Optimal control of quantum systems with Padé integrators and direct collocation: PICO.jl

Aaron Trowbridge

Abstract

Here we present a novel approach to the problem of quantum optimal control which takes advantage of a relationship between the implicit midpoint method and the Padé approximation of the matrix exponential function. We demonstrate this approach, using an accompanying Julia package, on an assortment of quantum systems including a single qubit system, two qubit system, and a system composed of a qubit coupled to an array of harmonic oscillators.

Contents

0 Introduction **1**

1 Problem Formulation **1**

2 Dynamics **2**

2.1 Padé integrators 3

3 Differentiation **4**

3.1 Objective Gradient 4

3.2 Dynamics Jacobian 5

3.3 Hessian of the Lagrangian 7

0 Introduction

1 Problem Formulation

Given a quantum system with a Hamiltonian of the form

$$H(\mathbf{a}(t), t) = H_{\text{drift}} + \sum_{j=1}^{c} a^j(t) H_{\text{drive}}^j$$

we solve the optimization problem

$$\begin{aligned}
\underset{\mathbf{x}_{1:T},\ \mathbf{u}_{1:T-1}}{\text{minimize}} \quad & \sum_{i=1}^{n} Q \cdot \ell(\tilde{\psi}_T^i, \tilde{\psi}_{\text{goal}}^i) + \frac{1}{2} \sum_{t=1}^{T-1} R_t \cdot \mathbf{u}_t^2 \\
\text{subject to} \quad & \mathbf{f}(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{u}_t) = \mathbf{0} \\
& \tilde{\psi}_1^i = \tilde{\psi}_{\text{init}}^i \\
& \tilde{\psi}_T^1 = \tilde{\psi}_{\text{goal}}^1 \quad \text{if} \quad \mathsf{pin_first_qstate} = \mathsf{true} \\
& \textstyle\int \mathbf{a}_1 = \mathbf{a}_1 = \mathrm{d}_t \mathbf{a}_1 = \mathbf{0} \\
& \textstyle\int \mathbf{a}_T = \mathbf{a}_T = \mathrm{d}_t \mathbf{a}_T = \mathbf{0} \\
& |a_t^j| \leq a_{\text{bound}}^j
\end{aligned}$$

The *state* vector $\mathbf{x}_t$ contains both the n ($\mathsf{nqstates}$) quantum isomorphism states $\tilde{\psi}_t^i$ (each of dimension $\mathsf{isodim} = 2{*}\mathsf{ketdim}$) and the augmented control states $\int \mathbf{a}_t$, $\mathbf{a}_t$, and $\mathrm{d}_t \mathbf{a}_t$ (the number of augmented state vector is augdim). The *action* vector $\mathbf{u}_t$ contains the second derivative of the *control* vector $\mathbf{a}_t$, which has dimension $\mathsf{ncontrols}$. Thus, we have:

$$\mathbf{x}_t = \begin{pmatrix} \tilde{\psi}_t^1 \\ \vdots \\ \tilde{\psi}_t^n \\ \int \mathbf{a}_t \\ \mathbf{a}_t \\ \mathrm{d}_t \mathbf{a}_t \end{pmatrix} \quad \text{and} \quad \mathbf{u}_t = \left(\mathrm{d}_t^2 \mathbf{a}_t\right) \tag{1}$$

In summary,

$$\begin{aligned}
\dim(\mathbf{x}_t) &= \mathsf{nstates} = \mathsf{nqstates} * \mathsf{isodim} + \mathsf{ncontrols} * \mathsf{augdim} \\
\dim(\mathbf{u}_t) &= \mathsf{ncontrols}
\end{aligned}$$

Additionally the loss function ℓ can be chosen somewhat liberally, the default is currently

$$\ell(\tilde{\psi}, \tilde{\psi}_{\text{goal}}) = 1 - |\langle \psi | \psi_{\text{goal}} \rangle|^2$$

2 Dynamics

Finally, $\mathbf{f}(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{u}_t)$ describes the dynamics of all the variables in the system, where the controls' dynamics are trivial and formally $\tilde{\psi}_t^i$ satisfies a discretized version of the isomorphic Schröedinger equation:

$$\frac{\mathrm{d}\tilde{\psi}^i}{\mathrm{d}t} = \widetilde{(-iH)}(\mathbf{a}(t), t)\tilde{\psi}^i$$

