Interplay between Optimal Control and Reinforcement Learning for Agile Locomotion and Dexterous Manipulation in Robotics

F. Schramm[1,2,3] N. Perrin-Gilbert[3] J. Carpentier[1,2]

[1]Inria Paris, France [2]Département d'informatique de l'ENS, PSL Research University, France [3]Sorbonne Université, France

Context and scientific objectives

The capability of modern robots to achieve dexterous manipulation and agile locomotion remains limited.

Goal: Learn new skills **efficiently** and develop **robust and versatile** controller for robotics.
Intuition: Humans solve complex tasks **without thinking** about the movement of each of their muscles separately.
↪ exploits **motor synergies** to control several degrees of freedom simultaneously like central-nervous system
↪ serial and parallel **composition** of skills in a **hierarchic** fashion

Optimal control (OC): impressive results (acrobatic motions by Boston Dynamics [1]), but online re-planning, adaptability and robustness to uncertainties or external perturbations remain very challenging.

Reinforcement learning (RL): flexible and robust against uncertainties, discover complex and rich solutions in the face of contact interactions [2, 3], but methods remain data-intensive.

Should we learn or optimize? Leverage the advantages of both worlds:

- founded on the same mathematical principles (Bellman and Hamilton-Jacobi-Bellman equations) [4]
- connect RL policies at low frequency for high-level decisions to optimal controllers operating at higher frequency for both precise and local adjustment in face of uncertainties
- learn from limited data on a **physical robot** and apply policy efficiently to complex robotic tasks

This thesis is an interdisciplinary project with three main scientific axes: **control**, **perception**, and **experimentation** on simulated and real physical robots.

Optimization in robotics

.. is used for optimal design, simulation, inverse kinematics and dynamics, estimation, trajectory optimization.

Specifically : solve quadratic programs (QPs), under **robustness**, **accuracy** and **speed** requirements

$$\min_{x \in \mathbb{R}^n} \left\{ q(x) := \frac{1}{2} x^T H x + g^T x \right\} \quad \begin{cases} Ax = b \\ Cx - u \leq 0 \end{cases} \tag{1}$$

$H \in \mathcal{S}_+(\mathbb{R}^d)$ (i.e., symmetric positive semi-definite), $A \in \mathbb{R}^{n_e \times d}$, $C \in \mathbb{R}^{n_i \times d}$

Contribution

- Open-source software **Proxsuite**
- **ProxQP** algorithm [5] with state-of-the-art performances and working guarantees
- Solver with state-of-the-art numerical back-end for linear algebra

QP Applications

- Kalman filtering
- Legged locomotion
- Friction-less contact modelling
- Constrained forward-dynamics

First-order trajectory optimization

Goal: **fast** trajectory optimization for systems with **non-smooth dynamics**.

- Build **differentiable simulator** based on pinocchio [6]
- Use randomized smoothing [7] with automatic noise scheduling
- Compare against state-of-the-art RL algorithms using xpag

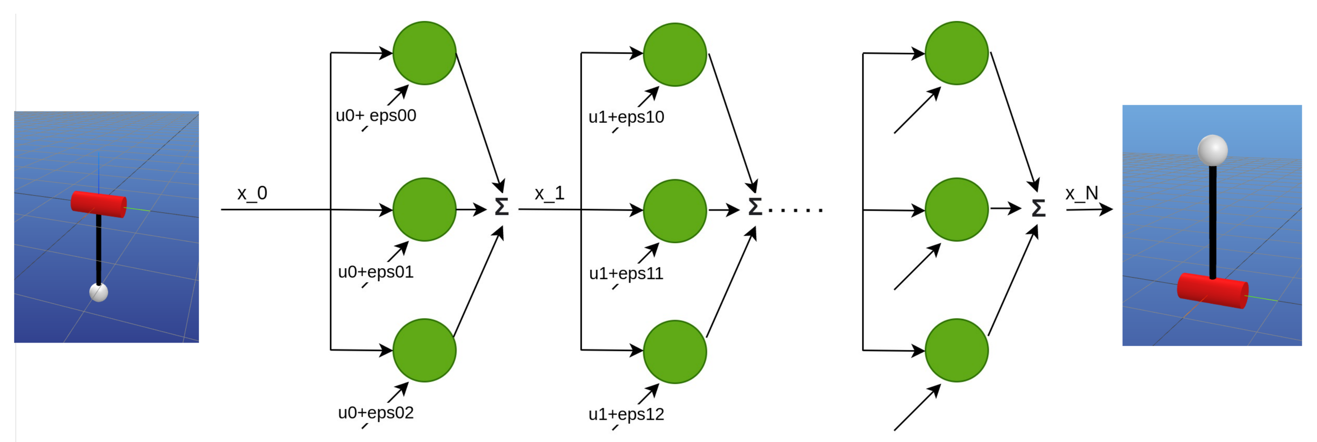
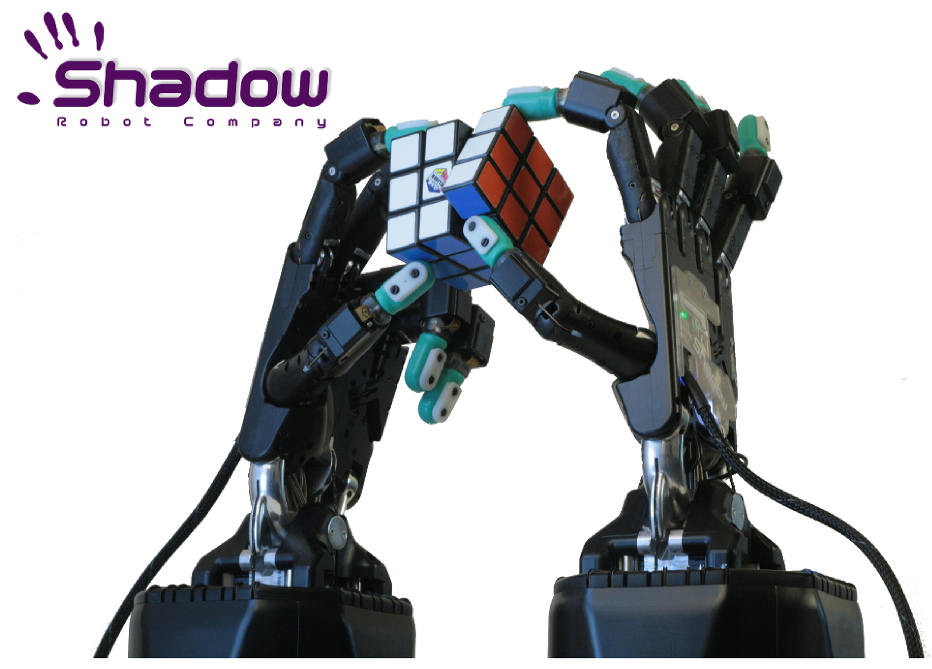


