Quantum Control via Discrete-Time Optimal Control

State-Action-Frequency Constrained Pontryagin Maximum Principle

Aneesh Anand Kamat

Indian Institute of Technology Bombay

aneesh.kamat@iitb.ac.in

Supervised Learning Project

November 9, 2023

Presentation Overview

Two-level Quantum System

- Consider a π pulse control driving a two-level quantum system from $|0\rangle$ to $e^{i\phi}|1\rangle$:

$$i\dot{\psi} = \begin{pmatrix} E_0 & \Omega(t) \\ \Omega^*(t) & E_1 \end{pmatrix} \psi,$$

- E_0 and E_1 are energies of $|0\rangle$ and $|1\rangle$, $\Omega(t)$ is the complex external field:

$$\Omega(t) = u(t)e^{i(E_1 - E_0)t}.$$

- Time-dependent change of variables to a rotating frame gives us

$$i\dot{\tilde{\psi}} = \begin{pmatrix} 0 & u(t) \\ u(t) & 0 \end{pmatrix} \tilde{\psi}.$$

Coupled Equations

- Let $c_1 = x_1 + iy_1$ and $c_2 = x_2 + iy_2$ be complex coordinates of $\tilde{\psi}$
- The Schrödinger equation is now equivalent to the coupled equations:

$$\dot{x}_1 = uy_2, \quad \dot{y}_1 = -ux_2,$$
$$\dot{x}_2 = uy_1, \quad \dot{y}_2 = -ux_1.$$

- The states correspond to $y_2 = \pm 1$ and the speed of evolution scales with the control amplitude $u \Rightarrow$ larger amplitudes, lesser transition time.
- Constraints and costs induce a need to optimize, bounding the amplitude allows for an optimal solution (albeit trivially constant at maximum amplitude)

Experimental Constraints

All controllable parameters have practical limitations: power, frequency, timing, etc.

- DACs have a finite rise-time, so pulses should start and end at 0.
- Fixed bandwidth operation due to complex control electronics.
- Amplitude is bounded by power constraints of the equipment.
- Particular frequencies within the bandwidth that we want to avoid.

Dynamics

- A finite-dimensional control system is a dynamical system governed by

$$\dot{q}(t) = f[q(t), u(t)], \tag{1}$$

  where $q : I \to M$ represents the state of the system, I is an interval in $\mathbb{R}$, and M is a smooth manifold, e.g., Bloch sphere for a qubit.
- The control law is $u : I \to U \subset \mathbb{R}^m$ and f is a smooth function such that $f(\cdot, \bar{u})$ is a vector field on M for every $\bar{u} \in U$.
- Piecewise continuous controls form a subset of *admissible controls* (regular enough), and are the only control laws reasonably applicable on experiment.

Time Evolution of Closed Quantum System

- The Schrödinger equation

$$i\dot{\psi}(t) = \left(H_0 + \sum_{j=1}^{m} u_j(t) H_j \right) \psi(t), \tag{2}$$

  - where ψ, the *wave function*, belongs to the unit sphere in $\mathbb{C}^N$,
  - $H_0, \ldots, H_m$ are $N \times N$ Hermitian matrices
  - the control parameters $u_j(t) \in \mathbb{R}$ are the components of the control $u(\cdot)$
- This control problem has the form of Eq. 1 with $n = 2N - 1$, $M = S^{2N-1} \subset \mathbb{C}^N$, $q = \psi$, and $f(\psi, u) = -i(H_0 + \sum_j^m u_j H_j)\psi$.
- The uncontrolled part - the H_0 term - is called the *drift* Hamiltonian.

Propagator Evolution

- The solution can also be expressed in terms of the unitary operator $\mathbf{U}(t, t_0)$, connecting the wave function at time t_0 to its value at t :

$$\psi(t) = \mathbf{U}(t, t_0)\psi(t_0).$$

- The *propagator* $\mathbf{U}(t, t_0)$ also satisfies the Schrödinger equation

$$i\dot{\mathbf{U}}(t, t_0) = \left(H_0 + \sum_{j=1}^{m} u_j(t) H_j \right) \mathbf{U}(t, t_0) \tag{3}$$

  with initial condition $\mathbf{U}(t_0, t_0) = \mathbb{I}_N$.
- In quantum computing, the control problem usually concerns the propagator $\mathbf{U}$.
- Eq. 3 has the form of Eq. 1 with $M = U(N) \subset \mathbb{C}^{N \times N}$ and $q = \mathbf{U}$.

