

---

## Contact Information

Work Address Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano,  
Via Camillo Golgi 39 (Edificio 21), 20133 Milan, Italy

Email [matteo.papini@polimi.it](mailto:matteo.papini@polimi.it)

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

---

## Highlights

Matteo Papini is an **Assistant Professor** (RTD-A) at **Politecnico di Milano**, Milan, Italy, in the Artificial Intelligence and Robotics Lab. His research is directed towards the development of intelligent systems, with a focus on the problem of *sequential decision making under uncertainty*, using the theory and algorithmic solutions of **reinforcement learning** and **contextual bandits**. In 2021 he earned a PhD in Information Technology (cum laude) from Politecnico di Milano, Milan, Italy, under the supervision of Marcello Restelli, 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, in Gergely Neu's team.

He has authored more than twenty peer-reviewed conference papers, including publications at top artificial intelligence and machine learning conferences such as **NeurIPS (5 papers)**, **ICML (5)**, **COLT (2)**, **AAAI (2)** and **IJCAI (2)**. Of these, one was awarded an oral presentation at NeurIPS 2018, which was only granted to the top 3% of the accepted papers. Moreover, he has published peer-reviewed manuscripts at renowned peer-reviewed machine learning journals — one on **JMLR** and two on Springer's **Machine Learning**. Among his recurring co-authors are Marcello Restelli (Politecnico di Milano), Matteo Pirota, Alessandro Lazaric (Meta), and Gergely Neu (UPF).

He is a member of **ELLIS** (European Laboratory for Learning and Intelligent Systems). He has worked as a teaching assistant for several classes at Politecnico di Milano and for two international reinforcement learning summer schools. He has served as a reviewer in the program committee of several conferences including NeurIPS, ICML and AAAI. In 2023, with Vincent Adam (UPF), he led the organization of the **Reinforcement Learning Summer School** (RLSS 2023) at UPF. He's going to be the **local chair** of ALT 2025 (International Conference on Algorithmic Learning Theory) in Milan.

---

## Faculty and Research Positions

- 2023–present **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 Pirodda. 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  
Supervisor: 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  
This school ranked first in Italy for several years according to Eduscopio (<https://eduscopio.it/>)

---

## Author Profile

- Google Scholar **h-index 13**, <https://scholar.google.it/citations?user=A2WxZlsAAAAJ>
- Scopus <https://www.scopus.com/authid/detail.uri?authorId=57202057824>
- dblp <https://dblp.uni-trier.de/pid/209/4897.html>

## Summary of Publications

Journal/Conference	# Papers	Rating	Homepage
JMLR	1	Scimago Q1, CORE A*	<a href="http://www.jmlr.org/">http://www.jmlr.org/</a>
Machine Learning	2	Scimago Q1, CORE A	<a href="https://www.springer.com/journal/10994">https://www.springer.com/journal/10994</a>
NeurIPS	5	CORE A*, GGS A++	<a href="https://nips.cc/">https://nips.cc/</a>
ICML	5	CORE A*, GGS A++	<a href="https://icml.cc/">https://icml.cc/</a>
AAAI	2	CORE A*, GGS A++	<a href="https://aaai.org/Conferences/AAAI-23/">https://aaai.org/Conferences/AAAI-23/</a>
IJCAI	2	CORE A*, GGS A++	<a href="https://www.ijcai.org/">https://www.ijcai.org/</a>
COLT	2	CORE A*, GGS A+	<a href="https://learningtheory.org/">https://learningtheory.org/</a>
AISTATS	2	CORE A, GGS A+	<a href="https://aistats.org/">https://aistats.org/</a>
IJCNN	1	CORE B, GGS A-	<a href="https://2023.ijcnn.org/">https://2023.ijcnn.org/</a>
ALT	2	CORE B, GGS B	<a href="http://algorithmiclearningtheory.org/">http://algorithmiclearningtheory.org/</a>

