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 2.1 Padé integrators	2 3
3	Differentiation 3.1 Objective Gradient	5
4	Minimum Time Problem	9
	4.1 Objective Gradient	
	4.2 Dynamics Jacobian	10
	4.3 Hessian of the Lagrangian	11

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 constrained optimization problem

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

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

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

$$(1)$$

In summary,

$$\dim(\mathbf{x}_t) = \mathsf{nstates} = \mathsf{nqstates} * \mathsf{isodim} + \mathsf{ncontrols} * \mathsf{augdim}$$

 $\dim(\mathbf{u}_t) = \mathsf{ncontrols}$

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

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

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:

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

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

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

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

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

The implicit dynamics constraint function f can be decomposed as follows:

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

2.1 Padé integrators

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

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

$$\mathbf{P}^{(4)}(\tilde{\psi}_{t+1}^{i}, \tilde{\psi}_{t}^{i}, \mathbf{a}_{t}) = \left(I - \frac{\Delta t}{2}G(\mathbf{a}_{t}) + \frac{(\Delta t)^{2}}{9}G(\mathbf{a}_{t})^{2}\right)\tilde{\psi}_{t+1}^{i}$$

$$- \left(I + \frac{\Delta t}{2}G(\mathbf{a}_{t}) + \frac{(\Delta t)^{2}}{9}G(\mathbf{a}_{t})^{2}\right)\tilde{\psi}_{t}^{i}$$

Where again

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

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

3 Differentiation

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

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

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

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

$$\nabla_{\mathbf{Z}} J(\mathbf{Z}) = \begin{pmatrix} \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_{1} \\ \vdots \\ \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_{t} \\ \vdots \\ \mathbf{0}_{x_{\text{dim}}} \\ R \cdot \mathbf{u}_{T-1} \\ Q \cdot \nabla_{\tilde{\psi}^{1}} \ell^{1} \\ \vdots \\ Q \cdot \nabla_{\tilde{\psi}^{n}} \ell^{n} \\ \mathbf{0} \end{pmatrix}$$

$$(4)$$

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

3.2 Dynamics Jacobian

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

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

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

This matrix has a block diagonal structure:

$$\frac{\partial \mathbf{F}}{\partial \mathbf{Z}} = \begin{pmatrix}
\frac{\partial \mathbf{f}_{1}}{\partial \mathbf{z}_{1}} & \frac{\partial \mathbf{f}_{1}}{\partial \mathbf{z}_{2}} & & & \\
& \ddots & \ddots & & \\
& & \frac{\partial \mathbf{f}_{t}}{\partial \mathbf{z}_{t}} & \frac{\partial \mathbf{f}_{t}}{\partial \mathbf{z}_{t+1}} & & \\
& & & \ddots & \ddots & \\
& & & \frac{\partial \mathbf{f}_{T}}{\partial \mathbf{z}_{T-1}} & \frac{\partial \mathbf{f}_{T}}{\partial \mathbf{z}_{T}}
\end{pmatrix}$$
(6)

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

f_t Jacobian expressions

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

$$\frac{\partial \mathbf{f}_{t}}{\partial \mathbf{z}_{t}} = \begin{pmatrix}
\ddots & \vdots & \vdots \\
\frac{\partial \mathbf{P}_{t}^{(m),i}}{\partial \tilde{\psi}_{t}^{i}} & \frac{\partial \mathbf{P}_{t}^{(m),i}}{\partial \mathbf{a}_{t}} \\
& \ddots & \vdots \\
-I_{c}^{\int \mathbf{a}_{t}} & -\Delta t I_{c}^{\mathbf{a}_{t}} \\
& & -I_{c}^{\mathbf{d}_{t}} & -\Delta t I_{c}^{\mathbf{d}_{t}} \\
& & \ddots & \vdots \\
& & & -I_{c}^{\mathbf{d}_{t}^{c-1} \mathbf{a}_{t}} & -\Delta t I_{c}^{\mathbf{d}_{t}}
\end{pmatrix}$$

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,

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

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

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

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

3.3 Hessian of the Lagrangian

The Lagrangian function is defined to be

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

where μ is a Z_{dim} -dimensional vector provided by the solver.

