The final gate objective functional

N. Anders Petersson*

July 22, 2019

1 Introduction

Let the wave functions $\boldsymbol{\psi}_j(t,\boldsymbol{\alpha}) \in [0,T] \times \mathbb{R}^D \to \mathbb{C}^N$ be governed by the Schrödinger equation,

$$\dot{\boldsymbol{\psi}}_j + iH(t,\boldsymbol{\alpha})\boldsymbol{\psi}_j = 0, \quad 0 \le t \le T, \quad \boldsymbol{\psi}_j(0) = \mathbf{e}_j, \quad j = 1,2,\ldots,N, \tag{1}$$

Here, $\mathbf{e}_j$ is the j^{th} canonical unit vector (zero, except element number j which is one). The Hamiltonian matrix satisfies

$$H(t,\boldsymbol{\alpha}) = H_0 + p(t,\boldsymbol{\alpha})H_c, \quad H(t,\boldsymbol{\alpha}) = H^\dagger(t,\boldsymbol{\alpha}),$$

where $p(t,\boldsymbol{\alpha})$ is a scalar function of time that depends on the parameter vector $\boldsymbol{\alpha} = [\alpha_1,\alpha_2,\ldots,\alpha_D]^T \in \mathbb{R}^D$. We introduce

$$\boldsymbol{\phi}_{jk}(t,\boldsymbol{\alpha}) = \frac{\partial \boldsymbol{\psi}_j(t,\boldsymbol{\alpha})}{\partial \alpha_k}, \quad k = 1,2,\ldots,D.$$

By differentiating (1) with respect to α_k,

$$\dot{\boldsymbol{\phi}}_{jk} + iH(t,\boldsymbol{\alpha})\boldsymbol{\phi}_{jk} = \mathbf{f}_{jk}(t,\boldsymbol{\alpha}), \quad 0 \le t \le T, \quad \boldsymbol{\phi}_{jk}(0) = 0, \tag{2}$$

$$\mathbf{f}_{jk}(t,\boldsymbol{\alpha}) = -i\frac{\partial p(t,\boldsymbol{\alpha})}{\partial \alpha_k} H_c \boldsymbol{\psi}_j(t,\boldsymbol{\alpha}). \tag{3}$$

Let $U_g = [\mathbf{d}_1,\mathbf{d}_2,\ldots,\mathbf{d}_N] \in \mathbb{C}^{N\times N}$, represent the target unitary matrix ($U_g^{-1} = U_g^\dagger$) and collect the wave functions for the different initial data in the unitary matrix $U = U(T,\alpha) \in \mathbb{C}^{N\times N}$, where

$$U = [\boldsymbol{\psi}_1,\boldsymbol{\psi}_2,\ldots,\boldsymbol{\psi}_N], \quad U^\dagger = U^{-1}. \tag{4}$$

An objective functional that measures the infidelity in the final unitary is given by

$$g_1(U(T,\boldsymbol{\alpha})) = 1 - \frac{1}{N^2}\left|S_T(\boldsymbol{\alpha})\right|^2, \quad S_T(\boldsymbol{\alpha}) = \left\langle U(T,\boldsymbol{\alpha}), U_g \right\rangle_F. \tag{5}$$

*Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, L-561, PO Box 808, Livermore CA 94551.

We define the Fröbenius matrix scalar product between square complex-valued matrices A and B by

$$\langle A, B\rangle_F = \mathrm{tr}\big(A^\dagger B\big) = \sum_{j=1}^{N} \langle \mathbf{a}_j, \mathbf{b}_j\rangle_2, \tag{6}$$

where $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N]$ and $B = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N]$. By differentiating (5) with respect to α_k,

$$\frac{\partial g_1}{\partial \alpha_k} = -\frac{2}{N^2}\mathrm{Re}\left(\frac{\partial S_T}{\partial \alpha_k}\overline{S}_T\right). \tag{7}$$

From (5) and (6),

$$S_T(\boldsymbol{\alpha}) = \sum_{j=1}^{N} \left\langle \boldsymbol{\psi}_j(T,\boldsymbol{\alpha}), \mathbf{d}_j\right\rangle_2, \quad \frac{\partial S_T(\boldsymbol{\alpha})}{\partial \alpha_k} = \sum_{j=1}^{N} \left\langle \boldsymbol{\phi}_{jk}(T,\boldsymbol{\alpha}), \mathbf{d}_j\right\rangle_2. \tag{8}$$

Inserting (8) into (7) results in

$$\frac{\partial g_1}{\partial \alpha_k} = -\frac{2}{N^2}\mathrm{Re}\left(\sum_{j=1}^{N} \left\langle \boldsymbol{\phi}_{jk}(T,\boldsymbol{\alpha}), \mathbf{d}_j\right\rangle_2 \overline{S}_T\right) = -2\,\mathrm{Re}\sum_{j=1}^{N} \left\langle \boldsymbol{\phi}_{jk}(T,\boldsymbol{\alpha}), \frac{\overline{S}_T}{N^2}\mathbf{d}_j\right\rangle_2. \tag{9}$$

