
Contact Information

Work Address Dipartimento di Informatica “Giovanni degli Antoni”, Università degli studi di Milano, via Celoria 18, 20133 Milan, Italy

Email matteo.papini@unimi.it

Website <https://t3p.github.io/>

Highlights

Matteo Papini is a **Tenure Track Assistant Professor (RTT)** at the **University of Milan**, Italy, with the LAILA Lab (Laboratory for Artificial Intelligence and Learning Algorithms) in the Department of Computer Science. His main line of research is **Reinforcement Learning** theory and applications. His research interests include online learning, multi-armed bandits, learning theory, and multi-agent systems.

In 2021 he earned a PhD in Information Technology (cum laude) from **Politecnico di Milano**, Milan, Italy, with a dissertation on safe policy optimization. In 2020 he worked as a student research intern at **Facebook AI Research** (now Meta). From 2021 to 2023 he was a postdoctoral researcher in the Artificial Intelligence and Machine Learning Research Group of **Universitat Pompeu Fabra** (UPF), Barcelona, Spain. From 2023 to 2025 he was a fixed-term assistant professor (RTDA) in Politecnico di Milano and a member of the RL3 group (Real Life Reinforcement Learning Research Lab) of AIRLab (Artificial Intelligence and Robotics Laboratory).

He has authored more than 20 peer-reviewed research papers, including at top artificial intelligence and machine learning conferences (**NeurIPS**, **ICML**, **COLT**, **AAAI**, **IJCAI**) and journals (**JMLR**, **Machine Learning**).

He is a member of **ELLIS** (European Laboratory for Learning and Intelligent Systems) since 2022. He has served as area chair for NeurIPS (since 2024) and ICML 2025, and is an action editor of **TMLR** (Transactions on Machine Learning Research) since 2024. In 2023 he was an organizer of **RLSS 2023** (Reinforcement Learning Summer School) in Barcelona. He was local chair of **ALT 2025** (International Conference on Algorithmic Learning Theory) in Milan.

Faculty and Research Positions

- October 2025–present **Tenure-track Assistant Professor (RTT)**, *Università degli Studi di Milano*, Milan, Italy
- 2023–2025 **Fixed-term Assistant Professor (RTDA)**, *Politecnico di Milano*, Milan, Italy
- 2021–2023 **Postdoctoral researcher**, *Universitat Pompeu Fabra*, Barcelona
- Mar-Apr 2021 **Research Assistant** (Attività di Supporto alla Ricerca), *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano*, Milan, Italy
Research Topic: Reinforcement Learning Techniques for Developing Artificial Test Drivers on a F1 Simulator, Principal Investigator: Marcello Restelli
- Sep-Dec 2020 **Research Intern at Facebook AI Research**, working (remotely) with the Paris FAIR team, under the supervision of Matteo Pirota. Research topic: representation learning for contextual bandits. The project led to the publication of a paper at ICML 2021 [C3]
- 2017–2021 **Research Assistant** (Assegno di Ricerca), *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano*, Milan, Italy
Research topic: *Study and Development of Reinforcement Learning Techniques in Industrial Settings*, Principal Investigator: Marcello Restelli.

Education

- 2017–2021 **Ph.D., Ingegneria dell'Informazione / Information Technology**, *Politecnico di Milano*, Milan, Italy
Advisor: Prof. Marcello Restelli
Dissertation: *Safe Policy Optimization* (<http://hdl.handle.net/10589/170196>)
[Awarded cum laude](#)
Date of degree: March 11, 2021
- 2015–2017 **M.Sc., Computer Science and Engineering - Ingegneria Informatica**, *Politecnico di Milano*, Milan, Italy
Master Thesis: *Adaptive Batch Size for Safe Policy Gradient Methods*, supervised by Prof. Marcello Restelli
Final mark: 110/110 cum laude
Date of degree: July 27, 2017
- 2012–2015 **B.Sc., Ingegneria Informatica (Computer Engineering)**, *Politecnico di Milano*, Milan, Italy
Final mark: 110/110 cum laude
Date of degree: July 24, 2015
- 2007–2012 **High School Diploma (Liceo Scientifico PNI)**, *Ist. Istr. Sup. P.L. Nervi*, Morbegno (SO), Italy, Final mark: 100/100
Top school in Italy for several years according to Eduscopio (<https://eduscopio.it/>)

Author Profile and Metrics (as of October 1, 2025)

Google Scholar **h-index 15**, <https://scholar.google.it/citations?user=A2WxZlsAAAAJ>

Scopus **h-index 12**, <https://www.scopus.com/authid/detail.uri?authorId=57202057824>