For the Hessian we have

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

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

Objective Hessian

With $\ell^i = \ell(\tilde{\psi}_T^i, \tilde{\psi}_{\mathrm{goal}}^i)$, we have

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

Dynamics Hessian

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

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

we have

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

$$(18)$$

with

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

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

and

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

$$\frac{\partial^{2} \mathbf{P}_{t}^{(4),i}}{\partial \tilde{\psi}_{t+1}^{i} \partial a_{t}^{j}} = -\frac{\Delta t}{2} G_{\text{drive}}^{j} + \frac{(\Delta t)^{2}}{9} \left(\left\{ G_{\text{drive}}^{j}, G_{\text{drift}} \right\} + \mathbf{a}_{t} \cdot \left\{ G_{\text{drive}}^{j}, \mathbf{G}_{\text{drive}} \right\} \right)$$
(21)

4 Minimum Time Problem

Once a solution has been found for a given *time horizon*, we can solve the time minimization problem below, initialized with the given solution.

For this problem we will define, with ${\bf Z}$ as defined before

$$\Delta \mathbf{t} = \begin{pmatrix} \Delta t_1 \\ \vdots \\ \Delta t_{T-1} \end{pmatrix} \quad \text{and} \quad \mathbf{\bar{Z}} = \begin{pmatrix} \mathbf{Z} \\ \Delta \mathbf{t} \end{pmatrix}$$

4.1 Objective Gradient

Let's write the objective function as

$$J = J_{\Delta t} + J_u + J_s \tag{23}$$

then

$$\nabla_{\mathbf{\bar{Z}}}J = \begin{pmatrix} \nabla_{\mathbf{Z}}J \\ \nabla_{\Delta \mathbf{t}}J \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{Z}}J_u + \nabla_{\mathbf{Z}}J_s \\ \nabla_{\Delta \mathbf{t}}J_{\Delta t} \end{pmatrix}$$
(24)

where

$$\nabla_{\mathbf{Z}} J_u = \begin{pmatrix} \vdots \\ \mathbf{0} \\ R_u \mathbf{u}_t \\ \vdots \end{pmatrix}$$
 (25)

and

$$\nabla_{\mathbf{Z}} J_{s} = \begin{pmatrix} \mathbf{0} \\ -R_{s}(\mathbf{u}_{2} - \mathbf{u}_{1}) \\ \mathbf{0} \\ R_{s}(\mathbf{u}_{2} - \mathbf{u}_{1}) - R_{s}(\mathbf{u}_{3} - \mathbf{u}_{2}) \\ \vdots \\ \mathbf{0} \\ R_{s}(\mathbf{u}_{t} - \mathbf{u}_{t-1}) - R_{s}(\mathbf{u}_{t+1} - \mathbf{u}_{t}) \\ \vdots \\ \mathbf{0} \\ -R_{s}(\mathbf{u}_{t-1} - 2\mathbf{u}_{t} + \mathbf{u}_{t+1}) \\ \vdots \\ \mathbf{0} \\ -R_{s}(\mathbf{u}_{t-1} - 2\mathbf{u}_{t} + \mathbf{u}_{t+1}) \end{pmatrix}$$

$$\vdots$$

$$\vdots$$

$$\mathbf{0} \\ -R_{s}(\mathbf{u}_{t-1} - 2\mathbf{u}_{t} + \mathbf{u}_{t+1})$$

$$\vdots$$

$$\mathbf{0} \\ -R_{s}(\mathbf{u}_{t-1} - 2\mathbf{u}_{t} + \mathbf{u}_{t+1}) \end{pmatrix}$$

$$(26)$$

4.2 Dynamics Jacobian

We then have, with **F** defined as before (but taking the corresponding Δt_t):

$$\frac{\partial \mathbf{F}}{\partial \bar{\mathbf{Z}}} = \begin{pmatrix} \frac{\partial \mathbf{F}}{\partial \mathbf{Z}} & \frac{\partial \mathbf{F}}{\partial \Delta \mathbf{t}} \end{pmatrix} \tag{27}$$

where

$$\frac{\partial \mathbf{F}}{\partial \Delta \mathbf{t}} = \begin{pmatrix} \frac{\partial \mathbf{f}_1}{\partial \Delta t_1} & & \\ & \ddots & \\ & & \frac{\partial \mathbf{f}_{T-1}}{\partial \Delta t_{T-1}} \end{pmatrix}$$
(28)

4.3 Hessian of the Lagrangian

As before we will first define the objective Hessian and then the dynamics Lagrangian Hessian

Objective Hessian