2 The adjoint state equation

Define the adjoint state equation according to

$$-\dot{\boldsymbol{\lambda}}_j - iH(t,\boldsymbol{\alpha})\boldsymbol{\lambda}_j = \mathbf{g}_j(t), \quad T \geq t \geq 0, \quad \boldsymbol{\lambda}_j(T) = \mathbf{q}_j, \quad j = 1, 2, \ldots, N. \tag{10}$$

Note that the adjoint equation is solved backwards in time and is subject to a terminal condition The forcing function $\mathbf{g}_j(t)$ and the terminal condition vector $\mathbf{q}_j$ will be determined below.

The solutions of the state equation (2) and the adjoint state equation (10) are related. To establish this relation, let's consider

$$\begin{aligned} I := \int_0^T \langle \boldsymbol{\phi}_{jk}, \mathbf{g}_j\rangle_2\, d\tau &= \int_0^T \langle \boldsymbol{\phi}_{jk}, -\dot{\boldsymbol{\lambda}}_j - iH(t,\boldsymbol{\alpha})\boldsymbol{\lambda}_j\rangle_2\, d\tau \\ &= -\int_0^T \langle \boldsymbol{\phi}_{jk}, \dot{\boldsymbol{\lambda}}_j\rangle_2\, d\tau + \int_0^T \langle iH(t,\boldsymbol{\alpha})\boldsymbol{\phi}_{jk}, \boldsymbol{\lambda}_j\rangle_2\, d\tau. \end{aligned} \tag{11}$$

By integration by parts,

$$\begin{aligned} \int_0^T \langle \boldsymbol{\phi}_{jk}, \dot{\boldsymbol{\lambda}}_j\rangle_2\, d\tau &= \big[\langle \boldsymbol{\phi}_{jk}, \boldsymbol{\lambda}_j\rangle_2\big]_0^T - \int_0^T \langle \dot{\boldsymbol{\phi}}_{jk}, \boldsymbol{\lambda}_j\rangle_2\, d\tau \\ &= \langle \boldsymbol{\phi}_{jk}(T), \mathbf{q}_j\rangle_2 - \int_0^T \langle \dot{\boldsymbol{\phi}}_{jk}, \boldsymbol{\lambda}_j\rangle_2\, d\tau, \end{aligned}$$

because $\boldsymbol{\phi}_{jk}(0) = 0$ and $\boldsymbol{\lambda}_j(T) = \mathbf{q}_j$. Thus,

$$I = -\langle \boldsymbol{\phi}_{jk}(T), \mathbf{q}_j \rangle_2 + \int_0^T \langle \dot{\boldsymbol{\phi}}_{jk} + iH(t,\boldsymbol{\alpha})\boldsymbol{\phi}_{jk}, \boldsymbol{\lambda}_j \rangle_2 \, d\tau$$
$$= -\langle \boldsymbol{\phi}_{jk}(T), \mathbf{q}_j \rangle_2 + \int_0^T \langle \mathbf{f}_{jk}, \boldsymbol{\lambda}_j \rangle_2 \, d\tau.$$

Therefore,

$$\langle \boldsymbol{\phi}_{jk}(T), \mathbf{q}_j \rangle_2 + \int_0^T \langle \boldsymbol{\phi}_{jk}, \mathbf{g}_j \rangle_2 \, d\tau = \int_0^T \langle \mathbf{f}_{jk}, \boldsymbol{\lambda}_j \rangle_2 \, d\tau. \tag{12}$$

where $\mathbf{f}_{jk}(t)$ is given by (3).

To evaluate (9), we set $\mathbf{q}_j = -\overline{S}_T \mathbf{d}_j / N^2$ and $\mathbf{g}_j(t) = 0$. Then,

$$\frac{\partial g_1}{\partial \alpha_k} = -2\text{Re}\sum_{j=1}^N \left\langle \boldsymbol{\phi}_{jk}(T,\boldsymbol{\alpha}), \frac{\overline{S}_T \mathbf{d}_j}{N^2} \right\rangle_2$$
$$= 2\text{Re}\sum_{j=1}^N \left\langle \boldsymbol{\phi}_{jk}(T,\boldsymbol{\alpha}), \mathbf{q}_j \right\rangle_2 = 2\text{Re}\sum_{j=1}^N \int_0^T \langle \mathbf{f}_{jk}, \boldsymbol{\lambda}_j \rangle_2 \, d\tau. \tag{13}$$