ORCID <https://orcid.org/0000-0002-3807-3171>
dblp <https://dblp.uni-trier.de/pid/209/4897.html>

Summary of Publications

Journal/Conference	# Papers	Scimago 24/CORE 23	Homepage
JMLR	1	Q1	http://www.jmlr.org/
Machine Learning	3	Q1	https://www.springer.com/journal/10994
NeurIPS	7	A*	https://nips.cc/
ICML	6	A*	https://icml.cc/
AAAI	2	A*	https://aaai.org/conference/aaai/
IJCAI	2	A*	https://www.ijcai.org/
COLT	2	A*	https://learningtheory.org/
AISTATS	2	A	https://aistats.org/
IJCNN	1	B	https://ijcnn.org/
ALT	2	B	http://algorithmiclearningtheory.org/

Selected Publications

- [1] **M. Papini**, D. Binaghi, G. Canonaco, M. Pirotta, M. Restelli: *Stochastic variance-reduced policy gradient*. **ICML** (2018) — **CORE A***
- [2] A.M. Metelli, **M. Papini**, F. Faccio, M. Restelli: *Policy optimization via importance sampling*. **NeurIPS (oral, top 3%)** (2018) — **CORE A***
- [3] L. Bisi, L. Sabbioni, E. Vittori, **M. Papini**, M. Restelli: *Risk-averse trust region optimization for reward-volatility reduction*. **IJCAI** (2020) — **CORE A***
- [4] A.M. Metelli, **M. Papini**, N. Montali, M. Restelli: *Importance Sampling Techniques for Policy Optimization*. **Journal of Machine Learning Research (JMLR)** 21.141 (2020) — **Scimago Q1**
- [5] **M. Papini**, M. Pirotta, M. Restelli: *Adaptive batch size for safe policy gradients*. **NeurIPS** (2017) — **CORE A***
- [6] P. D'Oro, A. M. Metelli, A. Tirinzoni, **M. Papini**, M. Restelli: *Gradient-aware model-based policy search*. **AAAI** (2020) — **CORE A***
- [7] **M. Papini**, A.M. Metelli, L. Lupo, M. Restelli: *Optimistic policy optimization via multiple importance sampling*. **ICML** (2019) — **CORE A***
- [8] **M. Papini**, M. Pirotta, M. Restelli: *Smoothing policies and safe policy gradients*. **Machine Learning** (2022) — **Scimago Q1**
- [9] **M. Papini**, A. Tirinzoni, M. Restelli, A. Lazaric, M. Pirotta: *Leveraging good representations in linear contextual bandits*. **ICML** (2021) — **CORE A***
- [10] **M. Papini**, A. Tirinzoni, A. Pacchiano, M. Restelli, A. Lazaric, M. Pirotta: *Reinforcement learning in linear mdps: Constant regret and representation selection*. **NeurIPS** (2021) — **CORE A***
- [11] G. Neu, J. Olkhovskaya, **M. Papini**, L. Schwartz: *Lifting the Information Ratio: An Information-Theoretic Analysis of Thompson Sampling for Contextual Bandits*. **NeurIPS** (2022) — **CORE A*** (authors in alphabetical order)

- [12] G. Paczolay., **M. Papini**, A. M. Metelli, I. Harmati, M. Restelli: *Sample complexity of variance-reduced policy gradient: weaker assumptions and lower bounds*. **Machine Learning** (2024) — [Scimago Q1](#)

Awards and other Recognitions

- Nov 16, 2022 **Recipient of Ayudas Juan de la Cierva-Formación 2021**, *research grant for postdoctoral researchers instructed by the Spanish research agency (Agencia Estatal de Investigación), funded by NextGenerationEU*
Role: researcher (Investigador). [Ranked third](#) in national call for the area of Information and Communication Technologies.
<https://www.aei.gob.es/convocatorias/buscador-convocatorias/ayudas-contratos-juan-cierva-formacion-2021>
- Feb 2022 **Member of ELLIS** , <https://ellis.eu/>
- 2021 **Neurips outstanding reviewer award**, [top 8%](#)
- 2020 **ICML top reviewer** certificate of appreciation, [top 33%](#)
- 2019 **ICML travel award**, \$1300 USD
- 2018 **Oral presentation at NeurIPS**, [top 3%](#)
- 2018 **Neurips travel award**, \$1000 USD
- 2018 **ICML travel award**, \$1500 USD
- 2017 **Neurips travel award**, \$1200 USD