Figure: Randomized smoothing applied to cart-pole system.

Advanced robotic platforms

- Experimental validation in simulation → sim-to-real gap
- Conduct experiments on state-of-the-art robotic platforms for both locomotion and manipulation
- Lay a new computational framework for robot control on **real hardware**

AI4IDF: Purchase of two Shadow Dexterous Humanoid Robot Hands → fine-grained manipulation tasks.



TIRREX (Technological Infrastructure for Robotics Research of Excellence): Acquisition of the humanoid robot Digit
→ navigate complex environments in dynamic fashion.

Open-source software

Development of open-source software enabling the learning and deployment of complex control policies on real robotic platforms. Software packages are **BSD-licensed** and easily installable on all major OS.

Pinocchio — Efficient and versatile rigid body dynamics algorithms

Pinocchio is an open-source and highly efficient framework for **simulation**, **planning** and **control** in robotics, bio-mechanics, civil engineering:

- efficient forward and inverse dynamics
- analytical derivatives of main algorithms
- automatic differentiation (CasADi, CppAD, . . .) and code generation

ProxSuite — The Advanced Proximal Optimization Toolbox

- fast: C++ implementation with custom Cholesky solver
- scalable: various backends for dense, sparse and matrix-free optimization
- easy-to-use: clear API, Python and Julia bindings

ProxSuite is a collection of numerically robust, precise and efficient solvers rooted in revisited primal-dual proximal algorithms. The first targeted application is Robotics, but ProxSuite can be used in other contexts without limits.

xpag

- fast: RL algorithms written in JAX using XLA to speed up differentiation calculations
- robotics-oriented: customized for **goal-conditioned RL**, particularly useful for robotics

xpag is a modular and efficient reinforcement learning (RL) library.

References

1. Kuindersma, S. *Recent Progress on Atlas, the World's Most Dynamic Humanoid Robot.* https://www.youtube.com/watch?v=EGABAx52GKI9. 2020. (2022).
2. Li, Z. *et al. Reinforcement learning for robust parameterized locomotion control of bipedal robots.* in *2021 IEEE International Conference on Robotics and Automation* (2021).
3. Hwangbo, J. *et al.* Learning agile and dynamic motor skills for legged robots. *Science Robotics* (2019).
4. Bertsekas, D. *Reinforcement learning and optimal control.* (Athena Scientific, 2019).
5. Bambade, A., El-Kazdadi, S., Taylor, A. & Carpentier, J. *PROX-QP: Yet another Quadratic Programming Solver for Robotics and beyond.* in *RSS 2022 - Robotics: Science and Systems* (New York, United States, June 2022). https://hal.inria.fr/hal-03683733.
6. Carpentier, J. & Mansard, N. *Analytical Derivatives of Rigid Body Dynamics Algorithms.* in *Robotics: Science and Systems (RSS 2018)* (Pittsburgh, United States, June 2018). https://hal.laas.fr/hal-01790971.
7. Lidec, Q. L., Montaut, L., Schmid, C., Laptev, I. & Carpentier, J. *Leveraging Randomized Smoothing for Optimal Control of Nonsmooth Dynamical Systems.* 2022.