

ODALRIC-AMBRYM MAILLARD

Curriculum Vitae

2025



October 12, 1984, Nouméa (France)

ORCID 0000-0001-7935-7026

[🌐 Personal website](#) [🔗 Gitlab](#) [📖 Google scholar](#)

Formulaire 1 — PARCOURS PROFESSIONNEL

Form 1 — PROFESSIONAL HISTORY

Information

Dans cette première fiche, je détaille mon parcours professionnel, d'un point de vue administratif.

1) Parcours Professionnel / *Professional history*

Situation professionnelle actuelle / *Current professional status*

Statut et fonction / *Position and Status*: CRCN Inria, échelon 7

Etablissement (ville - pays) / *Institution (city - country)*: Centre Inria de l'Université de Lille (Villeneuve d'Ascq - France)

Date d'entrée en fonction / *Start*: 2016

Expériences professionnelles antérieures / *Previous professional experiences*

Date début <i>Start</i>	Date fin <i>End</i>	Etablissement <i>Institution</i>	Fonction et statut ² <i>Position and status</i>
2019	Inria Lille - Nord Europe	Chargé de Recherche Habilité (CRCN HDR)
2016	2019	Inria Lille - Nord Europe	Chargé de Recherche (CRCN)
2015	2016	Inria Saclay - Île de France	Chargé de Recherche (CR2) Inria
2014	2015	EECS, Technion, Israel	Senior Researcher
2013	2014	EECS, Technion, Israel	Postdoctoral fellow
2011	2013	Information-Technology, Montanuniversität Leoben, Austria	Postdoctoral fellow
2008	2011	Université Lille 1 and Université Paul-Sabatier	Doctorant (Bourse ENS)

Nombre d'années d'exercice des métiers de la recherche après la thèse / *Number of years of professional research experience after the PhD*: 13 (PHD in october 2011)

2) Interruptions de carrière / *Career breaks*

Pas d'interruption officielle mais l'année COVID a eu de fortes répercussions personnelles

3) Encadrement d'étudiants et de jeunes chercheurs / *Supervision of students and early-stage researchers*

J'ai eu le plaisir d'encadrer plusieurs brillants esprits. En accord avec les instructions de cette section, j'indique ci-après les taux d'encadrements (supervision scientifique effective), lesquels diffèrent légèrement des taux de direction de thèse (supervision administrative), voir également [ici](#). J'indique, le cas échéant, un lien vers la page professionnelle ou le manuscrit de thèse de chacun.

Postdoctorants:

- [Tanguy Lefort](#) (Postdoc, 2024-ongoing, PEPR PI@ntAgroEco), Topic: *Bandit recommender systems for crowdsourcing*.
- [Tuan Dam](#) (Postdoc, 2022-2024, Chaire IA), Topic: *Planning in stochastic environments*. Now Lecturer/Assistant Professor at Hanoi University.
- [Timothée Mathieu](#) (Postdoc, 2021-2023, Inria A.Ex SR4SG), Topic: *Robust decision making to corruption*. Now permanent Inria researcher in Inria Scool team (hence he is now a colleague).
- [Mohammad Sadeqh Talebi](#) (Postdoc, 2018-2020, ANR BADASS), Topic: *Average regret minimization in Markov Decision Processes*. Now Assistant Professor at Copenhagen University.

Doctorants¹:

- Adrien Prevost (50%, 2024-ongoing, PEPR FOUNDRY), Topic: *Foundations of non-parametric, structured, groupsequential bandits*.
- Waris Radji (100%, 2024-ongoing, PEPR FOUNDRY), Topic: *Bandit strategies to revisit approximate and deep reinforcement learning*.
- Anthony Kobanda (50%, 2024-ongoing, CIFRE UBISOFT), Topic: *Continual reinforcement learning for real-world contexts*.

¹ Seuls les étudiants encadrés après mon habilitation (2019) sont indiqués.

- Sumit Vashishtha (100%, 2022-ongoing, Chaire IA fellow), Topic: *Efficient sub-sampling processes for complex bandit problems*.
- Frederick Fabre (33%, 2021-ongoing, U. Reunion & CIRAD), Topic: *Confident prediction of weeds in agroecology context*.
- [Fabien Pesquerel](#) (100%, 2020-2023, ENS fellow, now Postdoc at Inria Rennes) Topic: *Information per unit of interaction in stochastic sequential decision making*. Impact: First ever instance-dependent exact optimal regret efficient strategies for MDPs.
- [Patrick Saux](#) (50%, 2020-2023, ENS fellow) Topic: *Mathematics of statistical sequential decision making : concentration, risk-awareness and modelling in stochastic bandits, with applications to bariatric surgery*. Impact: **1 Lancet publication**, **Interdisciplinary PhD in Medicine**
- [Dorian Baudry](#) (50%, 2019-2022, Inria fellow, now Postdoc in Oxford) Topic: *Non-parametric algorithms for multi-armed bandits*. Impact: State-of-the-art non-parametric optimal bandit strategies, showcased in agronomy context.
- [Romain Gautron](#) (33%, 2019-2022, Cirad fellow, now Postdoc in CIAT) Topic: *Apprentissage par renforcement pour l'aide à la conduite des cultures des petits agriculteurs des pays du Sud: vers la maîtrise des risques*. Impact: **Interdisciplinary PhD in Agronomy**. Sequential version of DSSAT model for RL research. Initiated collaboration with CIRAD & INRAE.
- [Reda Ouhamma](#) (50%, 2019-2023, Ecole Polytechnique fellow, now Postdoc at EPFL) Topic: *Toward realistic reinforcement learning*. Impact: Removing theoretic-convenient but practically-limiting hypothesis in sequential decision making.
- [Hassan Saber](#) (100%, 2019-2022, Inria fellow, now at AFLEC Dubai) Topic: *Structure adaptation in bandit theory*. Impact: Unique proof technique for information-based structured bandits.
- [Edouard Laurent](#) (50%, 2017-2020, CIFRE Renault, now at Google Deepmind) Impact: *Safe and Efficient Reinforcement Learning for Behavioural Planning in Autonomous Driving*. Significance: **3 PhD awards**
Interdisciplinary PhD in Control

En amont de mon habilitation, j'ai également co-supervisé²: S. Akhtar (2017, ANR), R. Alami (2016-2020, CIFRE Orange Labs), R. Allesiaro (2015-2017, CIFRE Orange Labs), N. Galichet (2016, Université Orsay).

J'ai de plus encadré 18 étudiants de Master depuis 2015. Enfin, des doctorants étrangers me contactent régulièrement pour effectuer un séjour de recherche et travailler avec moi durant leur doctorat. De même, de jeunes chercheurs me contactent pour organiser des séjours. Vous trouverez ci-dessous une liste sélectionnée d'étudiants invités :

- Chuan-Zheng Lee, Stanford University, Stanford fellow, May-August 2019.
- Arun Verma, IIT Bombay, Raman-Charpak fellow, June-Nov 2019.
- Ramakrishnan K, IISc Bangalore, Raman-Charpak fellow, May-July 2022.
- Shubhada Agrawal, TIFR Bombay, funded by Inria School, March-May 2022.
- J. Komiyama (2019, Research Associate, University of Tokyo),
- A. Mehrabian (2019, Postdoctoral Researcher, McGill University).

Depuis l'arrivée de la ZRR, l'accueil d'étudiants étrangers pour de telles visites est devenu presque impossible, en raison des délais de traitement du FSD incompatibles avec le dépôt de demandes de bourses spécifiques.

4) Encadrement de développements technologiques (logiciel, matériel, robotique) / *Supervision of technological development (software, hardware, robotics)*

- Hernán Carvajal (Ingénieur, 2021-2023, Inria A.Ex SR4SG): [WeG@rden](#) (voir aussi la [documentation](#)), plateforme de collecte de données et recommandation de pratiques agroécologiques
- T. Carta (Stagiaire, 07-09 2021, Inria): [FarmGym](#) (v1). Simulateur gamifié d'agroécosystèmes, à destination de la recherche en apprentissage séquentiel et par renforcement.
- B. Driss (Ingénieur, 09/2023-09/2024, Inria): FarmGym [AgroRL](#) challenge. Compétition d'algorithmes d'apprentissage par renforcement sur le simulateur FarmGym.

5) Responsabilités collectives / *Responsibilities*

Information

Je liste mes participations détaillées aux comités (soutenance, recrutement, relecture, etc) sur [ma page personnelle](#).

RELECTURE:

Je suis régulièrement sollicité pour participer à des comités de programme de conférences ou pour effectuer des évaluations d'articles pour des revues internationales. Bien que je conserve la liste complète des articles évalués de manière privée, j'ai rendu publique la liste des comités (comités de relecture ou jurys) auxquels j'ai participé, accessible via [cette page](#). Celle-ci inclut un lien vers le site du comité lorsque celui-ci existe. Depuis 2019, date de mon HDR, j'ai évalué environ 60 articles pour

²Concernant les étudiants en thèse, le mot "encadré" ou "dirigé" étant réservés à l'après habilitation, j'utilise ici le mot 'supervisé'.

des revues ou conférences internationales. Entre 2011, début de ma thèse, et 2019, j'ai contribué à environ 55 évaluations dans le cadre de comités de conférences ou de revues scientifiques. Je fais également partie des éditeurs du Journal of Machine Learning Research, et rapporteur pour l'ANR (Agence Nationale de la Recherche) et son équivalent Hongrois, le Hungarian Scientific Research Fund (OTKA).