Participation to Research Projects

- 2024-2025 **Principal Investigator**, *Sviluppo di strategie automatiche di trading per il mercato dell'energia basate su Reinforcement Learning (development of automatic trading strategies for the energy market based on Reinforcement Learning)*, research contract between Politecnico di Milano and A2A S.p.A.
- 2024-2025 **Principal Investigator**, *CC Auto Tune: ottimizzazione automatica dei parametri del Performance Controller per Compressori Centrifughi (automatic optimization of performance controller parameters for centrifugal compressors)*, research contract between Politecnico di Milano and MADE S.C.A.R.L. (Competence Center for Industry 4.0)
Industrial research project with Baker Hughes Nuovo Pignone.
- 2023-present **Research Scientist**, *Artificial Intelligence Foundations for Sequential Decision Making*
Extended Partnership - Future Artificial Intelligence Research (FAIR). National Recovery and Resilience Plan, Mission 4 "Education and research" - Component 2 "From research to business" - Investment 1.3, funded by the European Union - NextGenerationEU
Principal Investigator: Nicola Gatti
- 2021-2022 **Research Scientist**, *Provably Efficient Algorithms for Large-Scale Reinforcement Learning*
ERC (Grant agreement No. 950180)
Principal Investigator: Gergely Neu

- 2021 **Research Scientist**, *Reinforcement learning techniques for developing artificial test drivers on a F1 simulator*, Ferrari
Industrial project.
Principal Investigator: Marcello Restelli
- 2017-2019 **Research Scientist**, *TOTAL EFFICIENCY 4.0, (POR FESR)*
In collaboration with Pirelli Tyre S.p.A.
Principal Investigator: Marcello Restelli

Editorial Activities

- 2024-present **Action Editor** for Transactions on Machine Learning Research (TMLR)
- 2024-present **Area Chair** for NeurIPS
- 2025 **Area Chair** for ICML, RLC (Reinforcement Learning Conference)
- 2025 Reviewer for AISTATS
- 2024-present **Expert reviewer** for Transactions on Machine Learning Research (TMLR)
- 2024 Reviewer for ICML, COLT, Senior Reviewer for RLC (Reinforcement Learning Conference)
- 2022-2024 Reviewer for Transactions on Machine Learning Research (TMLR)
- 2023 Reviewer for ICML, COLT, EWRL, NeurIPS, ICLR, AISTATS
- 2022 Reviewer for IEEE Transactions on Automatic Control (TACON)
- 2022 Reviewer for ICML, NeurIPS, EWRL
- 2021 **Outstanding Reviewer** (top 8%) for NeurIPS, **expert reviewer** for ICML, reviewer for AISTATS, emergency reviewer for AAAI
- 2020 **Top 33% reviewer** for ICML, reviewer for NeurIPS, AISTATS, AAAI, UAI, ECAI
- 2019 Reviewer for ICML, NeurIPS, UAI

Organization of Scientific Events

- 2025 **The 36th International Conference on Algorithmic Learning Theory (ALT 2025)**, February 24-27 2025, Milan, Italy. Role: Local Chair, with co-chair Giulia Clerici (Ellis Unit Milan), <http://algorithmiclearningtheory.org/alt2025/>
- 2023 **Reinforcement Learning Summer School (RLSS 2023)**, June 28–July 5 2023, Universitat Pompeu Fabra, Barcelona, Spain. Role: Organizer, with co-organizer Vincent Adam (UPF), <https://rlsummerschool.com/>

Invited and Contributed Talks

- Sep 18, 2025 **Contributed talk at EWRL 2025**, (*European Workshop on Reinforcement Learning*) Tübingen, Germany, "Accelerating Policy Gradient Algorithms with Data Reuse"
- Jan 12, 2024 **Invited talk at the "Mini-Workshop on Reinforcement Learning"**, University of Mannheim, Germany
- May 25, 2023 **Invited talk at the Theory of Reinforcement Learning Workshop**, University of Alberta, Edmonton, Canada, titled "Offline Primal-Dual Reinforcement Learning for Linear MDPs"

- May 11, 2023 **Invited talk at the AI Seminars**, *Politecnico di Milano, Milan, Italy*, titled "Large-Scale Offline Reinforcement Learning"
- Sep 16, 2022 **Invited talk at the Reinforcement Learning Conference**, *Technische Universität Dresden, Dresden, Germany*, titled "Lifting the Information Ratio"
- Jul 1, 2021 **Invited talk at Mathematical Statistics and Learning**, *Barcelona Graduate School of Economics, Barcelona, Spain*, titled "Leveraging Good Representations in Linear Contextual Bandits and MDPs"
- Nov 6, 2020 **Invited talk at Gerhard Neumann's research seminar**, *Karlsruhe Institute of Technology (delivered online)*, titled "Safe Policy Optimization"
- Sep 19, 2019 **Invited talk at the Workshop on Markets, Algorithms, Prediction and LEarning (MAPLE)**, *Politecnico di Milano, Milan, Italy*, titled "Optimistic Policy Optimization via Multiple Importance Sampling"