I will the use the notation $G(H)(\mathbf{a}(t), t) = \widetilde{(-iH)}(\mathbf{a}(t), t)$, to describe this operator (the Generator of time translation), which acts on the isomorphic quantum state vectors

$$\tilde{\psi} = \begin{pmatrix} \psi^{\mathrm{Re}} \\ \psi^{\mathrm{Im}} \end{pmatrix}$$

It can be shown that

$$G(H) = -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \otimes H^{\mathrm{Re}} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes H^{\mathrm{Im}}$$

where $\otimes$ is the Kronecker product. We then have the linear isomorphism dynamics equation:

$$\frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}t} = G(\mathbf{a}(t), t)\tilde{\psi}$$

where

$$G(\mathbf{a}(t), t) = G(H_{\mathrm{drift}}) + \sum_j a^j(t) G(H^j_{\mathrm{drive}})$$

The implicit dynamics constraint function $\mathbf{f}$ can be decomposed as follows:

$$\mathbf{f}(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{u}_t) = \begin{pmatrix} \mathbf{P}^{(m)}(\tilde{\psi}^1_{t+1}, \tilde{\psi}^1_t, \mathbf{a}_t) \\ \vdots \\ \mathbf{P}^{(m)}(\tilde{\psi}^n_{t+1}, \tilde{\psi}^n_t, \mathbf{a}_t) \\ \int \mathbf{a}_{t+1} - (\int \mathbf{a}_t + \mathbf{a}_t \cdot \Delta t_t) \\ \mathbf{a}_{t+1} - (\mathbf{a}_t + \mathrm{d}_t \mathbf{a}_t \cdot \Delta t_t) \\ \mathrm{d}_t \mathbf{a}_{t+1} - (\mathrm{d}_t \mathbf{a}_t + \mathbf{u}_t \cdot \Delta t_t) \end{pmatrix}$$

2.1 Padé integrators

We define (and implement) just the $m \in \{2, 4\}$ order Padé integrators $\mathbf{P}^{(m)}$:

$$
\begin{aligned}
\mathbf{P}^{(2)}(\tilde{\psi}^i_{t+1}, \tilde{\psi}^i_t, \mathbf{a}_t) &= \left(I - \frac{\Delta t}{2} G(\mathbf{a}_t)\right) \tilde{\psi}^i_{t+1} - \left(I + \frac{\Delta t}{2} G(\mathbf{a}_t)\right) \tilde{\psi}^i_t \\
\mathbf{P}^{(4)}(\tilde{\psi}^i_{t+1}, \tilde{\psi}^i_t, \mathbf{a}_t) &= \left(I - \frac{\Delta t}{2} G(\mathbf{a}_t) + \frac{(\Delta t)^2}{9} G(\mathbf{a}_t)^2\right) \tilde{\psi}^i_{t+1} \\
&\quad - \left(I + \frac{\Delta t}{2} G(\mathbf{a}_t) + \frac{(\Delta t)^2}{9} G(\mathbf{a}_t)^2\right) \tilde{\psi}^i_t
\end{aligned}
$$

Where again

$$
G(\mathbf{a}_t) = G_{\text{drift}} + \mathbf{a}_t \cdot \mathbf{G}_{\text{drive}}
$$

with $\mathbf{G}_{\text{drive}} = (G^1_{\text{drive}}, \ldots, G^c_{\text{drive}})^\top$, where $c = \mathsf{ncontrols}$

3 Differentiation

Our problem consists of $Z_{\text{dim}} = (\mathsf{nstates} + \mathsf{ncontrols}) \times \mathsf{T}$ total variables, arranged into a vector

$$\mathbf{Z} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{x}_T \\ \mathbf{u}_T \end{pmatrix} = \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_T \end{pmatrix} \tag{2}$$

where $\mathbf{z}_t = \begin{pmatrix} \mathbf{x}_t \\ \mathbf{u}_t \end{pmatrix}$ is referred to as a *knot point* and has dimension

$$z_{\text{dim}} = \mathsf{vardim} = \mathsf{nstates} + \mathsf{ncontrols}.$$

Also, as of right now, $\mathbf{u}_T$ is included in $\mathbf{Z}$ but is ignored in calculations.