SOUTENANCES:

- December, 02, 2024, Palaiseau (France): Jury PhD (examineur) of Emilia Siviero
- November, 19, 2024, Montpellier (France): Jury HDR (rapporteur) of Alice Cleynen
- September, 25, 2024, Toulouse (France): Jury PhD (examineur) of David Delande
- September, 20, 2024, Montpellier (France): Jury PhD (examineur) of Tanguy Lefort
- July, 01, 2024, Paris (France): Jury PhD (rapporteur) of David Brellmann
- September, 2023, Lyon (France): Jury PhD (rapporteur) of Antoine Barrier.
- June, 24, 2021, Orsay (France): Jury PhD (rapporteur) of Thibaut Cuvelier
- December, 2020, Orsay (France): Jury PhD (rapporteur) of Margaux Brégère
- October, 09, 2020, Paris (France): Jury PhD (examineur) of Robin Vogel
- February 20, 2018, Stockholm (Sweden): Jury PhD (opponent) of Stefan Magureanu (supervised by A. Proutiere).
- July 08, 2016, Villeneuve d'ascq (France): Jury PhD (examineur) of Hadrien Glaude.

RECRUTEMENTS:

- Spring, 2023, Lille (France): Jury MCF hiring committee of Université Lille.
- May, 2021, online (France): Jury CR hiring committee of the Mathnum department at Inrae.
- Spring, 2020, Lille (France): Jury CR hiring committee of Inria.
- Spring, 2019, Lille (France): Jury CR hiring committee of Inria.
- Spring, 2018, Lille (France): Jury CR hiring committee of Inria.
- 2021, 2022, 2023 Lille (France): Member of the local Commission Emploi Recherche (CER) of Inria Lille.

AUTRE:

- March 2021, Lille (France): Jury Project for Hackathon Biodimètre.

Concernant d'autres activités complémentaires, vous pouvez vous référer à cette [liste d'événements](#).

Je suis également animateur, en binôme avec Mikaela Keller, de l'[Axe Intelligence Artificielle](#) du laboratoire CRIStaL.

ORGANISATION D'ATELIERS SCIENTIFIQUES

- 2024, July, Toulouse (France) - Session organizer at RL4SN Informs conference "Challenges and progresses in Statistical Reinforcement Learning"
- 2020-2022: Tous les ateliers prévus ont été annulés en raison du Covid.
- 2019 - Organizing board of [Reinforcement Learning Summer School](#).
- 2018 - PI of "ANR BADASS Lecture Series", hosting P. Grünwald, at Inria Lille – Nord europe.
- 2017 - PI of Workshop "Sequential Structured Statistical Learning" at Institut des Hautes Études Scientifiques (IHES).
- 2014 - PI of NIPS Workshop "From Bad models to Good policies"

6) Management (si pertinent) / *Management (if relevant)*

EQUIPES DE RECHERCHE:

2022-2024 co-PI of Inria-Japan Associate team "[Real-Life Bandits \(RELIANT\)](#)" with J. Honda, 6 chercheurs permanents, 31 publications, 10k€/an couvrants les frais de missions. Voir également l'[article de vulgarisation](#) à ce sujet.

PROJETS DE RECHERCHE

2022 - 2025 co-PI of Inria-Inrae project "Three-risk proof Sequential Decision Making (3RDSM), 4k€."

2021 co-PI of RLRL "Real-Life Reinforcement Learning" project with B. De Saporta, 20k€.

2021 - 2024 co-PI of Chaire IA AppRenf on reinforcement learning with P. Preux, 500k€.

2019 - 2023 PI of INRIA Action Exploratoire "[Sequential Recommendation for Sustainable Gardening \(SR4SG\)](#)", 165k€: Gathering AI and AgroEcology research to recommend good gardening practice. [Article de vulgarisation](#)

2022 - 2024 Member (10%) of ANR "Bip-UP", with Lille Hospital, PI F. Pattou & P. Preux.

2023 - 2025 Member (5%) of ANR Epi-RL project, PI. F. Schwarzenruber

2023 - 2026 Member (5%) of PEPR PI@ntAgroEco project, PI. Joly&Bonnet

2019 - 2022 Member (10%) of I-SITE project "Bandits for Health (B4H)", with Lille Hospital, PI F. Pattou & P. Preux.

2019 - 2020 Member (5%) of "Learning Framework for Radio resource Allocation in LoRaWaAN", PI K. Khawam.

2016 - 2020 PI of ANR JCJC project "Bandits against non-Stationarity and Structure (BADASS)", 180k with E. Kaufmann and R. Combes. <https://project.inria.fr/badass/> Postdoctorant: M.S. Talebi, and PhD: S. W. Akhtar

2016 - co-PI IFCAM project "Contextual multi-armed bandits with hidden structure", with Aditya Gopalan (Bangalore), 7k.

2015 - PI Programme Invité Digiteo, 4k: Invitation of A. Gopalan for 3 weeks.

2015 - co-PI PEPS JCJC - INS2I project "PROMO" with R. Bardenet (CNRS, Lille), 7k.

2015 - co-PI INRIA-Carnot "Outil générique de prévision/contrôle de flux hydrauliques" with O. Teytaud and company prolog (1y engineer).

7) Mobilité (si pertinent) / *Mobility (if relevant)*

Geographically, I did my PhD in between Toulouse and Lille. I did a postdoc in Leoben, Austria, and a second one at Technion, Israel. I joined Inria Saclay then moved to Inria Lille. Thematically, I am operating a thematic shift towards ML for agroecology that started about 5 years ago, and continues today.

8) Enseignement (si pertinent) / *Teaching (if relevant)*

J'enseigne régulièrement un cours de **Statistical Reinforcement Learning** dans plusieurs programmes de Master :

- **École Polytechnique** : Master 2 "Artificial Intelligence & Advanced Visual Computing (Ai-ViC)" (contact: Marie-Paule Cani, Erwan Scornet)
 - Période : 2019–2023 (5 ans)
 - Volume horaire : 48 heures par an
 - Public : ~20 étudiants par an
- **École Centrale Supélec** : Master 2 Mention IA (contact: Céline Hudelot)
 - Période : 2021–2022
 - Volume horaire : 24 heures
 - Public : ~70 étudiants par an
- **École Centrale de Lille** : Master 2 "Données, Apprentissage, Décision (DAD)" (contact: Pierre Chainais)
 - Période : 2017–2018
 - Volume horaire : 18 heures
 - Public : ~40 étudiants

Je précise "Statistical" dans l'intitulé pour avertir les étudiants sur l'aspect théorique du cours. J'y aborde notamment des outils de concentration de la mesure, analyses de regret, apprentissage séquentiel et optimisation, régression statistique, équations de Bellman, programmation dynamique permettant d'établir les garanties théoriques de performance des algorithmes d'apprentissage. Une partie du cours s'appuie sur mon [manuscrit d'habilitation](#), servant de notes techniques et permettant d'approfondir certains sujets. Une autre partie plus informelle/visuelle est présente sur mon compte [github](#), on l'on y trouve des exemples de travaux pratiques. Une version plus élaborée néanmoins réservée aux chargés de TDs (corrigés, projets passés, etc) est présente sur mon compte [gitlab inria](#).

J'ai également enseigné un cours sur les Bandits manchots lors de la **Reinforcement Learning Summer School (RLSS)** organisée par mon équipe Inria, ([lien](#)):

- Période : 1^{er}–12 juillet 2019
- Volume horaire : 5 heures
- Public : ~300 étudiants

A cela s'ajoutent des [tutoriaux](#) dispensés sur des sujets plus applicatifs de l'apprentissage par renforcement. Enfin, durant ma thèse, j'ai exercé en tant que Moniteur, dispensant des travaux dirigés variés aux étudiants de Licence informatique.

9) Diffusion de l'information scientifique (si pertinent) / *Dissemination of scientific knowledge (if relevant)*

Mes activités de diffusion des savoirs s'étendent sur plusieurs axes, incluant des interventions en formation continue, ainsi que des présentations destinées à des publics non spécialistes, auxquelles s'ajoutent des initiatives plus récentes pour des cours "ouverts".