Student Supervision and Mentoring

As PhD co-Advisor

- 2025-present **Leonardo Cesani**, Information Technology, Politecnico di Milano
Co-supervisor with Marcello Restelli
- 2024-present **Andrea Menta**, Information Technology, Politecnico di Milano
Co-supervisor with Marcello Restelli
- 2023-present **Alessandro Montenegro**, Information Technology, Politecnico di Milano
Co-supervisor with Alberto Maria Metelli

As Master Thesis Advisor

- 2023-2025 **Master theses**, supervised 4 master students for M.Sc. Computer Science and Engineering, one for M.Sc. Mathematical Engineering, Politecnico di Milano
Relatore of 5 deposited M.Sc. theses (<https://www.politesi.polimi.it/>)

As Master Thesis co-Advisor

- 2018-2020 **Master theses**, co-supervised several Master theses (M.Sc. Computer Science and Engineering, Mathematical Engineering, Management Engineering), Politecnico di Milano
- 2023-2025 *Correlatore* of 9 deposited M.Sc. theses (<https://www.politesi.polimi.it/>)

Other Mentoring Activities

Since 2021, as a postdoctoral researcher, I have collaborated with several (5+) PhD students at Universitat Pompeu Fabra and Politecnico di Milano.

Teaching

- 2025 **Lecturer**, *Università degli Studi di Milano, Milan, Italy*
Computer science for cultural mediators (Dati, web e contenuti digitali), B.Sc. Mediazione Linguistica e Culturale, 30 hours, I semester 2025/2026 (scheduled)
- 2025 **Lecturer**, *Politecnico di Milano, Milan, Italy*
Algorithms and data structures (Algoritmi e Principi dell'Informatica – Mod. 2), B.Sc. Computer Science and Engineering, 30 hours, II semester 2024/2025

- 2025 **Lecturer**, *Politecnico di Milano, Milan, Italy*
Advanced Deep Learning, PhD Information Technology, 10 hours, 2024/2025
- 2024 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Foundations of Artificial Intelligence, resp. Prof. Francesco Amigoni and Pierluca Lanzi, 16 hours, I semester 2024/2025
- 2024 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Informatica (computer science for civil engineering students), resp. Prof. Marcello Restelli, 24 hours, II semester 2023/2024
- 2024 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Machine Learning, resp. Prof. Marcello Restelli, 20 hours, II semester 2023/2024
- 2023 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Foundations of Artificial Intelligence, resp. Prof. Francesco Amigoni and Pierluca Lanzi, 16 hours, I semester 2023/2024
- 11-15 Jul 2022 **Teaching Assistant**, *Reinforcement Learning Summer School, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands*
Role: delivered a tutorial on implementing policy-gradient algorithms with JAX, tutored students in practical sessions.
- 2022 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Intelligenza Artificiale (Artificial Intelligence — online class), resp. Prof. Andrea Bonarini, 6 hours, attività di didattica integrativa, II semester 2021/2022
- 2021 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Intelligenza Artificiale (Artificial Intelligence — online class), resp. Prof. Andrea Bonarini, 6 hours, attività di didattica integrativa, II semester 2020/2021
- 2020 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Intelligenza Artificiale (Artificial Intelligence — online class), resp. Prof. Andrea Bonarini, 6 hours, attività di didattica integrativa, II semester 2019/2020
- 2019 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Informatica B (computer science for mechanical engineering students), resp. Prof. Luca Cassano, 26 hours, attività di didattica integrativa (esercitatori), I semester 2019/2020
- 2018 **Teaching Assistant**, *Politecnico di Milano, Milan, Italy*
Informatica B (computer science for mechanical engineering students), resp. Prof. Luca Cassano, 28 hours, attività di didattica integrativa (esercitatori), I semester 2018/2019
- 1-12 Jul 2019 **Teaching Assistant**, *Reinforcement Learning Summer School, Inria Lille-Nord Europe, Lille, France*, Role: tutored students in practical sessions
- 2018 **Teaching Assistant**, *Politecnico di Milano (Polo Territoriale di Como), Como, Italy*
Web and Internet Economics, resp. Prof. Nicola Gatti, 10 hours, attività di didattica integrativa (esercitatori), II semester 2017-2018
- 2017 **Lab Assistant**, *Politecnico di Milano, Milan, Italy*
Informatica B (computer science for mechanical engineering students), resp. Prof. Luca Cassano, 9 hours, attività di didattica integrativa sperimentale (ex responsabili di laboratorio), I semester 2017-2018
- 2016 **Lab Tutor**, *Politecnico di Milano, Milan, Italy*
Prova Finale-Ingegneria del Software (software engineering: final project), resp. Prof. Carlo Ghezzi, 32 hours, attività di tutorato (ex tutor di laboratorio), II semester 2015/2016

