An Indirect Shooting Method for Stochastic Trajectory Optimization

Thomas Lew[1]

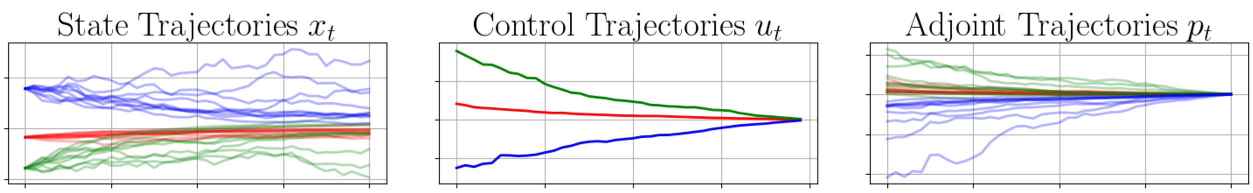

Fig. 1: Solution to an uncertainty-aware trajectory optimization problem computed using an indirect shooting method [1].

***Abstract*— We propose the first indirect shooting method for continuous-time stochastic trajectory optimization. This method leverages a new Pontryagin Maximum Principle (PMP) derived using rough path theory. In contrast to previous optimality conditions, this PMP does not rely on forward-backward stochastic differential equations, which is the key to unlock a practical indirect shooting method. We show that it converges 10× faster than a direct method on an example, thanks to exploiting the low-dimensional structure of solutions encoded in the PMP.**

I. Introduction

Accounting for uncertainty in the decision-making stack is key to achieving reliable robotics autonomy in complicated environments. In particular, algorithms for trajectory optimization under uncertainty underpin a wide range of applications [2]. Such methods typically solve stochastic trajectory optimization problems (**TOP**) of the form

$$\min_{(x,u)} \quad \mathbb{E}\left[\int_0^T \ell(x_t,u_t)\,\mathrm{d}t\right] \qquad (\textbf{TOP})$$

$$\text{s.t.} \quad \mathrm{d}x_t = b(x_t,u_t)\mathrm{d}t + \sigma(x_t)\mathrm{d}B_t, \quad t\in[0,T], \quad (1)$$

where (1) is a stochastic differential equation (SDE) in Stratonovich form, the states x_t are uncertain due to disturbances modeled as a Brownian motion B, and the control inputs u_t are open-loop. The dimensions of the state and control input are denoted by n and m.

Direct and indirect methods are two approaches for solving trajectory optimization problems. For deterministic problems, indirect methods typically converge faster to higher-accuracy solutions than direct methods, but they are more sensitive to the choice of initial guess. In the stochastic case, direct methods for solving **TOP** exist. However, there is currently no indirect method for solving **TOP**. Indeed, existing optimality conditions for solutions to **TOP** that an indirect method would rely on are formulated with forward-backward SDEs (FBSDE) (see e.g. [3], [4]) that introduce greater algorithmic and computational complexity if the dynamics are nonlinear.

[1] Toyota Research Institute, Los Altos, CA, USA

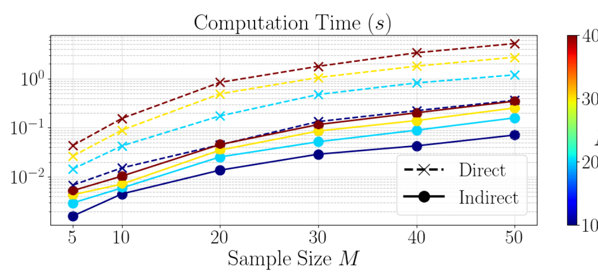

Fig. 2: Computation times to solve an instance of **TOP** using a direct and indirect method.

II. New Optimality Conditions

The following result provides new first-order optimality conditions for solutions to **TOP**. A more detailed formulation and a proof of the result are in [1].

*Theorem (Pontryagin Maximum Principle (**PMP**))*

Let (x,u) be an optimal solution to **TOP** and define the Hamiltonian $H(x,u,p) := p^\top b(x,u) - \ell(x,u)$.