3.1 Objective Gradient

Given the objective

$$J(\mathbf{Z}) = Q \sum_{i=1}^{n} \ell(\tilde{\psi}_T^i, \tilde{\psi}_{\text{goal}}^i) + \frac{R}{2} \sum_{t=1}^{T-1} \mathbf{u}_t^2 \tag{3}$$

we arrive at the gradient

$$\boldsymbol{\nabla}_{\mathbf{Z}} J(\mathbf{Z}) = \begin{pmatrix} \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_1 \\ \vdots \\ \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_t \\ \vdots \\ \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_{T-1} \\ Q \cdot \boldsymbol{\nabla}_{\tilde{\psi}^1} \ell^1 \\ \vdots \\ Q \cdot \boldsymbol{\nabla}_{\tilde{\psi}^n} \ell^n \\ \mathbf{0} \end{pmatrix} \tag{4}$$

where $\ell^i = \ell(\tilde{\psi}_T^i, \tilde{\psi}_{\text{goal}}^i)$. $\boldsymbol{\nabla}_{\tilde{\psi}^i} \ell^i$ is currently not calculated by hand, but at compile time via Symbolics.jl.

3.2 Dynamics Jacobian

Writing, $\mathbf{f}(\mathbf{z}_t, \mathbf{z}_{t+1}) = \mathbf{f}(\tilde{\psi}^i_{t+1}, \tilde{\psi}^i_t, \mathbf{a}_t)$, we can arrange the dynamics constraints into a vector

$$\mathbf{F} = \begin{pmatrix} \mathbf{f}(\mathbf{z}_1, \mathbf{z}_2) \\ \vdots \\ \mathbf{f}(\mathbf{z}_{T-1}, \mathbf{z}_T) \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_{T-1} \end{pmatrix} \tag{5}$$

where we have defined $\mathbf{f}_t = \mathbf{f}(\mathbf{z}_t, \mathbf{z}_{t+1})$.

The dynamics Jacobian matrix $\frac{\partial \mathbf{F}}{\partial \mathbf{Z}}$ then has dimensions

$$F_{\mathrm{dim}} \times Z_{\mathrm{dim}} = (f_{\mathrm{dim}} \cdot (T-1)) \times (z_{\mathrm{dim}} \cdot T)$$

This matrix has a block diagonal structure:

$$\frac{\partial \mathbf{F}}{\partial \mathbf{Z}} = \begin{pmatrix} \frac{\partial \mathbf{f}_1}{\partial \mathbf{z}_1} & \frac{\partial \mathbf{f}_1}{\partial \mathbf{z}_2} & & & & \\ & \ddots & \ddots & & & \\ & & \frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_t} & \frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_{t+1}} & & \\ & & & \ddots & \ddots & \\ & & & & \frac{\partial \mathbf{f}_T}{\partial \mathbf{z}_{T-1}} & \frac{\partial \mathbf{f}_T}{\partial \mathbf{z}_T} \end{pmatrix} \tag{6}$$

We just need the $f_{\mathrm{dim}} \times z_{\mathrm{dim}}$ Jacobian matrices $\frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_t}$ and $\frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_{t+1}}$.

$\mathbf{f}_t$ Jacobian expressions

With $\mathbf{P}^{(m),i}_t = \mathbf{P}^{(m)}(\tilde{\psi}^i_{t+1}, \tilde{\psi}^i_t, \mathbf{a}_t)$, we first have

$$\frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_t} = \begin{pmatrix} \ddots & & & & \vdots & & & & \\ & \frac{\partial \mathbf{P}^{(m),i}_t}{\partial \tilde{\psi}^i_t} & & & \frac{\partial \mathbf{P}^{(m),i}_t}{\partial \mathbf{a}_t} & & & & \\ & & \ddots & & \vdots & & & & \\ & & & -I_c^{\int \mathbf{a}_t} & -\Delta t I_c^{\mathbf{a}_t} & & & & \\ & & & & -I_c^{\mathbf{a}_t} & -\Delta t I_c^{\mathrm{d}_t \mathbf{a}_t} & & & \\ & & & & & \ddots & & \ddots & \\ & & & & & & & -I_c^{\mathrm{d}_t^{c-1} \mathbf{a}_t} & -\Delta t I_c^{\mathbf{u}_t} \end{pmatrix} \tag{7}$$

where, $c =$ ncontrols, and the diagonal dots in the bottom right indicate that the number of $-I_c$ blocks on the diagonal should equal augdim, which is set to 3 by default.