Steps to solve an OCP

Note, without proof, a closed quantum system is controllable if the matrix Lie algebra generated by the matrices $H_0, \ldots, H_m$ is SU(N).

1. Find conditions that guarantee the existence of solutions.
2. Apply first-order necessary conditions.
3. Selection of the best solution among all candidates.

Statement of the PMP

Consider the optimal control problem

$$\dot{q}(t) = f[q(t), u(t)],$$
$$q(0) = q_{\text{in}}, \quad q(T) \in \mathcal{T},$$
$$\int_0^T f^0[q(t), u(t)]dt + \phi[q(T)] \to \min,$$

where

- *M is a smooth manifold of dimension n, $U \subset \mathbb{R}^m$,*
- *$\mathcal{T}$ is a (non-empty) smooth submanifold of M; it can be reduced to a point (fixed terminal point) or coincide with M (free terminal point),*
- *f, f^0 are smooth and $u \in \mathcal{U}$,*
- *$q : [0, T] \to M$ is a continuous curve.*

The (Pre-)Hamiltonian

Define the function (called the pre-Hamiltonian)

$$\mathcal{H}(q, p, u, p^0) = \langle p, f(q, u) \rangle + p^0 f^0(q, u) \tag{4}$$

with

$$(q, p, u, p^0) \in T^*M \times U \times \mathbb{R}$$

If the pair $(q, u) : [0, T] \to M \times U$ *is optimal then there exists a never vanishing continuous pair* $(p, p^0) : [0, T] \ni t \mapsto [p(t), p^0] \in T^*_{q(t)}M \times \mathbb{R}$, *where* $p^0 \leq 0$ *is a constant, such that, for almost every* $t \in [0, T]$, *we have:*

Conditions

1. *q satisfies the Hamiltonian equation $\dot{q}(t) = \frac{\partial \mathcal{H}}{\partial p}[q(t), p(t), u(t), p^0]$;*
2. *p satisfies the Hamiltonian equation $\dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial q}[q(t), p(t), u(t), p^0]$;*
3. *the quantity $\mathcal{H}_M[q(t), p(t), p^0] := \max_{v \in U} \mathcal{H}[q(t), p(t), v, p^0]$ is well defined and*

$$\mathcal{H}[q(t), p(t), u(t), p^0] = \mathcal{H}_M[q(t), p(t), p^0],$$

   which corresponds to the maximization condition.
   Moreover,
4. *there exists a constant $c \geq 0$ such that $\mathcal{H}_M[q(t), p(t), p^0] = c$ on $[0, T]$, with $c = 0$ if the final time is free (value of the Hamiltonian);*
5. *for every $v \in T_{q(T)}\mathcal{T}$, we have $\langle p(T), v\rangle = p^0 \langle d\phi[q(T)], v\rangle$ (transversality condition), where $d\phi$ is the differential of the function ϕ.*

A few remarks

- The covector p is called the *adjoint state* while p^0 is the *abnormal multiplier*. The quantities $p(\cdot)$ and p^0 play the roles of Lagrange multipliers.
- A trajectory $q(\cdot)$ for which there exist $p(\cdot)$, $u(\cdot)$, and p^0 such that $[q(\cdot), p(\cdot), u(\cdot), p^0]$ satisfies all the conditions given by the PMP is called an *extremal trajectory*.
- The 4-uple $[q(\cdot), p(\cdot), u(\cdot), p^0]$ is called an *extremal* or, equivalently, an *extremal lift of* $q(\cdot)$. Such an extremal is called *normal* if $p^0 \neq 0$ and *abnormal* if $p^0 = 0$.
- The PMP is only a necessary condition for optimality. Not all (if any) extremal trajectories are optimal, the step of existence must be verified.

Use of the PMP

Applying the PMP is not straightforward as many conditions need to be satisfied that are coupled with each other. In practice:

1. Use the maximization condition (3.) to express, when possible, the control as a function of the state and of the covector, i.e., $u = w(q, p)$.
2. Insert the control found in the previous step into the Hamiltonian equations
3. Find p_{in} such that:
$$q(T; p_{\text{in}}, p^0) \in \mathcal{T}. \tag{5}$$
4. If Eq. 5 has a unique solution p_{in} and if we have *a priori* verified the existence of an optimal solution, then the optimal control problem is solved!

Current QOC Algorithms

Several quantum optimal control protocols have been developed in recent decades to address various challenges in this field.