Then, there exists a stochastic process p called *adjoint vector* starting from some random initial conditions p_0 such that:

i) **Adjoint equation**: the adjoint vector p solves the random rough differential equation (RDE)

$$\mathrm{d}p_t = -\frac{\partial H}{\partial x}(x_t,u_t,p_t)\mathrm{d}t - \frac{\partial \sigma}{\partial x}(x_t)^\top p_t \mathrm{d}B_t.$$

ii) **Transversality condition**: with probability one,

$$p_T = 0.$$

iii) **Maximality condition**: the optimal control input maximizes the average value of the Hamiltonian:

$$u_t = \operatorname*{arg\,max}_{v\in U} \mathbb{E}\left[H(x_t,v,p_t)\right], \quad t\in[0,T].$$

The similarity to the deterministic PMP is striking! The Hamiltonian H is unchanged, the adjoint equation is interpreted pathwise and has the same drift term $-\frac{\partial H}{\partial x}$, the transversality condition is identical, and the only difference in the maximality condition is an average. The reader is invited to compare **PMP** with [5, Theorem 2].

Importantly, the adjoint equation is *not* an FBSDE like in standard PMPs, but a random RDE defined pathwise using *rough path theory* [6]–[8][1]. This is the key to unlock a practical indirect method.

III. A New Indirect Shooting Method

PMP informs the design of a new indirect method for nonlinear stochastic trajectory optimization. That is, if we approximate all averages in **TOP** and **PMP** using Monte Carlo estimates for a sample size M, the search for approximate solutions to **TOP** consists of finding the initial values of the adjoint vector $(p_0^i)_{i=1}^M \in \mathbb{R}^{Mn}$ such that the transversality condition $(p_T^i)_{i=1}^M = 0$ holds. If the maximality condition gives a closed-form expression of the control u_t^M as a function of $(x_t^i, p_t^i)_{i=1}^M$ (which is often the case for control-affine systems), then we can use a root-finding Newton method to efficiently find a solution $(p_0^i)_{i=1}^M$ satisfying the transversality condition:

Algorithm (Indirect Shooting Method)

Parameters: Sample size M, tolerance ϵ
Inputs: Initial guess for $(p_0^i)_{i=1}^M$, samples $(B^i)_{i=1}^M$
While $\|F\|_\infty < \epsilon$:

1) Compute the transversality condition error and its Jacobian with respect to $(p_0^i)_{i=1}^M$

$$F := (p_T^1, \ldots, p_T^M)$$
$$J := \nabla_{(p_0^1,\ldots,p_0^M)}(p_T^1, \ldots, p_T^M)$$

by integrating the coupled differential equation

$$\begin{cases} \mathrm{d}x_t^i = b(x_t^i, u_t^M)\mathrm{d}t + \sigma(x_t^i)\mathrm{d}B_t^i, \\ \mathrm{d}p_t^i = -\frac{\partial H}{\partial x}(x_t^i, u_t^M, p_t^i)\mathrm{d}t - \frac{\partial \sigma}{\partial x}(x_t^i)^\top p_t^i \mathrm{d}B_t^i, \\ \text{with } u_t^M = \arg\max_{v \in U} \frac{1}{M}\sum_{i=1}^M H(x_t^i, v, p_t^i) \end{cases}$$

with initial conditions $((x_0, p_0^1), \ldots, (x_0, p_0^M))$.

2) Update $(p_0^i)_{i=1}^M$ by taking a Newton step

$$(p_0^1, \ldots, p_0^M) \leftarrow (p_0^1, \ldots, p_0^M) - J^{-1}F.$$

Return the control trajectory u^M by integrating the coupled differential equation above from $(p_0^i)_{i=1}^M$.

This approach is commonly known as an indirect shooting method in deterministic trajectory optimization [5] and is a natural extension to the stochastic setting using **PMP** and a Monte Carlo approximation [9]–[11]. To our knowledge, this method has not appeared in the literature yet, as previous optimality conditions rely on FBSDEs that introduce greater complexity. Indeed, previous indirect methods use deep learning to solve the FBSDEs from the classical stochastic PMP [12]–[14], whereas this indirect method does not require training a neural network and uses a Newton method instead.

IV. Results

We evaluate the method on a stabilization task for a system with nonlinear rigid body dynamics. The cost is a standard quadratic $\ell(x, u) = x^\top Q x + u^\top R u$. As a baseline, we use the direct method in [11] that solves a Monte Carlo reformulation of **TOP** via sequential quadratic programming. Details are in [1] and code is available at `github.com/ToyotaResearchInstitute/rspmp`.