FORMATION CONTINUE:

J'interviens régulièrement dans des programmes de formation continue :

- **École Polytechnique** : Executive Master Data-Science (contact: David Filliat)
 - Période : 2020–2024 (5 ans)
 - Volume horaire : 3 heures, deux fois par an
 - Public : ~50 participants par session

RESSOURCES PEDAGOGIQUES:

Récemment, j'ai initié des collaborations avec des structures visant à rendre la science accessible au plus grand nombre :

- **Inria Learning Lab** : Classe ouverte sur l'Apprentissage par Renforcement (projet en cours, A. Lagarrigue).
- **Inria Academy** : Cours d'Apprentissage par Renforcement pour des entreprises (projet en cours, A-C. Lamballe).

INTERVENTIONS AUPRES DE PUBLICS NON SPECIALISTE:

Je participe régulièrement à des événements et conférences destinés à des audiences non académiques:

- **Présentations hors communauté de recherche** : Consultez la [liste de mes interventions](#) pour plus de détails.
- **Fête de la Science** : Participation en octobre 2020 à une vidéo grand public organisée par Inria.
- **Journée "Petits débrouillards", 2015, LRI**: Organisation de présentation à destination des lycéens.

10) Visibilité (si pertinent) / *Visibility (if relevant)*

En 2022, j'ai eu l'heureuse surprise que des chercheurs indépendants³ **nomment un algorithme en mon honneur** ([Lien](#)).

En 2012, j'ai reçu le **1^{er} prix national de thèse** décerné par l'AFIA ([Lien](#)).

Avec mes co-auteurs, nous avons eu deux articles acceptés en **présentations orales à NeurIPS**⁴ ([Lien](#), [Lien](#))

Certaines de mes invitations et interventions sont listées [ici](#). En voici une sélection :

- June 28-30, 2023, Nancy (France): INFORMS Applied Probability Society Conference.
- June 04, 2019, Amsterdam (Netherlands): Centrum voor Wiskunde en Informatica (CWI) (groupe de P. Grünwald).
- May 13-14, 2019, Besançon (France): "Workshop on Empirical Processes and Applications to Statistics".
- March 07, 2019, Paris (France): "3rd nonstationary day", Institut Henri Poincaré (IHP).
- November 03, 2018, Sapporo (Japan): 20th Ibis conference (**plénière d'ouverture**) .
- June 03, 2016, Paris (France): GdR Isis.
- January 12, 2016, Lancaster (UK): "the Bandits Workshop"
- January 08, 2016, Orsay (France): journées "Statistics/Learning" at Paris-Saclay at IHES.
- November 09, 2015, Toulouse (France): Intervention au CIMI.
- January, 30, 2015 Paris (France): Journées "YSP" à l'Institut Henri Poincaré (IHP).

I was also invited to [organize](#) a session at RL4SN INFORMS conference, June 17-21, 2024, Toulouse (France).

11) Éléments divers / *Other relevant information*

En 2023, j'ai soumis une proposition ERC Consolidator Grant (CoG). Bien que j'ai réussi à passer la phase écrite (B1), le projet n'a finalement pas été financé, principalement jugé « trop ambitieux ». A vrai dire, je devais initialement bénéficier d'un accompagnement par l'Inria, mais le chercheur censé me superviser n'a offert aucune aide après la première phase et peu durant celle-ci, me laissant gérer seul la préparation de l'Oral.

En 2024, j'ai par la suite coordonné une soumission d'un ambitieux projet PEPR à l'interface Agroécologie-Numérique, réunissant environ 30 chercheurs de plusieurs instituts dont INRIA, CIRAD, INRAE, CNA (Centre National d'Agroécologie), ARMEFLHOR, et stimulant de nombreuses discussions. Ce projet est actuellement en révision.

Mon étudiant Édouard Leurent, a reçu **trois prix** suites à ses travaux de thèse (Abertis Prize for Best Doctoral Thesis, Best PhD Award awarded by CNRS (GdR MACS) & Club EEA, Second prize at "AI and the 3rd industrial revolution" Challenge)

Mon étudiante Mahsa Asadi, a obtenu un prix de **meilleur article étudiant** à ACML.

Entre 2016 et 2020, j'ai bénéficié de la Prime d'encadrement doctoral et de recherche (PEDR), mais je n'y postule plus désormais.

En 2015, j'ai été sélectionné parmi 200 jeunes chercheurs internationaux pour participer au Heidelberg Laureate Forum et rencontrer les lauréats Abel, Fields et Turing. La même année, j'ai participé à une délégation de l'Association Technion France (ATF) en Israël.

³de moi, c'est à dire avec qui je n'ai aucun lien, collaboration, étudiant, ni projet commun.

⁴Taux de sélection < 3%, < 1% respectivement.

Formulaire 2 — DESCRIPTION SYNTHÉTIQUE DE L'ACTIVITÉ ANTÉRIEURE

Form 2 — SUMMARY OF YOUR PAST ACTIVITY

Information

Dans cette partie, je présente un bref aperçu des lignes directrices formant mes travaux de recherche.

VISION GÉNÉRALE

Apport à la recherche Ma recherche se situe à l'intersection entre la **théorie et les applications de l'apprentissage automatique**, avec un intérêt particulier pour le développement de méthodologies pour la prise de décision séquentielle, la théorie des bandits multi-bras (MAB), et les statistiques mathématiques appliquées à l'apprentissage par renforcement (AR). Mes contributions abordent des défis fondamentaux tels que la **minimisation du regret, la quantification de l'incertitude et l'adaptabilité**, tout en visant un impact concret dans les sciences expérimentales et la planification sous incertitude. Le fil conducteur de mon travail est d'**offrir aux systèmes décisionnels la capacité d'apprendre efficacement et de manière robuste dans des environnements dynamiques, incertains et structurés**.

Vision Une conviction profonde dans ma recherche est que **fournir des garanties théoriques sur les capacités d'apprentissage** des algorithmes est essentiel à leur applicabilité, notamment lorsque ces garanties sont à la fois fines (les plus précises) et significatives (fondées sur des hypothèses pratiques et des quantités interprétables). Cette approche rigoureuse inspire non seulement le développement de nouveaux outils statistiques mais aussi la conception d'algorithmes innovants. Bien que cette rigueur semble parfois moins présente dans des domaines comme l'apprentissage profond, souvent dominés par des avancées à fort contenu d'ingénierie, je suis convaincu que l'établissement de bases théoriques solides est de plus en plus crucial. Par exemple, l'émergence des réglementations sur l'IA souligne le besoin croissant de certifications robustes, garantissant que les comportements algorithmiques respectent des normes éthiques et pratiques adaptées aux applications réelles.

Direction et synergies J'ai toujours poursuivi des orientations qui non seulement cherchent à faire progresser l'état de l'art, mais qui peuvent également créer des synergies entre les disciplines. Cela inclut l'identification des problèmes émergents dans les bandits manchots, l'analyse par concentration de la mesure et l'apprentissage robuste, ainsi que l'orientation de la recherche de mon groupe pour relever ces défis efficacement. En travaillant avec diverses équipes, j'ai cosigné plus de 60 publications, et j'accepte toujours des collaborations sur des défis ciblés. J'ai notamment développé un réseau interdisciplinaire de collaborateurs dans le domaine du contrôle, de la médecine et de l'agroécologie, en poussant à des initiatives collectives en open-source et à une recherche éthique. Je cherche de plus en plus à faire le lien entre les avancées théoriques et les applications marquantes, et ces collaborations m'orientent également vers la création de plateformes applicatives (telles que gym-dssat, farm-gym, ou wegarden).

THÉMATIQUES DE RECHERCHE

Mes activités de recherche s'articulent d'une part selon les thèmes et questions de recherches considérés, et d'autre part selon les projets de recherche que j'ai portés ou auxquels j'ai contribué. Les questions de recherches portant sur système dynamique incertain étant très riches, je procède à présent en alimentant une "pile" de questions de recherches que je formalise au fil de l'eau et que je dépile avec mes étudiants et collaborateurs, en fonction de leurs affinités et de l'orientation que prend leur réflexion. Ceci permet une plus grande flexibilité, et un meilleur engagement des étudiants, tout en résolvant un problème "lorsque c'est le moment". Ces projets se séparent eux même en deux grandes familles:

Fondements de l'apprentissage par renforcement statistique, à portée plus méthodologique. La ligne de force s'est d'abord articulée autour de mon ANR JCJC "[Bandits against non-Stationarity and Structure](#) (BADASS)", où l'accent a été mis sur la prise de décision séquentielle non stationnaire (tels que les points de changement [39, 40, 32, 81]) et l'exploitation de diverses notions de structure dans les bandits et les processus décisionnels de Markovs (PDMs) [9, 44, 46, 37] pour améliorer l'efficacité des stratégies d'apprentissage en terme de regret. Les résultats de ce projet ont par la suite influencé la Chaire IA "AppRenf" (avec Philippe Preux) ainsi que la définition du projet d'équipe Inria School (Sequential Continual and Online Learning), avec mes collègues Emilie Kaufmann et Philippe Preux. Par la suite, ces travaux ont eux même ouvert la porte à l'équipe associée Inria-Japon [Real-Life Bandits](#) (RELIANT) entre Inria School et le "groupe de Kyoto", que je co-dirige avec Junya Honda (University Kyoto), se concentrant sur les défis théoriques qui limitent l'applicabilité des bandits dans le monde réel, telles que les données corrompues, les contraintes d'équité ou les défis de calcul. C'est dans ce cadre que nous avons obtenu, par exemple, [22, 18] l'optimalité en minimisation du regret dans les PDMs. Ces travaux de nature fondamentale se poursuivent aujourd'hui avec différents collaborateurs.