Participation on Committees

- April 16 2025 **PhD Jury Member**, *University of Liège, Belgium*
- April 10 2025 **President of committee for PhD defense**, *PhD program in Information Technology, Politecnico di Milano*
- 2024 **Examiner of PhD thesis**, *University of Melbourne, Australia*
- June 27 2024 **President of committee for PhD defense**, *PhD program in Information Technology, Politecnico di Milano*
- May 2024 **Committee member for assignment of industrial research funding**, *Bando a cascata del progetto "Future Artificial Intelligence Research – FAIR", (Bando Imprese), codice PE0000013, PNRR, Missione 4, Componente 2, Investimento 1.3*
- 2023-2025 **Committee member for degree examinations**, *B. Sc. and M. Sc. in Computer Science and Engineering, Politecnico di Milano (several instances)*
- Mar 30 2023 **Board member for the Thesis Proposal Defenses**, *PhD program in Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain*
- 2023 **Jury Member for the CLAIRE R2Net 2022 Papers Highlights**
- 2021, 2022 **Evaluator** for the pre-screening of applicants to the ELLIS PhD program
- July 5 2022 **Jury member for the undergraduate thesis defenses**, *Universitat Pompeu Fabra, Barcelona, Spain*
- Mar 17 2022 **Board member for the Thesis Proposal Defense**, *PhD program in Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain*

Attendance to Conferences and Workshops

- 2025 **EWRL 18**, *Tübingen, Germany*, 1 contributed talk, 1 poster
- 2025 **RLC (Reinforcement Learning Conference)**, *Edmonton, Canada*, 1 poster + 1 spotlight talk at the Workshop on Programmatic Reinforcement Learning
- 2024 **NeurIPS**, *Vancouver, Canada*, 2 posters
- 2024 **ICML**, *Vienna, Austria*, 2 posters + 1 at the ARLET workshop
- 2024 **AISTATS**, *Valencia, Spain*, 1 poster
- 2024 **ALT**, *San Diego, USA*, 1 poster
- 2024 **Mini Workshop on Reinforcement Learning**, *Mannheim, Germany*, 1 invited talk
- 2023 **EWRL 16**, *Brussels, Belgium*, 2 posters
- 2023 **AISTATS**, *Valencia, Spain*
- 2023 **Upper Bound**, *Edmonton, Canada*, invited talk at the RL Theory Workshop
- 2023 **ALT**, *Singapore*, attended as co-author and served as **session chair**
- 2022 **NeurIPS**, *New Orleans, USA*, presented 2 posters in poster sessions
- 2022 **ELLIS ILIR Workshop**, *Feldberg, Germany*
- 2022 **EWRL 15**, *Milan, Italy*, presented 2 posters in poster sessions
- 2022 **Reinforcement Learning Conference**, *Dresden, Germany*, invited talk
- 2022 **ICML-2022 Workshop on Complex Feedback in Online Learning**, *Baltimore, USA*, presented 1 poster in poster session

- 2021 **NeurIPS**, *online edition*, presented 1 poster in online poster session and recorded 1 spotlight talk
- 2021 **Mathematical Statistics and Learning**, *Barcelona, Spain*, invited talk
- 2021 **ICML**, *online edition*, presented 1 poster in online poster session and recorded 1 talk
- 2021 **AAAI**, *online edition*, presented 1 poster in online poster session
- 2021 **IJCAI**, *online edition*, presented 1 poster in online poster session
- 2020 **NeurIPS**, *online edition*, attended online
- 2020 **AISTATS**, *online edition*, recorded 1 talk
- 2019 **Workshop MAPLE**, *Milan, Italy*, invited talk
- 2019 **ICML**, *Long Beach, USA*, presented 1 poster and delivered 1 oral presentation
- 2018 **NeurIPS**, *Montreal, Canada*, presented 1 poster and delivered 1 oral presentation
- 2018 **EWRL 14**, *Lille, France*, presented 1 poster
- 2018 **ICML**, *Stockholm, Sweden*, presented 1 poster and delivered 1 oral presentation
- 2017 **NeurIPS**, *Long Beach, USA*, presented 1 poster

