



**Offer #2025-09039**

## **Post-Doctoral Research Visit F/M Theoretical Foundations of Online Convex Reinforcement Learning**

**Contract type :** Fixed-term contract

**Level of qualifications required :** PhD or equivalent

**Fonction :** Post-Doctoral Research Visit

### **About the research centre or Inria department**

The Inria Grenoble research center groups together almost 600 people in 27 research teams and 8 research support departments.

Staff is present on three campuses in Grenoble, in close collaboration with other research and higher education institutions (University Grenoble Alpes, CNRS, CEA, INRAE, ...), but also with key economic players in the area.

Inria Grenoble is active in the fields of high-performance computing, verification and embedded systems, modeling of the environment at multiple levels, and data science and artificial intelligence. The center is a top-level scientific institute with an extensive network of international collaborations in Europe and the rest of the world.

### **Context**

This proposal is supported by the Inria Thoth (<https://team.inria.fr/thoth/>) project-team and may involve collaborations with the Inria Ghost(<https://team.inria.fr/ghost/>) project-team. It will be supervised by Pierre Gaillard.

The position will be based in the Inria Center at the University Grenoble-Alpes.

### **Assignment**

#### **Assignments**

The project will focus on theoretical aspects of **convex reinforcement learning (CURL)**. In recent years, deep reinforcement learning (RL) has seen remarkable success in fields such as language modeling, computer vision, and robotics.

However, RL relies on assumptions of linearity in the objective function, which are not always satisfied.

The CURL problem generalizes RL to a **convex objective**. More precisely, it consists in minimizing a convex function  $f$  over the state-action distributions  $\pi$  induced by an agent's policy  $\pi$  by solving:  $\min_{\pi} f(\pi)$

Beyond RL, CURL generalizes several frameworks in machine learning, including:

- **Pure exploration** [1],
- **Imitation learning** [2],
- Certain instances of **mean-field control** [3],
- **Mean-field games** [4],
- **Risk-averse reinforcement learning** [5].

The non-linearity of CURL breaks the linear structure inherent in standard RL, rendering the classical Bellman equations invalid. The theoretical performance analysis of algorithms in this general framework remains largely unexplored [6-8], and existing solutions rely on strong assumptions and require finite state and action spaces, leading to poor scalability as these spaces grow.

In this postdoctoral project, we aim to lift these restrictive assumptions and extend this line of work to **parametrized state and action spaces**. The main challenge will be to develop an efficient solution that adapts to the **effective dimension** of these spaces. We also anticipate that new research directions may emerge during the visit.

## **Skills**

A Phd degree in mathematics or theoretical computer science, with specialisation optimization, machine learning, statistical learning or game theory, as witnessed by publications in relevant venues including NeurIPS, COLT, ICML, ALT, AISTATS, FOCS, STOC, SODA, EC, JMLR, GEB.

## **References**

- [1] E. Hazan, S. Kakade, K. Singh et A. Van Soest. “Provably Efficient Maximum Entropy Exploration”. In : International Conference on Machine Learning. T. 97. Sept. 2019, p. 2681-2691.
- [2] J. W. Lavington, S. Vaswani et M. Schmidt. “Improved Policy Optimization for Online Imitation Learning”. In :

- Proceedings of The 1st Conference on Lifelong Learning Agents. Sous la dir. de S. Chandar, R. Pascanu et D. Precup. T. 199. Proceedings of Machine Learning Research. PMLR, 22–24 Aug 2022, p. 1146-1173.
- [3] A. Bensoussan, P. Yam et J. Frehse. Mean Field Games and Mean Field Type Control Theory. English. SpringerBriefs in Mathematics. Springer, 2013.
- [4] P. Lavigne et L. Pfeiffer. Generalized conditional gradient and learning in potential mean field games. 2023.
- [5] J. Garcia, Fern et o Fernandez. “A Comprehensive Survey on Safe Reinforcement Learning”. In : Journal of Machine Learning Research 16.42 (2015), p. 1437-1480.
- [6] B. M. Moreno, M. Bregere, P. Gaillard et N. Oudjane. “Efficient model-based concave utility reinforcement learning through greedy mirror descent”. In : International Conference on Artificial Intelligence and Statistics. PMLR. 2024, p. 2206-2214.
- [7] B. M. Moreno, M. Bregere, P. Gaillard et N. Oudjane. “MetaCURL : Non-stationary Concave Utility Reinforcement Learning”. In : NeurIPS’24 : Advances in Neural Information Processing Systems. 2024.
- [8] B. M. Moreno, K. Eldowa, P. Gaillard, M. Bregere et N. Oudjane. “Online Episodic Convex Reinforcement Learning”. In : arXiv preprint arXiv :2505.07303 (2025).

## Main activities

The research mission includes the production of both theoretical and practical contributions, to be enhanced by:

- publications and presentations in machine learning or optimization conferences or journals,
- creation of Python packages

## Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (90 days / year) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities

- Access to vocational training
- Complementary health insurance under conditions

## Remuneration

2788€ gross salary / month

## General Information

- **Theme/Domain** : Optimization, machine learning and statistical methods  
Statistics (Big data) (BAP E)
- **Town/city** : Montbonnot
- **Inria Center** : [Centre Inria de l'Université Grenoble Alpes](#)
- **Starting date** : 2025-10-01
- **Duration of contract** : 1 year, 6 months
- **Deadline to apply** : 2025-07-20

## Contacts

- **Inria Team** : [THOTH](#)
- **Recruiter** :  
Gaillard Pierre / [pierre.gaillard@inria.fr](mailto:pierre.gaillard@inria.fr)

## About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

## Instruction to apply

Applications must be submitted online on the Inria website.

Processing of applications sent by other channels is not guaranteed.

### Defence Security :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit,

following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

**Recruitment Policy :**

As part of its diversity policy, all Inria positions are accessible to people with disabilities.