Développement de l'AR pour assister les sciences expérimentales, à portée plus applicative. Plus particulièrement, ces travaux s'organisent autour de trois types de projets à portée applicative. D'une part la thèse brillante⁵ CIFRE d'Edouard Leurent sur le contrôle et la planification robuste pour les problèmes de navigation de véhicules autonomes, d'autre part une initiative personnelle marquée en direction de l'agroécologie, avec l'Action Exploratoire AEx "Sequential Recommendation for Sustainable Gardening (SR4SG)", ouvrant une collaboration interdisciplinaire très fructueuses avec le CIRAD et l'INRAE et plusieurs projets à l'interface Agronomie-numérique. Je cite en particulier la thèse interdisciplinaire remarquable de Romain Gautron, avec le CIRAD sur l'applicabilité de l'apprentissage aux agroécosystèmes. Ici, le vif intérêt que j'ai développé pour l'agroécologie, et qui représente une voie de recherche originale dans le domaine de l'apprentissage séquentiel, vient d'une conviction profonde que les défis qui s'y posent sont à la fois de la plus haute importance sociétale et fournissent une source fraîche de belles questions de recherche en statistique mathématique et en apprentissage par renforcement. C'est pourquoi j'ai beaucoup poussé dans ces directions ces dernières années, en développant de plus en plus de connexions et de collaborations avec des chercheurs en agroécologie. J'ai d'ailleurs coordonné en 2024 une soumission d'un ambitieux projet PEPR à l'interface Agroécologie-Numérique, réunissant environs 30 chercheurs de plusieurs instituts dont INRIA, CIRAD, INRAE, CNA (Centre National d'Agroécologie), et stimulant de nombreuses de discussions. Ce projet est actuellement en révision.

Enfin, je participe à l'élaboration de compagnons de décisions en médecine, dans le projet ANR "Bip-UP", porté par F. Pattou (Hopital de Lille) & P. Preux (Inria Scool), suivant le projet "Bandits for Health (B4H)", se concentrant tous deux sur le suivi de patients dans un cadre clinique réel post-opératoire. Ici, le doctorat de Patrick Saux a permis d'appliquer des méthodes de concentration avancée [19] pour élaborer des régions de confiance en prédiction du poids formées sur une grande cohorte de patients. Ces travaux ont finalement été publiés dans la revue médicale *Lancet* par l'équipe médicale. En lien avec ces questions, j'ai également collaboré avec B. De Saporta (U. Montpellier) via le projet « Real-Life Reinforcement Learning » de nature plus fondamentale concernant l'apprentissage de processus de Markov déterministes par morceaux via l'apprentissage par renforcement, un modèle de contrôle fortement motivé par la formalisation d'une question médicale.

ENCADREMENTS ET COLLABORATIONS

Succès Tout au long de ma carrière, j'ai eu le privilège d'encadrer des personnes brillantes, chacune pour des raisons uniques. À chaque fois, je suis émerveillé et conforté par l'importance du lien humain dans la réussite de ces parcours, bien au-delà de la simple maîtrise scientifique. Je reste en contact étroit avec plusieurs d'entre eux, qui évoluent aujourd'hui dans divers pays (Colombie, Émirats arabes unis, Danemark, Finlande, Inde, etc.). Leur enthousiasme spontané pour les mathématiques ou l'agroécologie, qu'ils partagent encore avec moi, me donne le sentiment que le "message est passé". De plus, de nouvelles collaborations voient le jour (Ubisoft, CNA, ESA, etc.), ouvrant des perspectives prometteuses et créant des opportunités inédites pour l'avenir.

Echecs Je dois admettre quelques échecs concernant des doctorants ayant interrompu leur thèse : Dans chaque cas (trois au total), les étudiants n'avaient pas effectué de stage avec moi avant de commencer leur doctorat, en raison de contraintes imposées par l'école doctorale et de contingences liées au financement. J'ai toujours plaidé pour que les stages soient obligatoires avant d'autoriser les étudiants à entamer leur parcours doctoral. Cependant, chaque année, les calendriers et mécanismes d'attribution de bourse contraignent à admettre des doctorants sans pouvoir effectuer de stage préalable, ce qui engendre alors des difficultés pour toutes les parties concernées.

Conclusion Pour conclure, mes travaux de recherche s'inscrivent dans une trajectoire claire visant à concevoir des systèmes d'apprentissage automatique robustes, adaptatifs et théoriquement solides, répondant à des enjeux fondamentaux tout en prenant en compte les complexités du monde réel. Ce focus, associé à une capacité démontrée à encadrer et diriger des collaborations pluridisciplinaires, me positionne en situation de direction de recherche pour définir et conduire des agendas de recherche impactants tant pour la société que pour l'avenir de ces domaines. C'est donc tout naturellement que je poursuis cet agenda en candidatant sur poste de directeur de recherche.

⁵soulignée par l'obtention de 2 prix de thèse, et d'un oral à Neurips.

Formulaire 3 — LISTE COMPLÈTE DES CONTRIBUTIONS

Form 3 — COMPLETE LIST OF CONTRIBUTIONS

1. Publications caractéristiques/*Representative publications*

- [1] Sayak Ray Chowdhury, Patrick Saux, **ODALRIC-AMBRYM MAILLARD**, Aditya Gopalan. **BREGMAN DEVIATIONS OF GENERIC EXPONENTIAL FAMILIES** In *Conference On Learning Theory (COLT)*, Jul 2023, Bangalore, India. ([Link](#))

Ce travail "clôture" en quelque sorte un long parcours de recherche sur la construction de régions de confiance uniformes en temps pour des familles exponentielles génériques, en étendant la méthode de Laplace de manière élégante et générale. Le théorème obtenu, lorsqu'il est appliqué à de nombreuses familles exponentielles classiques, que nous prenons soins de détailler, génère des régions de confiance explicites surpassant les techniques spécifiques existantes. Il permet notamment de traiter des familles exponentielles de dimension supérieure, incluant l'exemple emblématique des lois normales à moyenne et variance inconnues, un cas parfois cité comme particulièrement difficile.

La méthode de preuve combine avec finesse géométrie, dualité et arguments statistiques, et est le fruit de ma contribution personnelle, en combinant les propriétés de dualités des divergences de Bregman avec un choix judicieux de mesure de référence dans l'application de la méthode de Laplace. Les étudiants Sayak et Patrick en particulier, ont pu dériver la spécialisation du gain d'information de Bregman et des bornes de concentration à une riche famille de lois. Travailler sur ce projet a également été l'occasion agréable de renouveler ma collaboration avec Aditya Gopalan (IISc Bangalore); nous nous connaissons depuis notre postdoctorat commun au Technion.

- [2] Fabien Pesquerel and **Odalric-Ambrym Maillard**. **IMED-RL: REGRET OPTIMAL LEARNING OF ERGODIC MARKOV DECISION PROCESSES**. In *NeurIPS 2022 - Thirty-sixth Conference on Neural Information Processing Systems*, Thirty-sixth Conference on Neural Information Processing Systems, New-Orleans, United States, November 2022. ([Link](#))

Cet article introduit le premier algorithme de minimisation du regret pour les processus de décision markoviens (MDP) offrant des performances de minimization de regret prouvées optimales, au sens "instance-dependent". Un point remarquable, est que, en plus des garanties théoriques d'optimalité, la stratégie proposée est également efficace sur le plan numérique, y compris en dehors des hypothèses d'analyse. Cet article s'inscrit dans une série de travaux intensifs dans lequel je me suis investi pleinement, embarquant étudiants et co-auteurs, et poursuivis également par plusieurs groupes visant à intégrer des stratégies issues des bandits dans les MDP. Il fait suite en particulier à mes contributions précédentes [88, 44, 31]. Dans cette série de travaux, j'ai agi comme chercheur principal, apportant des innovations clés et développant des outils d'analyse précis, notamment des bornes uniformes en temps et basées sur la divergence de Kullback-Leibler, tout en affinant les quantités problème-dépendantes, notamment pour l'estimation des noyaux de transition et leur impact sur les estimations des fonctions de valeur. Ces travaux auraient eu une portée toute différente sans l'apport de chercheurs talentueux, que j'ai pris à plaisir à encadrer sur ces sujets: Dr. Sadegh Talebi (ancien postdoc, maintenant professeur assistant à l'Université de Copenhague), dont la clarté rédactionnelle et la "regularisation" apportée par nos discussions a permis de mieux communiquer des résultats complexes; Fabien Pesquerel (ancien doctorant, actuellement postdoctorant à Inria Rennes), qui a adapté des techniques d'amélioration de politique pour IMED-RL, essentielles pour les bornes indépendantes du problème dans les MDP ergodiques. Ce projet représentait un défi important, et Fabien était la personne idéale pour le mener à bien. En ce sens, c'est à lui que l'on doit la preuve finale de cet algorithme; Enfin Hassan Saber (ancien doctorant, actuellement à Abu Dhabi), dont les techniques de preuves sur les bandits structurés m'ont inspiré pour analyser, à la suite de cet article, des MDP génériques avec transitions connues, en mettant l'accent sur des stratégies d'exploration efficaces [18]. Nous étendons désormais cette recherche d'une part aux MDP à états continus pour l'apprentissage approché et profond avec mon étudiant Waris Radji, et aux MDPs génériques en collaboration avec le groupe de Bruno Gaujal (Inria Grenoble).