Similar to before, the cost of computing all components of the gradient of the objective function is almost independent of the number of components. We start by solving the Schrödinger equation forwards in time to obtain the terminal state $\boldsymbol{\psi}_j(T) =: \mathbf{w}_j$ for $j = 1, 2, \ldots, N$. The time stepping is then reversed to integrate (1) backwards in time,

$$\dot{\boldsymbol{\psi}}_j + iH(t,\boldsymbol{\alpha})\boldsymbol{\psi}_j = 0, \quad T \geq t \geq 0, \quad \boldsymbol{\psi}_j(T) = \mathbf{w}_j, \quad j = 1, 2, \ldots, N. \tag{14}$$

The adjoint wave equation (10) (with $\mathbf{g}_j = 0$) is simulataneously solved backwards in time to calculate $\boldsymbol{\lambda}_j(t)$. At each time step, (3) is evaluated to calculate $\mathbf{f}_{jk}(t)$ and combined with $\boldsymbol{\lambda}_j(t)$, accumulate the integral (13) to compute the gradient of the objective functional.

3 Discouraging population of higher energy states

The actual Hamiltonian system is in general infinite-dimensional. To make the dimensionality of the system finite we introduce an additional term in the objective functional to discourage population of highly energetic states. Let $N_g \geq 0$ denote the number of guard states and expand the wave function such that $\boldsymbol{\psi}_j(t,\boldsymbol{\alpha}) \in [0,T] \times \mathbb{R}^D \to \mathbb{C}^{N_t}$, where $N_t = N + N_g$. The Schrödinger equation (1) still governs the evolution of $\boldsymbol{\psi}_j(t,\boldsymbol{\alpha})$, but now also models the evolution of the guard states. As a result the Hamiltonian matrix H is now of size $N_t \times N_t$. The forcing function $\mathbf{f}_{jk}(t,\boldsymbol{\alpha})$ is now a vector with N_t elements. The matrices $U(t,\boldsymbol{\alpha})$ and U_g are now represented by rectangular matrices with N_t rows and N columns,

$$U = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_N], \quad U_g = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \ldots & \mathbf{d}_N \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \end{bmatrix}.$$

Here, the zero vector $\mathbf{0} \in \mathbb{R}^{N_g}$. Thus, the definition of the objective function g_1 in (5) still holds. Furthermore, only the first N elements of $\boldsymbol{\psi}_j$ and $\boldsymbol{\phi}_{jk}$ matter because the last N_g rows of U_g are identically zero. To accomodate for the guard levels, the terminal condition for the adjoint wave equation (10) becomes

$$\boldsymbol{\lambda}_j(T) = -\frac{\overline{S}_T}{N^2} \begin{bmatrix} \mathbf{d}_j \\ \mathbf{0} \end{bmatrix}. \tag{15}$$

To measure the population of the guard levels we define the functional

$$g_2(U(\boldsymbol{\alpha})) = \int_0^T \sum_{j=1}^{N} \left\langle \boldsymbol{\psi}_j(\tau, \boldsymbol{\alpha}), W \boldsymbol{\psi}_j(\tau, \boldsymbol{\alpha}) \right\rangle_2 \, d\tau, \tag{16}$$

where W is a real diagonal matrix with zero entries in the first N rows and columns,

$$W = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & w_1 & & \\ & & & & \ddots & \\ & & & & & w_{N_g} \end{bmatrix}, \quad 0 < w_1 \leq w_2 \leq \ldots \leq w_{N_g}.$$

The gradient of g_2 satisfies

$$\frac{\partial g_2}{\partial \alpha_k} = 2\,\mathrm{Re} \sum_{j=1}^{N} \int_0^T \left\langle \boldsymbol{\phi}_{jk}(\tau, \boldsymbol{\alpha}), W \boldsymbol{\psi}_j(\tau, \boldsymbol{\alpha}) \right\rangle_2 \, d\tau.$$

The gradient of the combined objective functional

$$g(\boldsymbol{\alpha}) = g_1(\boldsymbol{\alpha}) + g_2(\boldsymbol{\alpha}),$$

therefore satisfies

$$\frac{\partial g}{\partial \alpha_k} = 2\mathrm{Re} \sum_{j=1}^{N} \left(-\left\langle \boldsymbol{\phi}_{jk}(T, \boldsymbol{\alpha}), \frac{\overline{S}_T}{N^2} \mathbf{d}_j \right\rangle_2 + \int_0^T \left\langle \boldsymbol{\phi}_{jk}(\tau, \boldsymbol{\alpha}), W \boldsymbol{\psi}_j(\tau, \boldsymbol{\alpha}) \right\rangle_2 \, d\tau \right).$$

Note that each term in the sum is of the type (13). Thus, we can compute the gradient using the previous approach by solving the adjoint equation (10) with terminal condition (15) and forcing function

$$\mathbf{g}_j = W \boldsymbol{\psi}_j.$$