



**Offre n°2025-09039**

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

**Type de contrat :** Fixed-term contract

**Niveau de diplôme exigé :** PhD or equivalent

**Fonction :** Post-Doctoral Research Visit

### **A propos du centre ou de la direction fonctionnelle**

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.

### **Contexte et atouts du poste**

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.

### **Mission confiée**

#### **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).

## Principales activités

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

## Avantages

- 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

## Rémunération

2788€ gross salary / month

## Informations générales

- **Thème/Domaine** : Optimization, machine learning and statistical methods Statistics (Big data) (BAP E)
- **Ville** : Montbonnot
- **Centre Inria** : [Centre Inria de l'Université Grenoble Alpes](#)
- **Date de prise de fonction souhaitée** : 2025-10-01
- **Durée de contrat** : 1 year, 6 months
- **Date limite pour postuler** : 2025-07-20

## Contacts

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

## A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

**Attention:** Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

## Consignes pour postuler

Applications must be submitted online on the Inria website.

Processing of applications sent by other channels is not guaranteed.

### Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle

que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

**Politique de recrutement :**

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.