- [3] Edouard Leurent, Denis Efimov, and **Odalric-Ambrym Maillard**. **ROBUST-ADAPTIVE CONTROL OF LINEAR SYSTEMS: BEYOND QUADRATIC COSTS**. In *34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020. ([Link](#)) **Oral, 1% acceptance rate**

Ce travail résulte d'une collaboration avec une autre équipe Inria spécialisée en contrôle, l'équipe VALSE, ainsi qu'avec le constructeur automobile Renault, dans le cadre d'une thèse CIFRE portant sur la planification comportementale robuste pour la navigation autonome. Nous avons combiné la construction de régions de confiance empiriques issues pour la régression avec la planification robuste en contrôle linéaire ainsi qu'une approche pertinente de propagation d'intervalles. Cette thèse a également permis de relever plusieurs défis en planification, notamment l'adaptation de la recherche d'arbre Monte Carlo aux graphes pour la planification en boucle fermée, et la correction des limitations des stratégies de planification optimistes en boucle ouverte.

Dans cette collaboration, j'ai apporté mon expertise sur la planification et la construction de régions de confiance, tout en assurant un suivi rapproché d'Édouard Leurent. Concernant les ensembles de confiance, une procédure d'estimation des paramètres a été utilisée pour modéliser le comportement incertain des autres véhicules dans la scène. Cela a permis de définir des régions de confiance sur leurs trajectoires, utilisées ensuite pour adapter la planification du véhicule contrôlé. J'ai guidé Édouard dans les raffinements et analyses théoriques de ces questions, que nous avons majoritairement développées ensemble. La mise en œuvre finale et les simulations sont dues à ses excellentes compétences en programmation.

Cette collaboration fructueuse a permis à Édouard Leurent de recevoir trois prix de thèse et une publication orale à NeurIPS. Par ailleurs, cette thèse a démontré l'intérêt combiné des outils de statistiques séquentielles développés au sein de l'équipe Scool et des outils issus de la théorie du contrôle développés par l'équipe VALSE.

2. Politique de publication/*Publication policy*

Politique sur les versions courtes/étendues Dans la liste ci-dessous, certaines publications apparaissent d'abord sous forme de workshop, avant d'être complétées et publiées en conférence ou journal. Les workshops, dans notre domaine, ne comportent pas d'actes, contrairement aux conférences et journaux. Les conférences en apprentissage automatique sont quant à elles sélectives, avec un taux d'acceptation entre 20 et 30%, et représentent l'équivalent de journaux dans d'autres domaines.

ArXiv ou non ? J'ai toujours privilégié les conférences et journaux, seules sources validées par un comité de relecture. Avant 2020, je m'abstenais de publier sur ArXiv, car ce n'est pas une plateforme certifiée. Les rares exceptions étaient sous l'insistance de co-auteurs. Aujourd'hui, bien que la publication d'articles de conférences sur ArXiv soit devenue une pratique courante assurant une visibilité accrue à nombre de travaux, je reste hésitant à suivre cette pratique, ce qui diminue malheureusement leur portée.

3. Publications

Conventions pour l'ordre des auteurs Il est presque impossible de déduire le rôle exact d'un auteur à partir de sa position dans la liste, car les conventions varient selon les contextes : communautés scientifiques, institutions, rôle en tant que superviseur ou non, ou encore selon des règles ad hoc. Souvent, l'ordre alphabétique est utilisé, parfois combiné avec les positions académiques, ce qui donne peu d'indications sur les contributions réelles. Pour clarifier mon rôle, j'utilise des styles de police distincts pour mon nom dans la bibliographie ci-dessous (d'autres co-auteurs peuvent avoir des rôles similaires, mais je n'applique cette convention qu'à moi-même). Les rôles sont définis comme suit :

- **PRINCIPAL CONTRIBUTEUR**: Responsable de l'idée principale, apportant des développements théoriques significatifs, la rédaction et/ou les expériences. Indiqué en majuscules et en gras.
- **Contributeur clé**: Apporte une idée ou un résultat/preuve clé, sans être nécessairement l'auteur principal. Indiqué en gras.
- **Contributeur**: Contributeur standard. Indiqué en police normale.

En complément, ces rôles peuvent être combinés avec celui de Superviseur clé, Indiqué par une police soulignée, témoignant d'un encadrement actif : discussions, contributions aux résultats, etc., de manière approfondie. Cette convention définit six rôles distincts identifiés par des styles de police correspondants.

3.1 Notable Publications/*Notable publications*

- [1] Edouard Leurent, Denis Efimov, and **Odalric-Ambrym Maillard**. **ROBUST-ADAPTIVE CONTROL OF LINEAR SYSTEMS: BEYOND QUADRATIC COSTS**. In *34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020. **Oral, 1% acceptance rate**
- [2] Mahsa Asadi, Mohammad Sadegh Talebi, Hippolyte Bourel, and **ODALRIC-AMBRYM MAILLARD**. **MODEL-BASED REINFORCEMENT LEARNING EXPLOITING STATE-ACTION EQUIVALENCE**. In *Asian Conference on Machine Learning (ACML)*, pages 204–219. PMLR, 2019. **Best student paper award**
- [3] **ODALRIC-AMBRYM MAILLARD**, Timothy A Mann, and Shie Mannor. **HOW HARD IS MY MDP?" THE DISTRIBUTION-NORM TO THE RESCUE**. *Advances in Neural Information Processing Systems (NIPS)*, volume 27, pages 1835–1843, 2014. **Oral, 3% acceptance rate**

3.2 Revues internationales/*International journals*

- [1] Timothée Mathieu, Debabrota Basu, **Odalric-Ambrym Maillard**. **BANDITS CORRUPTED BY NATURE: LOWER BOUNDS ON REGRET AND ROBUST OPTIMISTIC ALGORITHMS**. In *Transactions on Machine Learning Research Journal (TMLR)*, 2024.
- [2] Timothée Mathieu, Riccardo Della Vecchia, Alena Shilova, Matheus Medeiros Centa, Hector Kohler, Odalric-Ambrym Maillard, Philippe Preux. **ADA STOP: SEQUENTIAL TESTING FOR EFFICIENT AND RELIABLE COMPARISONS OF DEEP RL AGENTS**. In *Transactions on Machine Learning Research Journal (TMLR)*, 2024.
- [3] Kinda Khawam, Hassan Fawaz, Samer Lahoud, Odalric-Ambrym Maillard and Steven Martin **A CHANNEL SELECTION GAME FOR MULTI-OPERATOR LoRaWAN DEPLOYMENTS**. In *Computer Networks*, Volume 216, October 2022.

- [4] Mahsa Asadi, Aurélien Bellet, **Odalric-Ambrym Maillard** and Marc Tommasi. **COLLABORATIVE ALGORITHMS FOR ONLINE PERSONALIZED MEAN ESTIMATION**. In *Transactions on Machine Learning Research*, 2022.
- [5] Romain Gautron, **Odalric-Ambrym Maillard**, Philippe Preux, Marc Corbeels, and Régis Sabbadin. **REINFORCEMENT LEARNING FOR CROP MANAGEMENT**. *Computers and Electronics in Agriculture*, 200:107182, July 2022.
- [6] **ODALRIC-AMBRYM MAILLARD**. **LOCAL DVORETZKY-KIEFER-WOLFOWITZ CONFIDENCE BANDS**. *Mathematical Methods of Statistics*, 2022.
- [7] Lilian Besson, Emilie Kaufmann, **Odalric-Ambrym Maillard**, and Julien Seznec. **EFFICIENT CHANGE-POINT DETECTION FOR TACKLING PIECEWISE-STATIONARY BANDITS**. *Journal of Machine Learning Research*, March 2022.
- [8] **ODALRIC-AMBRYM MAILLARD**. **BOUNDARY CROSSING PROBABILITIES FOR GENERAL EXPONENTIAL FAMILIES**. *Mathematical Methods of Statistics*, 27(1):1–31, 2018.
- [9] Audrey Durand, **Odalric-Ambrym Maillard** and Joëlle Pineau. **STREAMING KERNEL REGRESSION WITH PROVABLY ADAPTIVE MEAN, VARIANCE, AND REGULARIZATION**. *The Journal of Machine Learning Research*, 19(1):650–683, 2018.
- [10] Robin Allesiardo, Raphaël Féraud, and **Odalric-Ambrym Maillard**. **THE NON-STATIONARY STOCHASTIC MULTI-ARMED BANDIT PROBLEM**. In *International Journal of Data Science and Analytics*, 3(4):267–283, 2017.
- [11] Rémi Bardenet and **ODALRIC-AMBRYM MAILLARD**. **CONCENTRATION INEQUALITIES FOR SAMPLING WITHOUT REPLACEMENT**. *Bernoulli*, 21(3):1361–1385, 2015.
- [12] Olivier Cappé, Aurélien Garivier, **ODALRIC-AMBRYM MAILLARD**, Rémi Munos and Gilles Stoltz. **KULLBACK–LEIBLER UPPER CONFIDENCE BOUNDS FOR OPTIMAL SEQUENTIAL ALLOCATION**. *The Annals of Statistics*, 41(3):1516–1541, 2013.
- [13] **ODALRIC-AMBRYM MAILLARD** and Rémi Munos. **LINEAR REGRESSION WITH RANDOM PROJECTIONS**. *The Journal of Machine Learning Research*, 13:2735–2772, 2012.