One of the most ubiquitous is GRadient Ascent Pulse Engineering (GRAPE)

- It is a first-order gradient-based optimization algorithm
- It can be derived from the necessary conditions of the PMP for a fixed control time without constraints on the final state and control

A recent algorithm is the Gradient Optimization of Analytic conTrols (GOAT)

- uses control ansatzes to find pulses described by only a few parameters and has an efficient computation of the gradient using the chain rule
- It incorporates experimental constraints using bounding functions and window functions for amplitude constraints and smooth start and finish.

Consideration of Constraints

As outlined earlier, several experimental constraints must be included for a successful quantum optimal control algorithm.

- Ideally, these should be included right at the synthesis stage as compared to using filter functions etc. after the pulse has been designed.
- Such post-processing will inevitably deviate from the optimal trajectory and induce errors that may be detrimental in highly sensitive nonlinear systems.
- Paruchuri, Pradyumna et al. (2020) presented ”A Frequency-Constrained Geometric Pontryagin Maximum Principle on Matrix Lie Groups”.
- This incorporates state-action-frequency constraints right at synthesis, allowing particular frequencies to be filtered out in addition to bandwidth constraints.

Problem Setup

Fix a positive integer N that plays the role of a time horizon, and for $\mathcal{N} \in \mathbb{N}^*$, set $[\mathcal{N}] := \{0, \ldots, \mathcal{N} - 1\}$ and $[\mathcal{N}]^* := [\mathcal{N}] \setminus \{0\}$. Let G be a matrix Lie group with $\mathfrak{g}$ its Lie algebra. Consider a controlled discrete-time system evolving partly on a fixed matrix Lie group G and partly on $\mathbb{R}^d$, given by

$$\begin{cases} q_{t+1} &= q_t s_t(q_t, x_t) \\ x_{t+1} &= f_t(q_t, x_t, u_t) \end{cases} \quad \text{for} \quad t \in [N]^*, \tag{6}$$

where $(q_t, x_t) \in G \times \mathbb{R}^d$ is the vector of states and $u_t \in \mathbb{R}^m$ the vector of control actions of the system at a discrete-time instant t. These two maps are smooth:

- $s_t : G \times \mathbb{R}^d \to G$ (dynamics of the states q_t on the matrix Lie group G), and
- $f_t : G \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ (dynamics of state x_t in $\mathbb{R}^d$)

Frequency Constraints

Let $\mathbb{R}^N \ni u^{(k)} := (u_t^{(k)})_{t=0}^{N-1}$ be the trajectory of the k^{th} component of the control. The subscript on u denotes the stage and the superscript denotes the component of the control. The hat on top of a variable denotes its frequency representation. The discrete Fourier transform (DFT) of $u^{(k)}$ is defined by

$$\mathbb{C}^N \ni \widehat{u^{(k)}} := Fu^{(k)} \quad \text{for } k = 1, \dots, m,$$

$$\text{where} \quad F := \frac{1}{\sqrt{N}} \left(\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \cdots & \omega^{(N-1)(N-1)} \end{bmatrix} \right) \in \mathbb{C}^{N \times N}$$

for $\omega := e^{\frac{-i2\pi}{N}}$.

DFT of Control Trajectory

Let u denote the stacked vector $\left((u^{(1)})^\top \dots (u^{(m)})^\top\right)^\top$, and define the DFT of a control trajectory by the vector

$$\mathbb{C}^{mN} \ni \hat{u} := \left(\begin{bmatrix} \widehat{u^{(1)}} \\ \vdots \\ \widehat{u^{(m)}} \end{bmatrix}\right) = \left(\begin{bmatrix} Fu^{(1)} \\ \vdots \\ Fu^{(m)} \end{bmatrix}\right) = \mathcal{F}\left(\begin{bmatrix} u^{(1)} \\ \vdots \\ u^{(m)} \end{bmatrix}\right),$$

where $\mathcal{F}$ is a block diagonal matrix with the standard DFT matrix F being each block. Note here that $(\widehat{u^{(k)}})_j \in \mathbb{C}$ represents the $(2\pi(j-1)/N)^{\text{th}}$ frequency component of the trajectory $u^{(k)}$. Hence, if elimination of the $(2\pi(j-1)/N)^{\text{th}}$ frequency component of $u^{(k)}$ is desired, it can be ensured by introducing the constraint

$$0 = (\widehat{u^{(k)}})_j = F_j u^{(k)},$$

where F_j is the j^{th} row of the DFT matrix defined above.