Complete List of Papers

Journal Papers

- [J1] L. Civitavecchia, M. Papini. "Exploration-Free Reinforcement Learning with Linear Function Approximation." *Reinforcement Learning Journal*, 2025 (to appear).
- [J2] G. Tedeschi, M. Papini, A. M. Metelli, and M. Restelli. Search or Split: Policy Gradient with Adaptive Policy Space. *Machine Learning* 114, 186 (2025). <https://doi.org/10.1007/s10994-025-06820-2>
- [J3] G. Paczolat, M. Papini, A. M. Metelli, I. Harmati, and M. Restelli. Sample Complexity of Variance-Reduced Policy Gradient: Weaker Assumptions and Lower Bounds. *Machine Learning* 113, 6475–6510 (2024). <https://doi.org/10.1007/s10994-024-06573-4>
- [J4] M. Papini, G. Manganini, A. M. Metelli, and M. Restelli. "Policy Gradient with Active Importance Sampling." *Reinforcement Learning Journal*, vol. 2, 2024, pp. 645–675. <https://rlj.cs.umass.edu/2024/papers/Paper90.html>
- [J5] M. Papini, M. Pirotta, and M. Restelli. Smoothing policies and safe policy gradients. *Machine Learning* 111, 4081–4137 (2022). <https://doi.org/10.1007/s10994-022-06232-6>
- [J6] A. M. Metelli, M. Papini, N. Montali, M. Restelli. Importance Sampling Techniques for Policy Optimization. *Journal of Machine Learning Research* 21(141):1–75, 2020. <https://www.jmlr.org/papers/v21/20-124.html>

Conference Papers

- [C1] A. Montenegro, M. Mussi, M. Papini, and A. M. Metelli. Convergence analysis of policy gradient methods with dynamic stochasticity. In *ICML*, 2025 (to appear).
- [C2] G. Neu, M. Papini, and L. Schwartz. Optimistic information directed sampling. In

- COLT*, volume 247 of *Proceedings of Machine Learning Research*, pages 3970–4006. PMLR, 2024.
- [C3] A. Montenegro, M. Mussi, M. Papini, and A. M. Metelli. Last-iterate global convergence of policy gradients for constrained reinforcement learning. In *NeurIPS*, 2024.
 - [C4] A. Montenegro, M. Mussi, A. M. Metelli, and M. Papini. Learning optimal deterministic policies with stochastic policy gradients. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pages 36160–36211. PMLR, 21–27 Jul 2024.
 - [C5] D. Maran, A. M. Metelli, M. Papini, and M. Restelli. Projection by convolution: Optimal sample complexity for reinforcement learning in continuous-space mdps. In *COLT*, volume 247 of *Proceedings of Machine Learning Research*, pages 3743–3774. PMLR, 2024.
 - [C6] D. Maran, A. M. Metelli, M. Papini, and M. Restelli. No-regret reinforcement learning in smooth MDPs. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pages 34760–34789. PMLR, 21–27 Jul 2024.
 - [C7] D. Maran, A. M. Metelli, M. Papini, and M. Restelli. Local linearity: the key for no-regret reinforcement learning in continuous mdps. In *NeurIPS*, 2024.
 - [C8] G. Gabbianelli, G. Neu, M. Papini, and N. Okolo. Offline primal-dual reinforcement learning for linear mdps. In *AISTATS*, volume 238 of *Proceedings of Machine Learning Research*, pages 3169–3177. PMLR, 2024.
 - [C9] G. Gabbianelli, G. Neu, and M. Papini. Importance-weighted offline learning done right. In *ALT*, volume 237 of *Proceedings of Machine Learning Research*, pages 614–634. PMLR, 2024.
 - [C10] F. Bacchiocchi, F. E. Stradi, M. Papini, A. M. Metelli, and N. Gatti. Online learning with off-policy feedback in adversarial mdps. In K. Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 3697–3705. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Main Track.
 - [C11] G. Gabbianelli, G. Neu, and M. Papini. Online learning with off-policy feedback. In *ALT*, volume 201 of *Proceedings of Machine Learning Research*, pages 620–641. PMLR, 2023.
 - [C12] A. Tirinzoni, M. Papini, A. Touati, A. Lazaric, and M. Pirodda. Scalable representation learning in linear contextual bandits with constant regret guarantees. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:2307–2319, 2022.
 - [C13] G. Neu, I. Olkhovskaia, M. Papini, and L. Schwartz. Lifting the information ratio: An information-theoretic analysis of thompson sampling for contextual bandits. *Advances in Neural Information Processing Systems (NeurIPS)*, 35:9486–9498, 2022.