3.3 Conférence internationales avec comité de lecture/Reviewed international conferences

- [14] Hassan Saber, **Odalric-Ambrym Maillard**. **BANDITS WITH MULTIMODAL STRUCTURE**. In *RLC 2024 - Reinforcement Learning Conference*, Aug 2024, Amherst Massachusetts, United States. pp.39.
- [15] Tuan Quang Tuan Dam, **Odalric-Ambrym Maillard**, Emilie Kaufmann. **POWER MEAN ESTIMATION IN STOCHASTIC MONTE-CARLO TREE SEARCH**. In *Uncertainty in Artificial Intelligence*, Jul 2024, Barcelona, Spain.
- [16] Shubhada Agrawal, Timothée Mathieu, Debabrota Basu, **Odalric-Ambrym Maillard**. **CRIMED: LOWER AND UPPER BOUNDS ON REGRET FOR BANDITS WITH UNBOUNDED STOCHASTIC CORRUPTION**. In *International Conference on Algorithmic Learning Theory (ALT)*, Feb 2024, San Diego (CA), United States. pp.74-124.
- [17] Dorian Baudry, Fabien Pesquerel, Rémy Degenne, **Odalric-Ambrym Maillard**. **FAST ASYMPTOTICALLY OPTIMAL ALGORITHMS FOR NON-PARAMETRIC STOCHASTIC BANDITS**. In *NeurIPS 2023 - Thirty-seventh Conference on Neural Information Processing Systems*, Dec 2023, New Orleans (Louisiana), United States.
- [18] Hassan Saber, Fabien Pesquerel, **ODALRIC-AMBRYM MAILLARD**, Mohammad Sadegh Talebi. **LOGARITHMIC REGRET IN COMMUNICATING MDPs: LEVERAGING KNOWN DYNAMICS WITH BANDITS**. In *Asian Conference on Machine Learning (ACML)*, Nov 2023, Istanbul, Turkey.
- [19] Sayak Ray Chowdhury, Patrick Saux, **ODALRIC-AMBRYM MAILLARD**, Aditya Gopalan. **BREGMAN DEVIATIONS OF GENERIC EXPONENTIAL FAMILIES** In *Conference On Learning Theory (COLT)*, Jul 2023, Bangalore, India.
- [20] Patrick Saux and **Odalric-Ambrym Maillard**. **RISK-AWARE LINEAR BANDITS WITH CONVEX LOSS**. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR 206, Valencia, Spain, April 2023.
- [21] Reda Ouhamma, Debabrota Basu and **Odalric-Ambrym Maillard**. **BILINEAR EXPONENTIAL FAMILY OF MDPs: FREQUENTIST REGRET BOUND WITH TRACTABLE EXPLORATION & PLANNING** In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 37, 8:9336-9344, Washington DC, United States, Feb 2023.
- [22] Fabien Pesquerel and **Odalric-Ambrym Maillard**. **IMED-RL: REGRET OPTIMAL LEARNING OF ERGODIC MARKOV DECISION PROCESSES**. In *NeurIPS 2022 - Thirty-sixth Conference on Neural Information Processing Systems*, Thirty-sixth Conference on Neural Information Processing Systems, New-Orleans, United States, November 2022.
- [23] Hassan Saber, Pierre Ménard, and **Odalric-Ambrym Maillard**. **INDEXED MINIMUM EMPIRICAL DIVERGENCE FOR UNIMODAL BANDITS**. In *35th International Conference on Neural Information Processing Systems (NeurIPS)*, Sydney, Australia, December 2021.

- [24] Dorian Baudry, Patrick Saux, and Odalric-Ambrym Maillard. **FROM OPTIMALITY TO ROBUSTNESS: DIRICHLET SAMPLING STRATEGIES IN STOCHASTIC BANDITS**. In *35th International Conference on Neural Information Processing Systems (NeurIPS)*, Sydney, Australia, December 2021.
- [25] Fabien Pesquerel, Hassan Saber, and Odalric-Ambrym Maillard. **STOCHASTIC BANDITS WITH GROUPS OF SIMILAR ARMS**. In *NeurIPS 2021 - Thirty-fifth Conference on Neural Information Processing Systems*, Stochastic bandits with groups of similar arms, Sydney, Australia, December 2021.
- [26] Hassan Saber, Léo Saci, **Odalric-Ambrym Maillard**, and Audrey Durand. **ROUTINE BANDITS: MINIMIZING REGRET ON RECURRING PROBLEMS**. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*, Bilbao, Spain, September 2021.
- [27] Dorian Baudry, Romain Gautron, Émilie Kaufmann and **Odalric-Ambrym Maillard**. **OPTIMAL THOMPSON SAMPLING STRATEGIES FOR SUPPORT-AWARE CVAR BANDITS**. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*. PMLR, 2021.
- [28] Mohammad Sadegh Talebi, Anders Jonsson, and Odalric-Ambrym Maillard. **IMPROVED EXPLORATION IN FACTORED AVERAGE-REWARD MDPs**. In *International Conference on Artificial Intelligence and Statistics (AI&STATS)*, volume 130. PMLR, 2021.
- [29] Sayak Ray Chowdhury, Aditya Gopalan, and **ODALRIC-AMBRYM MAILLARD**. **REINFORCEMENT LEARNING IN PARAMETRIC MDPs WITH EXPONENTIAL FAMILIES**. In *International Conference on Artificial Intelligence and Statistics (AI&STATS)*, pages 1855–1863. PMLR, 2021.
- [30] Yannis Flet-Berliac, Reda Ouhamma, Odalric-Ambrym Maillard, and Philippe Preux. **LEARNING VALUE FUNCTIONS IN DEEP POLICY GRADIENTS USING RESIDUAL VARIANCE**. In *International Conference on Learning Representations (ICLR)*, 2021.
- [31] Hippolyte Bourel, **ODALRIC-AMBRYM MAILLARD**, and Mohammad Sadegh Talebi. **TIGHTENING EXPLORATION IN UPPER CONFIDENCE REINFORCEMENT LEARNING**. In *International Conference on Machine Learning (ICML)*, pages 1056–1066. PMLR, 2020.
- [32] Réda Alami, **Odalric-Ambrym Maillard**, and Raphaël Féraud. **RESTARTED BAYESIAN ONLINE CHANGE-POINT DETECTOR ACHIEVES OPTIMAL DETECTION DELAY**. In *International Conference on Machine Learning (ICML)*, pages 211–221. PMLR, 2020.
- [33] Edouard Leurent, Denis Efimov, and **Odalric-Ambrym Maillard**. **ROBUST-ADAPTIVE CONTROL OF LINEAR SYSTEMS: BEYOND QUADRATIC COSTS**. In *34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
- [34] Edouard Leurent, Denis Efimov, and Odalric-Ambrym Maillard. **ROBUST-ADAPTIVE INTERVAL PREDICTIVE CONTROL FOR LINEAR UNCERTAIN SYSTEMS**. In *59th IEEE Conference on Decision and Control (CDC)*, pages 1429–1434. IEEE, 2020.
- [35] Edouard Leurent and **Odalric-Ambrym Maillard**. **MONTE-CARLO GRAPH SEARCH: THE VALUE OF MERGING SIMILAR STATES**. In *Asian Conference on Machine Learning (ACML)*, pages 577–592. PMLR, 2020.
- [36] Edouard Leurent, Denis Efimov, and **Odalric-Ambrym Maillard**. **ROBUST ESTIMATION, PREDICTION AND CONTROL WITH LINEAR DYNAMICS AND GENERIC COSTS**. In *34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
- [37] Mahsa Asadi, Mohammad Sadegh Talebi, Hippolyte Bourel, and **ODALRIC-AMBRYM MAILLARD**. **MODEL-BASED REINFORCEMENT LEARNING EXPLOITING STATE-ACTION EQUIVALENCE**. In *Asian Conference on Machine Learning (ACML)*, pages 204–219. PMLR, 2019.
- [38] Edouard Leurent and **Odalric-Ambrym Maillard**. **PRACTICAL OPEN-LOOP OPTIMISTIC PLANNING**. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD)*, pages 69–85. Springer, Cham, 2019.
- [39] **ODALRIC-AMBRYM MAILLARD**. **SEQUENTIAL CHANGE-POINT DETECTION: LAPLACE CONCENTRATION OF SCAN STATISTICS AND NON-ASYMPTOTIC DELAY BOUNDS**. In *Algorithmic Learning Theory (ALT)*, pages 610–632. PMLR, 2019.
- [40] Subhojyoti Mukherjee and Odalric-Ambrym Maillard. **IMPROVED CHANGEPOINT DETECTION FOR PIECEWISE IID BANDITS**. In *22nd International Conference on Artificial Intelligence and Statistics (AI&STATS)*, 2019.
- [41] Ronald Ortner, Matteo Pirota, Ronan Fruit, Alessandro Lazaric, and Odalric-Ambrym Maillard. **REGRET BOUNDS FOR LEARNING STATE REPRESENTATIONS IN REINFORCEMENT LEARNING**. In *Advances in Neural Information Processing Systems (NIPS)*, volume 32, pages 12738–12748, 2019.
- [42] Mohammad Sadegh Talebi and **Odalric-Ambrym Maillard**. **LEARNING MULTIPLE MARKOV CHAINS VIA ADAPTIVE ALLOCATION**. In *Advances in Neural Information Processing Systems (NIPS)*, Volume 32, 2019.
- [43] Nicolas Carrara, Edouard Leurent, Romain Laroche, Tanguy Urvoy, **Odalric-Ambrym Maillard**, and Olivier Pietquin. **BUDGETED REINFORCEMENT LEARNING IN CONTINUOUS STATE SPACE**. In *Advances in Neural Information Processing Systems (NIPS)*, Volume 32, 2019.