## Selected Publications (12)

- [1] **M. Papini**, D. Binaghi, G. Canonaco, M. Pirotta, M. Restelli: *Stochastic variance-reduced policy gradient*. **ICML** (2018) — **CORE A\*, GGS A++**
- [2] A.M. Metelli, **M. Papini**, F. Faccio, M. Restelli: *Policy optimization via importance sampling*. **NeurIPS (oral, top 3%)** (2018) — **CORE A\*, GGS 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\*, GGS 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\*, GGS A++**
- [6] P. D'Oro, A. M. Metelli, A. Tirinzoni, **M. Papini**, M. Restelli: *Gradient-aware model-based policy search*. **AAAI** (2020) — **CORE A\*, GGS A++**
- [7] **M. Papini**, A.M. Metelli, L. Lupo, M. Restelli: *Optimistic policy optimization via multiple importance sampling*. **ICML** (2019) — **CORE A\*, GGS A++**
- [8] **M. Papini**, M. Pirotta, M. Restelli: *Smoothing policies and safe policy gradients*. Springer's **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\*, GGS 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\*, GGS 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\*, GGS 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*. Springer's **Machine Learning** (2024) — **Scimago Q1**

---

## Awards, Grants, Scholarships, 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. Total amount granted for 2 years: 64.800€  
<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 in Funded Projects

- 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

---

## Organization of Scientific Events

- 2025 **International Conference on Algorithmic Learning Theory (ALT)**, *Politecnico di Milano, Milan, Italy. Role: Local Chair (to be announced)*,  
<http://algorithmiclearningtheory.org/>

- 2023 **Reinforcement Learning Summer School (RLSS)**, June 28–July 5 2023, *Universitat Pompeu Fabra, Barcelona, Spain*. Role: Main Organizer (co-organized with Vincent Adam), <https://rlsummerschool.com/>
- 2023 **ELLIS Pre-NeurIPS Fest**, December 4 2023, *ELLIS Unit of Milan*. Role: co-organizer, <https://www.ellis-milan.eu/2023/12/05/ellis-pre-neurips-fest-2023-in-milan/>

## Invited Talks

- 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"

## Teaching Experience

- 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

- June 2024 **Committee member for PhD defense**, *PhD program in Information Technology, Politecnico di Milano (nominated)*
- 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 (nominated)*
- Dec 19 2023 **Committee member for degree examination**, *M. Sc. in Computer Science and Engineering, Politecnico di Milano*
- 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*

## Editorial Activities

- 2024 **Area Chair** for NeurIPS
- 2024 Reviewer for ICML, COLT, Senior Reviewer for RLC (Reinforcement Learning Conference)

- 2022-present 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

## Attendance to Conferences and Workshops

- 2024 **ICML**, *Vienna, Austria*, 2 posters (planned)
- 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



---

## 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

---

## Student Supervision and Mentoring

### Supervision of PhD Students

- 2023-present **Alessandro Montenegro**, Information Technology, Politecnico di Milano  
Co-supervisor: Alberto Maria Metelli

### Supervision of Master Students

- 2023-present **Currently (co-)supervising 5+ master students**, Politecnico di Milano

### Co-supervision/Mentoring

- 2023-present **5+ PhD students**, Politecnico di Milano
- 2021-2023 **3 PhD students**, Universitat Pompeu Fabra, Barcelona
- 2018-2020 **8 master students**, Politecnico di Milano  
*Correlatore* of 8 deposited M.Sc. theses.

---

## Complete List of Papers

### Journal Papers

- [J1] G. Paczoly, M. Papini, A. M. Metelli, I. Harmati, and M. Restelli. Sample Complexity of Variance-Reduced Policy Gradient: Weaker Assumptions and Lower Bounds. *Machine Learning*, 1-36 2024, (<https://doi.org/10.1007/s10994-024-06573-4>)
- [J2] M. Papini, M. Pirotta, and M. Restelli. Smoothing policies and safe policy gradients. *Machine Learning*, 2022, 1-57 (<https://rdcu.be/c0a7S>)
- [J3] A. M. Metelli, M. Papini, N. Montali, M. Restelli. Importance Sampling Techniques for Policy Optimization. *Journal of Machine Learning Research (JMLR)* 21.141., pp. 1-75, 2020