The Effective Constraint

Hence, in general, control frequency constraints can be enforced by a collection of affine equality conditions in the control action variables, and this is represented abstractly by one equality constraint

$$\sum_{t=0}^{N-1} \widetilde{F}_t u_t = 0 \quad \text{where } \widetilde{F}_t \text{ are suitable matrices.} \tag{7}$$

It is imperative to point out here that how frequency constraints have been assimilated into the problem formulation enables the designer to cancel particular frequencies in the control inputs, a feature distinctly absent in other control synthesis schemes.

Constrained OCP

Collecting the definitions above, our constrained optimal control problem in discrete-time is:

$$\begin{aligned} \underset{(u_t)_{t=0}^{N-1}}{\text{minimize}} \quad & \textstyle\sum_{t=0}^{N-1} c_t(q_t, x_t, u_t) + c_N(q_N, x_N) \\ \text{subject to} \quad & \begin{cases} \text{dynamics (6)}, \\ u_t \in \mathbb{U}_t \quad \text{for each } t \in [N], \\ \varphi_t(q_t, x_t) \leqslant 0 \text{ for each } t \in [N+1]^* \\ (q_0, x_0) = (\bar{q}, \bar{x}), \\ F(u_0, \ldots, u_{N-1}) = 0, \end{cases} \end{aligned} \tag{8}$$

Some remarks

1. $(\bar{q}, \bar{x}) \in G \times \mathbb{R}^d$ and $N \in \mathbb{N}$ are fixed;
2. the maps $c_t : G \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ (for each $t \in [N]$ defining the cost-per-stage) and $c_N : G \times \mathbb{R}^d \to \mathbb{R}$ (accounting for the final stage cost) are smooth;
3. the maps $\varphi_t : G \times \mathbb{R}^d \to \mathbb{R}^{n_t}$ for $t \in [N+1]^*$ denote constraints on the states and are smooth;
4. the set of admissible control actions $\mathbb{U}_t \subset \mathbb{R}^m$ is convex and compact for each $t \in [N]$;
5. the linear map $\mathbb{R}^{mN} \ni (u_0, \ldots, u_{N-1}) \mapsto F(u_0, \ldots, u_{N-1}) := \sum_{t=0}^{N-1} \tilde{F}_t u_t \in \mathbb{R}^\ell$ represents constraints on the frequency components of the control profile $(u_t)_{t=0}^{N-1}$.

Solving the OCP

- In spirit, the first-order necessary conditions for optimality are similar to the classical Euler's necessary conditions for optimality (that states that the gradient of a smooth function defined on an open set must vanish at an extremum point).
- Numerical algorithms are thereafter needed to arrive at optimal solutions starting from the necessary conditions given by the PMP.
- Our problem of finding a solution of the OCP characterized by the PMP can be reduced to finding a zero of a *nonlinear and implicit function*.
- We use shooting techniques to solve this root-finding problem posed by the resulting two-point boundary value problem distilled from the PMP.

Newton-Raphson

- The algorithms typically employed in computing a zero of a nonlinear map Φ are based on the Newton-Raphson (NR) iterative scheme and continuation methods.
- The NR iterates start with the intention of finding a zero of the first-order approximation of Φ near a zero ζ of Φ, i.e., from the affine map

$$z' \mapsto \Phi(z) + \Phi'(z)(z' - z)$$

for z, z' sufficiently close to ζ, leading to the recursion

$$z_{k+1} = z_k - \Phi'(z_k)^{-1}\Phi(z_k) \quad \text{for } k = 0, 1, \ldots$$

- Under standard hypotheses the sequence $(z_k)_{k\in\mathbb{N}}$ of iterates converges to ζ.
- If the NR scheme *does* converge, then it converges *quadratically*!

Requirements for NR Scheme

The effectiveness of the NR scheme is highly dependent on:

- the map Φ being sufficiently smooth,
- the availability of a good initial guess of the joint state-adjoint variables at one of the boundary points of the interval,
- the need for the derivative of Φ to be invertible everywhere sufficiently close to a zero of Φ, and
- the accuracy of the numerical computation of the derivative of Φ via finite-difference schemes.

Robbins-Monro

- The need for a derivative-free root-finding algorithm brings us to the Robbins-Monro (RM) scheme, also known as Stochastic Approximation (SA).
- This only relies on the ability to evaluate the function at given points.
- Its popularity can be attributed to the pervasiveness of ‘noise’ or randomness in engineering systems and the scheme’s adaptability to work in noisy situations.
- Its *incremental* nature has several advantages including lower per-iterate computation and memory requirements.
- SA ‘adapts’ to the needs of the problem.