- [44] Mohammad Sadegh Talebi and **Odalric-Ambrym Maillard**. [VARIANCE-AWARE REGRET BOUNDS FOR UNDISCOUNTED REINFORCEMENT LEARNING IN MDPs](#). In *Algorithmic Learning Theory (ALT)*, pages 770–805. PMLR, 2018.
- [45] Réda Alami, **Odalric-Ambrym Maillard**, and Raphaël Féraud. [MEMORY BANDITS: A BAYESIAN APPROACH FOR THE SWITCHING BANDIT PROBLEM](#). In *31st Conference on Neural Information Processing Systems (NIPS)*, 2017.
- [46] Borja Balle and **ODALRIC-AMBRYM MAILLARD**. [SPECTRAL LEARNING FROM A SINGLE TRAJECTORY UNDER FINITE-STATE POLICIES](#). In *International Conference on Machine Learning (ICML)*, pages 361–370. PMLR, 2017.
- [47] **ODALRIC-AMBRYM MAILLARD**. [BOUNDARY CROSSING FOR GENERAL EXPONENTIAL FAMILIES](#). In *International Conference on Algorithmic Learning Theory (ALT)*, pages 151–184. PMLR, 2017.
- [48] Jaouad Mourtada and **Odalric-Ambrym Maillard**. [EFFICIENT TRACKING OF A GROWING NUMBER OF EXPERTS](#). In *International Conference on Algorithmic Learning Theory (ALT)*, pages 517–539. PMLR, 2017.
- [49] Akram Erraqabi, Michal Valko, Alexandra Carpentier, and **Odalric-Ambrym Maillard**. [PLIABLE REJECTION SAMPLING](#). In *International Conference on Machine Learning (ICML)*, pages 2121–2129. PMLR, 2016.
- [50] Akram Baransi, **Odalric-Ambrym Maillard**, and Shie Mannor. [SUB-SAMPLING FOR MULTI-ARMED BANDITS](#). In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD)*, pages 115–131. Springer, Berlin, Heidelberg, 2014.
- [51] **ODALRIC-AMBRYM MAILLARD**, Timothy A Mann, and Shie Mannor. [HOW HARD IS MY MDP?" THE DISTRIBUTION-NORM TO THE RESCUE"](#). In *Advances in Neural Information Processing Systems (NIPS)*, pages 1835–1843, 2014.
- [52] **ODALRIC-AMBRYM MAILLARD** and Shie Mannor. [LATENT BANDITS](#). In *International Conference on Machine Learning (ICML)*, pages 136–144. PMLR, 2014.
- [53] Ronald Ortner, **ODALRIC-AMBRYM MAILLARD**, and Daniil Ryabko. [SELECTING NEAR-OPTIMAL APPROXIMATE STATE REPRESENTATIONS IN REINFORCEMENT LEARNING](#). In *International Conference on Algorithmic Learning Theory (ALT)*, pages 140–154. Springer, Cham, 2014.
- [54] **ODALRIC-AMBRYM MAILLARD**. [ROBUST RISK-AVERSE STOCHASTIC MULTI-ARMED BANDITS](#). In *International Conference on Algorithmic Learning Theory (ALT)*, pages 218–233. Springer, Berlin, Heidelberg, 2013.
- [55] **Odalric-Ambrym Maillard**, Phuong Nguyen, Ronald Ortner, and Daniil Ryabko. [OPTIMAL REGRET BOUNDS FOR SELECTING THE STATE REPRESENTATION IN REINFORCEMENT LEARNING](#). In *International Conference on Machine Learning (ICML)*, pages 543–551. PMLR, 2013.
- [56] Phuong Nguyen, **Odalric-Ambrym Maillard**, Daniil Ryabko, and Ronald Ortner. [COMPETING WITH AN INFINITE SET OF MODELS IN REINFORCEMENT LEARNING](#). In *Artificial Intelligence and Statistics (AI&STATS)*, pages 463–471. PMLR, 2013.
- [57] **ODALRIC-AMBRYM MAILLARD**. [HIERARCHICAL OPTIMISTIC REGION SELECTION DRIVEN BY CURIOSITY](#). In *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)*, volume 1, pages 1448–1456, 2012.
- [58] **ODALRIC-AMBRYM MAILLARD** and Alexandra Carpentier. [ONLINE ALLOCATION AND HOMOGENEOUS PARTITIONING FOR PIECEWISE CONSTANT MEAN-APPROXIMATION](#). In *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)*, volume 2, pages 1961–1969, 2012.
- [59] Alexandra Carpentier, **Odalric-Ambrym Maillard**, and Rémi Munos. [SPARSE RECOVERY WITH BROWNIAN SENSING](#). In *Advances in Neural Information Processing System (NIPS)*, pages 1782–1790, 2011.
- [60] **ODALRIC-AMBRYM MAILLARD**, Rémi Munos, et al. [ADAPTIVE BANDITS: TOWARDS THE BEST HISTORY-DEPENDENT STRATEGY](#). In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AI&STATS)*; pages 570–578. JMLR Workshop and Conference Proceedings, 2011.
- [61] **ODALRIC-AMBRYM MAILLARD**, Rémi Munos, and Gilles Stoltz. [A FINITE-TIME ANALYSIS OF MULTI-ARMED BANDITS PROBLEMS WITH KULLBACK-LEIBLER DIVERGENCES](#). In *Proceedings of the 24th annual conference on learning theory (COLT)*, pages 497–514. JMLR Workshop and Conference Proceedings, 2011.
- [62] Mohammad Ghavamzadeh, Alessandro Lazaric, **Odalric-Ambrym Maillard**, and Rémi Munos. [LSTD WITH RANDOM PROJECTIONS](#). In *Advances in Neural Information Processing System (NIPS)*, pages 721–729, 2010.
- [63] **ODALRIC-AMBRYM MAILLARD** and Rémi Munos. [ONLINE LEARNING IN ADVERSARIAL LIPSCHITZ ENVIRONMENTS](#). In *Joint european conference on machine learning and knowledge discovery in databases (ECML-PKDD)*, pages 305–320. Springer, Berlin, Heidelberg, 2010.
- [64] **ODALRIC-AMBRYM MAILLARD** and Rémi Munos. [SCRAMBLED OBJECTS FOR LEAST-SQUARES REGRESSION](#). In *Advances in Neural Information Processing System (NIPS)*, pages 1549–1557, 2010.

- [65] **Odalric-Ambrym Maillard**, Rémi Munos, Alessandro Lazaric, and Mohammad Ghavamzadeh. [FINITE-SAMPLE ANALYSIS OF BELLMAN RESIDUAL MINIMIZATION](#). In *Proceedings of 2nd Asian Conference on Machine Learning (ACML)*, pages 299–314. JMLR Workshop and Conference Proceedings, 2010.
- [66] **ODALRIC-AMBRYM MAILLARD** and Rémi Munos. [COMPRESSED LEAST-SQUARES REGRESSION](#). In *Advances in Neural Information Processing System (NIPS)*, 2009.
- [67] **ODALRIC-AMBRYM MAILLARD** and Nicolas Vayatis. [COMPLEXITY VERSUS AGREEMENT FOR MANY VIEWS](#). In *International Conference on Algorithmic Learning Theory (ALT)*, pages 232–246. Springer, Berlin, Heidelberg, 2009.