- [C14] M. Papini, A. Tirinzoni, M. Restelli, A. Lazaric, and M. Pirotta. Leveraging good representations in linear contextual bandits. In *ICML*, volume 139 of *Proceedings of Machine Learning Research*, pages 8371–8380. PMLR, 2021.
- [C15] M. Papini, A. Tirinzoni, A. Pacchiano, M. Restelli, A. Lazaric, and M. Pirotta. Reinforcement learning in linear mdps: Constant regret and representation selection. In *NeurIPS*, pages 16371–16383, 2021.
- [C16] A. M. Metelli, M. Papini, P. D’Oro, and M. Restelli. Policy optimization as online learning with mediator feedback. In *AAAI*, pages 8958–8966. AAAI Press, 2021.
- [C17] M. Papini, A. Battistello, and M. Restelli. Balancing learning speed and stability in policy gradient via adaptive exploration. In *AISTATS*, volume 108 of *Proceedings of Machine Learning Research*, pages 1188–1199. PMLR, 2020.
- [C18] P. D’Oro, A. M. Metelli, A. Tirinzoni, M. Papini, and M. Restelli. Gradient-aware model-based policy search. In *AAAI*, pages 3801–3808. AAAI Press, 2020.
- [C19] L. Bisi, L. Sabbioni, E. Vittori, M. Papini, and M. Restelli. Risk-averse trust region optimization for reward-volatility reduction. In *IJCAI*, pages 4583–4589. ijcai.org, 2020.
- [C20] M. Papini, A. M. Metelli, L. Lupo, and M. Restelli. Optimistic policy optimization via multiple importance sampling. In *ICML*, volume 97 of *Proceedings of Machine Learning Research*, pages 4989–4999. PMLR, 2019.
- [C21] M. Beraha, A. M. Metelli, M. Papini, A. Tirinzoni, and M. Restelli. Feature selection via mutual information: New theoretical insights. In *IJCNN*, pages 1–9. IEEE, 2019.
- [C22] M. Papini, D. Binaghi, G. Canonaco, M. Pirotta, and M. Restelli. Stochastic variance-reduced policy gradient. In *ICML*, volume 80 of *Proceedings of Machine Learning Research*, pages 4023–4032. PMLR, 2018.
- [C23] A. M. Metelli, M. Papini, F. Faccio, and M. Restelli. Policy optimization via importance sampling. In *NeurIPS*, pages 5447–5459, 2018.
- [C24] M. Papini, M. Pirotta, and M. Restelli. Adaptive batch size for safe policy gradients. In *NeurIPS*, pages 3591–3600, 2017.

Workshop Papers

- [W1] D. Salaorni, V. De Paola, S. Delpero, G. Dispoto, P. Bonetti, A. Russo, G. Calcagno, F. Trovò, M. Papini, A. M. Metelli, M. Mussi, M. Restelli. Gym4ReaL: A Benchmark Suite for Evaluating Reinforcement Learning in Realistic Domains. EWRL 18 (European Workshop on Reinforcement Learning), Tübingen, Germany, 2025.
- [W2] P. Olivieri, F. Lasca, A. Gianola, M. Papini. Do it for HER: First-order Logic Reward Specification in Reinforcement Learning. RLC Workshop on Programmatic Reinforcement Learning, Edmonton, Canada, 2025.
- [W3] L. Bolis, S. Livella, S. Patania, D. Ognibene, M. Papini, K. Morita. Balancing Benefits and Risks: RL Approaches for Addiction-Aware Social Media Recommenders. RLDM (Multidisciplinary Conference on Reinforcement Learning and Decision Making), Dublin, Ireland, 2025.