Results in Figure 1 (right) show that the adjoint vector trajectories p^i start from different initial conditions p_0^i and are all zero at the final time ($p_T^i = 0$) to satisfy the transversality condition of **PMP**. Also, results in Figure 2 show that the indirect method is about $10\times$ faster than the direct method, thanks to leveraging the structure of the problem encoded in PMP to optimize over only the Mn variables p_0^i for the indirect method, versus optimizing over the $N(Mn+m)$ variables $(x_{k\Delta t}^i, u_{k\Delta t}^i)$ for the direct method, where N is the number of timesteps used to discretize the differential equation in **TOP**.

However, indirect methods typically have higher numerical sensitivity to the choice of initial guess. This tradeoff is well-known in the deterministic trajectory optimization literature, motivating the development of multiple shooting and homotopy methods [5], [15] for uncertainty-aware trajectory optimization. For further results on feedback optimization, we refer to [1].

V. Conclusion and Outlook

The optimality conditions in **PMP** provide new insights onto the structure of solutions to stochastic trajectory optimization problems. The main motivation for deriving **PMP** is the development of new algorithms for uncertainty-aware control and planning that can more easily borrow ideas from the deterministic optimal control literature, such as indirect shooting methods.

In the recent years, robotics research has shown that training policies *offline* via reinforcement learning by heavily randomizing the environment can yield robust controllers. Designing new trajectory optimization and MPC algorithms to solve similar uncertainty-aware problems *online* would potentially also yield robust controllers, but without the need to learn a policy offline. To do so, leveraging low-dimensional characterizations of solutions such as **PMP** and recent advances in numerical optimization are promising avenues for research.

[1] The adjoint equation cannot be interpreted as an Itô or Stratonovich SDE, because p_0 depends on the entire path of B. Using rough path theory to derive a stochastic PMP is new and the key to avoid FBSDEs.

Acknowledgements: We thank Marcus Greiff for his helpful feedback on this extended abstract.

References

[1] T. Lew, "Rough stochastic Pontryagin maximum principle and an indirect shooting method," 2025, available at https://arxiv.org/abs/2502.06726.

[2] T. Lew, R. Bonalli, and M. Pavone, "Risk-averse trajectory optimization via sample average approximation," *IEEE Robotics and Automation Letters*, vol. 9, no. 2, pp. 1500–1507, 2024.

[3] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian systems and HJB equations*. Springer New York, 1999.

[4] R. Bonalli and B. Bonnet, "First-order Pontryagin maximum principle for risk-averse stochastic optimal control problems," *SIAM Journal on Control and Optimization*, vol. 61, no. 3, pp. 1881–1909, 2023.

[5] E. Trélat, "Optimal control and applications to aerospace: Some results and challenges," *Journal of Optimization Theory & Applications*, vol. 154, no. 3, pp. 713–758, 2012.

[6] T. Lyons, M. Caruana, and T. Lévy, *Differential Equations Driven by Rough Paths: Lectures from the 34th Summer School on Probability Theory held in Saint-Flour*. Springer Berlin Heidelberg, 2007.

[7] P. K. Friz and M. Hairer, *A course on rough paths: with an introduction to regularity structures*. Springer International Publishing, 2020.

[8] A. Allan, "Lecture notes on rough path theory," ETH Zürich, 2021.

[9] C. Phelps, J. O. Royset, and Q. Gong, "Optimal control of uncertain systems using sample average approximations," *SIAM Journal on Control and Optimization*, vol. 54, no. 1, pp. 1–29, 2016.

[10] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, 3rd ed. Society for Industrial and Applied Mathematics, 2021.

[11] T. Lew, R. Bonalli, and M. Pavone, "Sample average approximation for stochastic programming with equality constraints," *SIAM Journal on Optimization*, vol. 34, no. 4, pp. 3506–3533, 2024.

[12] Z. Pereira, M. A. Wang, I. Exarchos, and E. A. Theodorou, "Learning deep stochastic optimal control policies using forward-backward SDEs," in *Robotics: Science and Systems*, 2019.

[13] R. Carmona and M. Laurire, "Deep learning for mean field games and mean field control with applications to finance," 2021, available at https://arxiv.org/abs/2107.04568.

[14] W. E, J. Han, and A. Jentzen, "Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations," *Communications in Mathematics and Statistics*, vol. 5, no. 4, pp. 349–380, 2017.

[15] R. Bonalli, "Optimal control of aerospace systems with control-state constraints and delays," Ph.D. dissertation, Sorbonne Université & ONERA - The French Aerospace Lab, 2018.