3.4 Ateliers internationaux avec comité de lecture/*Reviewed international workshops*

- [68] **ODALRIC-AMBRYM MAILLARD**, Timothée Mathieu and Debabrota Basu. [FARM-GYM: A MODULAR REINFORCEMENT LEARNING PLATFORM FOR STOCHASTIC AGRONOMIC GAMES](#). In *AIAFS 2023 - 2nd AAAI Workshop on AI for Agriculture and Food Systems*, Washington DC, United States, February 2023.
- [69] Romain Gautron, Emilio J. Padrón, Philippe Preux, Julien Bigot, Odalric-Ambrym Maillard, Gerrit Hoogenboom and Julien Teigny. [LEARNING CROP MANAGEMENT BY REINFORCEMENT: GYM-DSSAT](#). In *AIAFS 2023 - 2nd AAAI Workshop on AI for Agriculture and Food Systems*, Washington DC, United States, February 2023.
- [70] Patrick Saux and Odalric-Ambrym Maillard. [RISK-AWARE LINEAR BANDITS WITH CONVEX LOSS](#). In *European Workshop on Reinforcement Learning*, Milan, Italy, September 2022.
- [71] Reda Ouhamma, Debabrota Basu, and Odalric-Ambrym Maillard. [BILINEAR EXPONENTIAL FAMILY OF MDPs: FREQUENTIST REGRET BOUND WITH TRACTABLE EXPLORATION & PLANNING](#). In *EWRL 2022 – European Workshop on Reinforcement Learning*, Milan, Italy, September 2022.
- [72] Nicolas Carrara, Edouard Leurent, Romain Laroche, Tanguy Urvoy, Odalric-Ambrym Maillard, and Olivier Pietquin. [SCALING UP BUDGETED REINFORCEMENT LEARNING](#). *CoRR*, abs/1903.01004, 2019.
- [73] **ODALRIC-AMBRYM MAILLARD**, Rémi Coulom, and Philippe Preux. [PARALLELIZATION OF THE TD \(\$\lambda\$ \) LEARNING ALGORITHM](#). In *The Seventh European Workshop on Reinforcement Learning*, 2005.

3.5 Livres et chapitres de livre/*Books and book chapters*

- [74] **ODALRIC-AMBRYM MAILLARD**. [Mathematics of statistical sequential decision making](#). Habilitation thesis, Université de Lille, Sciences et Technologies, 2019.
- [75] **ODALRIC-AMBRYM MAILLARD**. [Apprentissage Séquentiel: Bandits, Statistique et Renforcement](#). PhD thesis, Université des Sciences et Technologie de Lille-Lille I, 2011.

3.6 Autres publications internationales (posters, articles courts)/*Other international publications (posters, short papers)*

- [76] Réda Alami, Odalric-Ambrym Maillard, and Raphaël Féraud. [MEMORY BANDITS: TOWARDS THE SWITCHING BANDIT PROBLEM BEST RESOLUTION](#). In *MLSS 2018-Machine Learning Summer School*, 2018.

3.7 Revues nationales/*National journals*

3.8 Conférence nationales avec comité de lecture/*Reviewed national conferences*

- [77] Frédéric Fabre Ferber, Dominique Gay, Jean-Christophe Soulié, Jean Diatta, Odalric-Ambrym Maillard. [ÉVALUATION DE CRITÈRES DE SÉLECTION DE NOYAUX POUR LA RÉGRESSION RIDGE À NOYAU DANS UN CONTEXTE DE PETITS JEUX DE DONNÉES](#) In *24ème conférence francophone sur l'Extraction et la Gestion des Connaissances (EGC) 2024*, Jan 2024, Dijon, France.

3.9 Ateliers nationaux avec comité de lecture/*Reviewed national workshops*

3.10 Rapports de recherche et articles soumis/*Research reports and publications under review*

- [78] Romain Gautron, Emilio J. Padrón, Philippe Preux, Julien Bigot, Odalric-Ambrym Maillard, and David Emukpere. [GYM-DSSAT: A CROP MODEL TURNED INTO A REINFORCEMENT LEARNING ENVIRONMENT](#). Research Report RR-9460, Inria Lille, July 2022.
- [79] Debabrota Basu, Odalric-Ambrym Maillard, and Timothée Mathieu. [BANDITS CORRUPTED BY NATURE: LOWER BOUNDS ON REGRET AND ROBUST OPTIMISTIC ALGORITHM](#). working paper or preprint, March 2022.

- [80] Dorian Baudry, Emilie Kaufmann, and **Odalric-Ambrym Maillard**. **SUB-SAMPLING FOR EFFICIENT NON-PARAMETRIC BANDIT EXPLO-
RATION**. *arXiv preprint arXiv:2010.14323*, 2020.
- [81] Lilian Besson, Emilie Kaufmann, **Odalric-Ambrym Maillard**, and Julien Seznec. **EFFICIENT CHANGE-POINT DETECTION FOR
TACKLING PIECEWISE-STATIONARY BANDITS**. 2020.
- [82] Yannis Flet-Berliac, Reda Ouhamma, Odalric-Ambrym Maillard, and Philippe Preux. **IS STANDARD DEVIATION THE NEW STANDARD?
REVISITING THE CRITIC IN DEEP POLICY GRADIENTS**. *arXiv preprint arXiv:2010.04440*, 2020.
- [83] Hassan Saber, Pierre Ménard, and **Odalric-Ambrym Maillard**. **FORCED-EXPLORATION FREE STRATEGIES FOR UNIMODAL BANDITS**. *arXiv
preprint arXiv:2006.16569*, 2020.
- [84] Hassan Saber, Pierre Ménard, and **Odalric-Ambrym Maillard**. **OPTIMAL STRATEGIES FOR GRAPH-STRUCTURED BANDITS**. *arXiv
preprint arXiv:2007.03224*, 2020.
- [85] **ODALRIC-AMBRYM MAILLARD**, Timothy Mann, Ronald Ortner, and Shie Mannor. **ACTIVE ROLL-OUTS IN MDP WITH IRREVERSIBLE
DYNAMICS**. 2019.
- [86] Edouard Leurent, Yann Blanco, Denis Efimov, and **Odalric-Ambrym Maillard**. **APPROXIMATE ROBUST CONTROL OF UNCERTAIN
DYNAMICAL SYSTEMS**. *arXiv preprint arXiv:1903.00220*, 2019.
- [87] Subhojyoti Mukherjee and **Odalric-Ambrym Maillard**. **DISTRIBUTION-DEPENDENT AND TIME-UNIFORM BOUNDS FOR PIECEWISE IID BANDITS**.
arXiv preprint arXiv:1905.13159, 2019.
- [88] Mohammad Sadegh Talebi and **Odalric-Ambrym Maillard**. **KL-UCRL REVISITED: VARIANCE-AWARE REGRET BOUND**. 2018.
- [89] **ODALRIC-AMBRYM MAILLARD**. **BASIC CONCENTRATION PROPERTIES OF REAL-VALUED DISTRIBUTIONS**. 2017. [https://hal.
archives-ouvertes.fr/cel-01632228](https://hal.archives-ouvertes.fr/cel-01632228)
- [90] **ODALRIC-AMBRYM MAILLARD**. **SELF-NORMALIZATION TECHNIQUES FOR STREAMING CONFIDENT REGRESSION**. 2016.
- [91] Robin Allesiardo, Raphaël Féraud, and **Odalric-Ambrym Maillard**. **RANDOM SHUFFLING AND RESETS FOR THE NON-STATIONARY
STOCHASTIC BANDIT PROBLEM**. *arXiv preprint arXiv:1609.02139*, 2016.
- [92] Aditya Gopalan, **Odalric-Ambrym Maillard**, and Mohammadi Zaki. **LOW-RANK BANDITS WITH LATENT MIXTURES**. *arXiv preprint
arXiv:1609.01508*, 2016.
- [93] Rémi Bardenet and **ODALRIC-AMBRYM MAILLARD**. **A NOTE ON REPLACING UNIFORM SUBSAMPLING BY RANDOM PROJECTIONS IN MCMC
FOR LINEAR REGRESSION OF TALL DATASETS**. *arXiv preprint*, 2015.
- [94] **ODALRIC-AMBRYM MAILLARD**, Rémi Munos, and Daniil Ryabko. **SELECTING THE STATE-REPRESENTATION IN REINFORCEMENT
LEARNING**. *arXiv preprint arXiv:1302.2552*, 2013.
- [95] **ODALRIC-AMBRYM MAILLARD** and Rémi Munos. **BROWNIAN MOTIONS AND SCRAMBLED WAVELETS FOR LEAST-SQUARES REGRESSION**.
2010.