Lastly,

$$\frac{\partial \mathbf{f}_t}{\partial \mathbf{z}_{t+1}} = \begin{pmatrix} \frac{\partial \mathbf{P}_t^{(m),1}}{\partial \tilde{\psi}_{t+1}^1} & & & \\ & \ddots & & \\ & & \frac{\partial \mathbf{P}_t^{(m),n}}{\partial \tilde{\psi}_{t+1}^n} & \\ & & & I_{c\cdot\mathsf{augdim}} \end{pmatrix} \tag{8}$$

$\mathbf{P}_t^{(m),i}$ Jacobian expressions

For the $\tilde{\psi}^i$ components, we have, for $m = 2$,

$$\frac{\partial \mathbf{P}_t^{(2),i}}{\partial \tilde{\psi}_t^i} = -\left(I + \frac{\Delta t}{2} G(\mathbf{a}_t)\right) \tag{9}$$

$$\frac{\partial \mathbf{P}_t^{(2),i}}{\partial \tilde{\psi}_{t+1}^i} = I - \frac{\Delta t}{2} G(\mathbf{a}_t) \tag{10}$$

and, for $m = 4$,

$$\frac{\partial \mathbf{P}_t^{(4),i}}{\partial \tilde{\psi}_t^i} = -\left(I + \frac{\Delta t}{2} G(\mathbf{a}_t) + \frac{(\Delta t)^2}{9} G(\mathbf{a}_t)^2\right) \tag{11}$$

$$\frac{\partial \mathbf{P}_t^{(4),i}}{\partial \tilde{\psi}_{t+1}^i} = I - \frac{\Delta t}{2} G(\mathbf{a}_t) + \frac{(\Delta t)^2}{9} G(\mathbf{a}_t)^2. \tag{12}$$

Now, for the $\mathbf{a}_t$ components, we have, for $m = 2$,

$$\frac{\partial \mathbf{P}_t^{(2),i}}{\partial a_t^j} = \frac{-\Delta t}{2} G_{\text{drive}}^j \left(\tilde{\psi}_{t+1}^i + \tilde{\psi}_t^i\right) \tag{13}$$

and, for $m = 4$,

$$\frac{\partial \mathbf{P}_t^{(4),i}}{\partial a_t^j} = \frac{-\Delta t}{2} G_{\text{drive}}^j \left(\tilde{\psi}_{t+1}^i + \tilde{\psi}_t^i\right) + \frac{(\Delta t)^2}{9} \left\{G_{\text{drive}}^j, G(\mathbf{a}_t)\right\} \left(\tilde{\psi}_{t+1}^i - \tilde{\psi}_t^i\right) \tag{14}$$

where $\{A, B\} = AB + BA$ is the anticommutator.

3.3 Hessian of the Lagrangian

The Lagrangian function is defined to be

$$\mathscr{L}(\mathbf{Z}; \sigma, \boldsymbol{\mu}) = \sigma \cdot J(\mathbf{Z}) + \boldsymbol{\mu} \boldsymbol{\cdot} \mathbf{F}(\mathbf{Z}) \tag{15}$$

where $\boldsymbol{\mu}$ is a Z_{dim}-dimensional vector provided by the solver.

For the Hessian we have

$$\nabla^2 \mathscr{L} = \sigma \cdot \nabla^2 J + \boldsymbol{\mu} \boldsymbol{\cdot} \nabla^2 \mathbf{F}. \tag{16}$$

We will look at $\nabla^2 J$ and $\boldsymbol{\mu} \boldsymbol{\cdot} \nabla^2 \mathbf{F}$ separately.