RM Iterations

- Consider the problem of finding the root(s) of a nonlinear function $h : \mathbb{R}^d \mapsto \mathbb{R}^d$ given only noisy measurements.
- It can be considered as a black box that on input $x \in \mathbb{R}^d$ gives out $h(x) + \text{noise}$.
- The Robbins-Monro scheme is the d-dimensional iteration

$$x_{n+1} = x_n + a(n)[h(x_n) + M_{n+1}], \tag{9}$$

  initiated at some x_0.
- Here, the 'noise' $\{M_n\}$ is considered to be uncorrelated with the past which is intuitively consistent.
- Note here that this iteration scales only *linearly* with the problem size!

RM Steps

- The master stroke of Robbins and Monro was to choose the step-size sequence $a(n) > 0$ such that

$$\sum_n a(n) = \infty, \quad \sum_n a(n)^2 < \infty. \tag{10}$$

- This can be understood as having a 'fat tail' and a finite energy.
- The iteration (9) can be considered a noisy discretization (or 'Euler scheme') for

$$\dot{x}(t) = h(x(t)) \tag{11}$$

with decreasing step-sizes $\{a(n)\}$ and noise $\{M_n\}$.

Asymptotic convergence

- One can show that with probability one, x_n as $n \to \infty$ will have the same asymptotic behaviour as (11).
- (10) allows us to treat $\{a(n)\}$ as discrete time steps and the entire time axis is covered when we track the asymptotic behaviour of (11) as time tends to infinity.
- This also ensures that the errors due to discretization and noise are asymptotically negligible because it is ensured that $a(n)$ decreases to zero at a certain minimum rate.
- Note that $a(n)$ is both the discretization step *and* a weight for the noise at time n.
- SA 'generalizes' the strong law of large numbers: it 'averages out' the noise.

The best of both!

- A recent recursive algorithm that combines the SA and NR schemes to find a zero of Φ has been proposed. This hybrid algorithm combines some of the best features of both:
  - the exploration of space to find a zero, the ability to progress without derivative computations, etc. of the SA algorithm, with
  - the fast (quadratic) rate of convergence of the NR scheme.
- As the NR iterates must converge quadratically, it is clear after only a few observations of its iterates whether they show signs of convergence.

The Hybrid Protocol

1. The SA algorithm is first employed to converge sufficiently close to a zero of Φ. This serves an exploratory purpose to find a suitable neighbourhood of a zero of Φ to settle down and provide a warm start for the next step.
2. Switch to the NR scheme (or a suitable variant) with the final iterate of the SA algorithm being the initial condition of the NR iterations.
3. If these iterates indeed converge, continue with the NR iteration to obtain a zero. Otherwise, simply revert to the SA algorithm in Step 1. above, continue with the iterations with a smaller threshold of error, and repeat until convergence.

A Brief Summary

In this Supervised Learning Project, I have:

- Formulated Quantum Optimal Control problems
- Introduced the Pontryagin Maximum Principle
- Motivated the need for better consideration of constraints
- Discussed the problem set up for Discrete-Time PMP
- Presented gradient, stochastic and hybrid shooting techniques

Future Plans

Over the coming months, we plan to:

- Implement the frequency-constrained discrete PMP on standard quantum OCPs
- Compare the performance of the three shooting techniques (gradient, stochastic, hybrid)
- If simulations are successful, test the protocol at the QuMaC lab, TIFR Mumbai on real superconducting qubits

References

- Boscain, U., M. Sigalotti, and D. Sugny (Sept. 2021). “Introduction to the Pontryagin Maximum Principle for Quantum Optimal Control”.
- Paruchuri, Pradyumna et al. (2020). “A Frequency-Constrained Geometric Pontryagin Maximum Principle on Matrix Lie Groups”.
- Kumar, Yogesh et al. (2022). “Sparse optimal control problems with intermediate constraints: Necessary conditions”.
- Borkar, Vivek S. (Dec. 2013). “Stochastic Approximation”.
- Hesse, Helke Karen (2008). “Multiple Shooting and Mesh Adaptation for PDE Constrained Optimization Problems”. PhD thesis.
- Wilhelm, Frank K. et al. (Mar. 2020). An Introduction into Optimal Control for Quantum Technologies.

Acknowledgements

- I would like to thank my supervisor Prof. Debasish Chatterjee for his guidance, resourcefulness and patience throughout this project.
- I would also like to thank Siddhartha Ganguly and Pradyumna Paruchuri for helpful discussions and resources.