### Conference Papers

- [C1] M. Papini, G. Manganini, A. M. Metelli, and M. Restelli. Policy gradient with active importance sampling. In *Reinforcement Learning Conference (RLC)*, 2024 (to appear).
- [C2] G. Neu, M. Papini, and L. Schwartz. Optimistic information directed sampling. In *COLT*, 2024 (to appear).
- [C3] A. Montenegro, M. Mussi, A. Metelli, and M. Papini. Learning optimal deterministic policies with stochastic policy gradients. In *ICML*, 2024 (to appear).
- [C4] D. Maran, A. Metelli, M. Papini, and M. Restelli. Projection by convolution: Optimal sample complexity for rl in continuous-space mdps. In *COLT*, 2024 (to appear).



- [C5] D. Maran, A. Metelli, M. Papini, and M. Restelli. No-regret reinforcement learning in smooth mdps. In *ICML*, 2024 (to appear).
- [C6] F. Bacchiocchi, F. E. Stradi, M. Papini, A. M. Metelli, and N. Gatti. Online learning with off-policy feedback in adversarial mdps. In *IJCAI*, 2024 (to appear).
- [C7] 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.
- [C8] 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.
- [C9] 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.
- [C10] A. Tirinzoni, M. Papini, A. Touati, A. Lazaric, and M. Pirotta. Scalable representation learning in linear contextual bandits with constant regret guarantees. *Advances in Neural Information Processing Systems*, 35:2307–2319, 2022.
- [C11] 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*, 35:9486–9498, 2022.
- [C12] 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.
- [C13] 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.
- [C14] 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.
- [C15] 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.
- [C16] 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.
- [C17] 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.
- [C18] 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.

- [C19] 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.
- [C20] 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.
- [C21] A. M. Metelli, M. Papini, F. Faccio, and M. Restelli. Policy optimization via importance sampling. In *NeurIPS*, pages 5447–5459, 2018.
- [C22] M. Papini, M. Pirotta, and M. Restelli. Adaptive batch size for safe policy gradients. In *NeurIPS*, pages 3591–3600, 2017.

#### Workshop Papers

- [W1] G. Neu, M. Papini, L. Schwartz. Optimistic Information Directed Sampling. FoRLaC (Foundations of Reinforcement Learning and Control) workshop at ICML, Vienna, Austria, 2024.
- [W3] 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
- [W4] 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
- [W5] 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
- [W6] 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
- [W7] G. Gabbianelli, M. Papini, G. Neu. Online Learning with Off-Policy Feedback. ICML-2022 workshop on Complex Feedback in Online Learning, Baltimore, USA, 2022
- [W8] 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
- [W9] 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
- [W10] 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
- [W11] M. Papini, A. Battistello, and M. Restelli. Safely Exploring Policy Gradient. 14th European Workshop on Reinforcement Learning, Lille, France, 2018

#### Papers In Preparation

- [P1] A. Montenegro, M. Mussi, M. Papini, A. M. Metelli. Last-Iterate Global Convergence of Policy Gradients for Constrained Reinforcement Learning. 2024.
- [P2] D. Maran, A. M. Metelli, M. Papini, M. Restelli. Local Linearity: the Key for No-regret Reinforcement Learning in Continuous MDPs. 2024.
- [P3] G. Tedeschi, M. Papini, A. M. Metelli, M. Restelli. Search or Split: Policy Gradient with Adaptive Policy Spaces. 2024.
- [P4] G. Paczolat, M. Papini, A. M. Metelli, I. Harmati, M. Restelli. Stabilizing Policy Gradient with Active Importance Sampling. 2024.

## 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** and **PyTorch**, **LaTeX** (typesetting of several manuscripts), **C++** (one research project, one mobile robotics project as master student), **C**, **MATLAB** (taught as TA), **Java** (final project of B.Sc.).

## Hobbies

Climbing, piano.