- [W4] M. Molaei, G. Paczolat, M. Papini, A. M. Metelli, M. Restelli. Actor-Critic with Active Importance Sampling. RLDM (Multidisciplinary Conference on Reinforcement Learning and Decision Making), Dublin, Ireland, 2025.
- [W5] G. Tedeschi, M. Papini, A. M. Metelli, M. Restelli. Policy Gradient Methods with Adaptive Policy Spaces. EWRL 17 (European Workshop on Reinforcement Learning), Toulouse, France, 2024.
- [W6] G. Neu, M. Papini, L. Schwartz. Optimistic Information Directed Sampling. FoRLaC (Foundations of Reinforcement Learning and Control) workshop at ICML, Vienna, Austria, 2024.
- [W7] G. Tedeschi, M. Papini, A. M. Metelli, M. Restelli. Policy Gradient Methods with Adaptive Policy Spaces. ARLET (Aligning Reinforcement Learning Experimentalists and Theorists) workshop at ICML, Vienna, Austria, 2024.
- [W8] F. Bacchiocchi, F. E. Stradi, M. Papini, A. M. Metelli, N. Gatti. Online Adversarial MDPs with Off-Policy Feedback and Known Transitions. 16th European Workshop on Reinforcement Learning, Brussels, Belgium, 2023
- [W9] G. Gabbianelli, G. Neu, N. Okolo, M. Papini. Offline Primal-Dual Reinforcement Learning for Linear MDPs. 16th European Workshop on Reinforcement Learning, Brussels, Belgium, 2023
- [W10] G. Neu, J. Olkhovskaya, M. Papini and L. Schwartz. Lifting the Information Ratio: An Information-Theoretic Analysis of Thompson Sampling for Contextual Bandits. 15th European Workshop on Reinforcement Learning, Milan, Italy, 2022
- [W11] A. Tirinzoni, M. Papini, A. Touati, A. Lazaric, and M. Pirotta. Scalable Representation Learning in Linear Contextual Bandits with Constant Regret Guarantees. 15th European Workshop on Reinforcement Learning, Milan, Italy, 2022
- [W12] G. Gabbianelli, M. Papini, G. Neu. Online Learning with Off-Policy Feedback. ICML-2022 workshop on Complex Feedback in Online Learning, Baltimore, USA, 2022
- [W13] A. Gianola, M. Montali, and M. Papini. Automated Reasoning for Reinforcement Learning Agents in Structured Environments. OVERLAY workshop on fOrmal VERification, Logic, Automata and sYNthesis, Padova, Italy, 2021
- [W14] M. Papini, A. Tirinzoni, A. Pacchiano, M. Restelli, A. Lazaric, and M. Pirotta. Reinforcement Learning in Linear MDPs: Constant Regret and Representation Selection. ICML Workshop on Reinforcement Learning Theory, virtual, 2021
- [W15] M. Papini, A. Battistello, and M. Restelli. Safe Exploration in Gaussian Policy Gradient. NeurIPS-2019 Workshop on Safety and Robustness in Decision Making, Vancouver, Canada, 2019
- [W16] M. Papini, A. Battistello, and M. Restelli. Safely Exploring Policy Gradient. 14th European Workshop on Reinforcement Learning, Lille, France, 2018

Papers In Preparation

- [P1] With A. Montenegro, F. Mansutti, M. Mussi, A. M. Metelli. Reusing Trajectories in Policy Gradients Enables Fast Convergence. Preprint: <https://arxiv.org/abs/2506.06178>
- [P2] With D. Salaorni, V. De Paola, S. Delpero, G. Dispoto, P. Bonetti, A. Russo, G. Calcagno, F. Trovò, M. Papini, A. M. Metelli, M. Mussi, M. Restelli. Gym4ReaL: Towards Real-World Reference Environments for Reinforcement Learning. Preprint: <https://arxiv.org/abs/2507.00257>
- [P3] With P. Olivieri, F. Lasca, A. Gianola. Do it for HER: First-order Logic Reward Specification in Reinforcement Learning
- [P4] With A. Montenegro, L. Cesani, M. Mussi, A. M. Metelli. Learning Deterministic Policies with Policy Gradients in Constrained Markov Decision Processes. Preprint: <https://arxiv.org/abs/2506.05953>

Attendance to Summer Schools and Exchange Programs

- Jul 24–Aug 2 2018 **CIFAR Deep Learning and Reinforcement Learning Summer School**, Toronto, Canada, presented 1 poster
- 7-14 Oct 2017 **ACAI Summer School on Reinforcement Learning**, Nieuwpoort, Belgium
- Autumn 2016 **Erasmus Programme**, KTH Royal Institute of Technology, Stockholm, Sweden

Languages

Italian, *native speaker*

English, *fluent (Cambridge FCE, Grade A — C1 CEFR level)*

Spanish, *intermediate (completed course of level B1 at UPF Barcelona)*

Programming Languages

Python (several research projects) including experience with deep learning libraries **TensorFlow**, **PyTorch** and **JAX**, **LaTeX** (typesetting of several manuscripts), **C++** (one research project, one robotics project as master student), **C**, **MATLAB** (taught as TA), **Java** (final project of B.Sc.).

Other Interests

Climbing, running, piano, science fiction.