. F. Krotov, I. N. Feldman, Izv. Akad. Nauk SSSR, Tekh. Kibern. 2, 160-168 (1983)
- A. I. Konnov and V. F. Krotov, Automation and Remote Control 60, 1427 (1999)
- D.J. Tannor, V. Kazakov and V. Orlov, In: *Time Dependent Quantum Molecular Dynamics*, J. Broeckhove and L. Lathouwers, eds., 347-360 (Plenum, 1992).
- Y. Maday and G. Turinici, J. Chem. Phys. 118, 8191 (2003)
- N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrueggen, S. Glaser, J. Mag. Resonance 172, 296 (2005)
- H. Rabitz, T.-S. Ho, M. Hsieh, R. Kosut, and M. Demiralp, Phys. Rev. A 74, 012721 (2006)