Objective Hessian

With $\ell^i = \ell(\tilde{\psi}^i_T, \tilde{\psi}^i_{\text{goal}})$, we have

$$\nabla^2 J(\mathbf{Z}) = \begin{pmatrix} \ddots & & & & & & & \\ & \mathbf{0} & & & & & & \\ & & R_t I_c & & & & & \\ & & & \ddots & & & & \\ & & & & \ddots & & & \\ & & & & & Q \cdot \nabla^2 \ell^i & & \\ & & & & & & \ddots & \\ & & & & & & & \mathbf{0} \end{pmatrix} \tag{17}$$

where $\nabla^2 \ell^i$ is again calculated using Symbolics.jl.

Dynamics Hessian

With $\boldsymbol{\mu} = (\vec{\mu}_1, \ldots, \vec{\mu}_T)$, $\vec{\mu}_t = \left(\mu_t^1, \ldots, \mu_t^{z_{\text{dim}}}\right)$, and using

$$\vec{\mu}_t^{\tilde{\psi}^i} = \left(\mu_t^{(i-1)\cdot\tilde{\psi}_{\text{dim}}+1}, \ldots, \mu_t^{i\cdot\tilde{\psi}_{\text{dim}}}\right)$$

we have

$$\boldsymbol{\mu} \cdot \nabla^2 \mathbf{F} = \begin{pmatrix} & \vdots & & & & \\ & \left(\frac{\partial^2 \mathbf{P}_t^{(m),i}}{\partial \tilde{\psi}_t^i \partial a_t^j}\right)^\top \vec{\mu}_t^{\tilde{\psi}^i} & & & & \\ & \vdots & & & & \\ \ddots & \mathbf{0} & & & & \\ & \sum_{i=1}^{n} \vec{\mu}_t^{\tilde{\psi}^i} \cdot \frac{\partial^2 \mathbf{P}_t^{(4),i}}{\partial a_t^k \partial a_t^j} & \mathbf{0} & \ldots & \left(\vec{\mu}_t^{\tilde{\psi}^i}\right)^\top \frac{\partial^2 \mathbf{P}_t^{(m),i}}{\partial a_t^k \partial \tilde{\psi}_{t+1}^i} & \ldots \\ & & \ddots & & & \end{pmatrix} \tag{18}$$

with

$$\frac{\partial^2 \mathbf{P}_t^{(4),i}}{\partial a_t^k \partial a_t^j} = \frac{(\Delta t)^2}{9} \left\{G_{\text{drive}}^j, G_{\text{drive}}^k\right\} \left(\tilde{\psi}_{t+1}^i - \tilde{\psi}_t^i\right) \tag{19}$$

with, again, $\{A, B\} = AB + BA$, being the anticommutator.

since

$$x \cdot (Ay) = x^\top Ay = (A^\top x)^\top y$$

For the mixed partials we have:

$$\frac{\partial^2 \mathbf{P}_t^{(2),i}}{\partial \tilde{\psi}_t^i \partial a_t^j} = \frac{\partial^2 \mathbf{P}_t^{(2),i}}{\partial \tilde{\psi}_{t+1}^i \partial a_t^j} = -\frac{\Delta t}{2} G_{\text{drive}}^j \tag{20}$$

and

$$\frac{\partial^2 \mathbf{P}_t^{(4),i}}{\partial \tilde{\psi}_t^i \partial a_t^j} = -\frac{\Delta t}{2} G_{\text{drive}}^j - \frac{(\Delta t)^2}{9} \left(\left\{G_{\text{drive}}^j, G_{\text{drift}}\right\} + \mathbf{a}_t \cdot \left\{G_{\text{drive}}^j, \mathbf{G}_{\text{drive}}\right\}\right) \tag{21}$$

$$\frac{\partial^2 \mathbf{P}_t^{(4),i}}{\partial \tilde{\psi}_{t+1}^i \partial a_t^j} = -\frac{\Delta t}{2} G_{\text{drive}}^j + \frac{(\Delta t)^2}{9} \left(\left\{G_{\text{drive}}^j, G_{\text{drift}}\right\} + \mathbf{a}_t \cdot \left\{G_{\text{drive}}^j, \mathbf{G}_{\text{drive}}\right\}\right) \tag{22}$$