

Dr.-Ing. Samuele Tosatto

9837 110 St. NW (Apt 604), T5K 2L8, Edmonton, Canada • Email: samuele.tosatto@gmail.com

Postdoc Researcher

RESEARCH INTERESTS

My goal is to enable robots to learn in the real-world. To this end, the learning scheme must be extremely sample efficient. I achieve sample efficiency by working on the theory of **policy gradient** algorithms, by using specific modelling of robotic actions **movement primitives**, and by considering **hierarchical architectures** that abstract time and action-space. My research can be seen as an intersection between reinforcement learning, machine learning and robotics.

ACADEMIC CARRIER

2021-Present	University of Alberta, Canada Postdoc-fellow , supervised by Prof. Rupam Mahmood
2017-2020	Technische Universität Darmstadt, Germany, Ph.D. In Computer Science (Magna cum laude) Supervisor: Prof. Jan Peters , external committee member Prof. Martha White . thesis , diploma
2016	ERASMUS Johannes Kepler Universität, Linz, Austria
2013-2017	Polytechnic University of Milan, Italy, Master in Software Engineering (110/110) thesis , diploma , grades , supplement
2009-2012	Polytechnic University of Milan, Italy, Bachelor in Software Engineering (96/110) diploma

ADDITIONAL EXPERIENCE

2017 SPP 1527 [Summer School on Robotics & Autonomous Learning](#)

PROGRAMMING SKILLS

I am a passionate programmer, and, thanks to my strong background in software engineering, I am able to produce clean code both for numerical simulations of experiments, and to interface with robots. Some of my libraries are currently used by the students. Some examples:

HeRL (Reinforcement Learning Helper) - A library that helps analyze and develop reinforcement learning algorithms.

robopy - A library that interfaces with ROS, executes robotic trajectory, reads sensory input, and interacts with the robot using voice commands.

LAMPO (LAtent Movement Policy Optimization) - Algorithm and numerical evaluation of our publication [C3].

Programming Languages: Python; Java; C#;

Libraries/Frameworks/Revision Controls: Numpy; Tensorflow; PyTorch; Scipy; Scikit-Learn; CVXPY; Cython; V-REP/Coppelia; ROS (and rospy); Git.

GEOGRAPHIC AND THEMATIC MOBILITY

I worked in three research laboratories: Artificial Intelligence and Robotic ([AIRLab](#)), Politecnico di Milano (Italy), under the supervision of [Prof. Marcello Restelli](#); Intelligent Autonomous Systems ([IAS](#)), Technische Universität Darmstadt (Germany), under the supervision of [Prof. Jan Peters](#); Reinforcement Learning and Artificial Intelligence ([RLAI](#)), University of Alberta (Canada), under the supervision of [Prof. Rupam Mahmood](#), which is formed by some of the most influential professors in my field, such as: [Prof. Richard Sutton](#), [Prof. Csaba Szepesvari](#), [Prof. Michael Bowling](#), [Prof. Martha White](#), to mention a few.

I published paper on a large range of topics in my field, such as sub-fields of reinforcement learning like: **exploration/exploitation** [C5], **value-based** methods [C7], **policy gradients** [W4,C4,J1,C1,C2,W3,C8,W1], **continuous-time** [C10], **hierarchical reinforcement learning** [C3, C10]; machine learning tools for robotics such as **movement primitives** [C3,C9]; and **statistical machine learning**, analyzing the asymptotic bias of a particular nonparametric technique [J2].

RESEARCH NETWORK

I have published papers with many researcher across Europe and North America: [Prof. Jan Peters](#), [Prof. Georgia Chalvatzaki](#), [Dr. Carlo D'Eramo](#), TU Darmstadt, Germany; [Prof. Marcello Restelli](#), Politecnico di Milano, Italy; [Prof. Rupam Mahmood](#) and [Prof. Martha White](#), University of Alberta, Canada; [Prof. Elmar Rückert](#), University of Leoben, Austria; [Prof. Joni Pajarinen](#), Aalto University, Finland; [Dr. Riad Akrouf](#), Research Scientist in INRIA, France; [Dr. Matteo Pirotta](#), Research Scientist at FAIR, Facebook, France. I am keeping active my research network, with active projects and thesis with other groups. I also have experience with industries, since my Ph.D. was funded by [Bosch](#), and I am currently co-supervising a research project with [Huawei](#).

NATURAL LANGUAGES

Italian (Mother tongue), English (C1), German (B1).

STUDENT SUPERVISION

2022	Master Thesis	Analysis of an Alternate Policy Gradient Estimator for Softmax Policies Shivam Garg
2020	Bachelor Thesis	Dimensionality Reduction of Movement Primitives in Parameter Space Jonas Stadtmüller
2019	Master Thesis	A Nonparametric Off-Policy Policy Gradient Joao Carvalho
2019	Master Thesis	Improving Sample-Efficiency with a Model-Based Deterministic Policy Gradient Saoud Hussam
2018	Master Thesis	Boosted Deep Q-Network Jeremy Tschirner

I currently co-supervise three Ph.D. students: [Qingfeng Lan](#), [Blanca Miller](#), and [Michael Przystupa](#).

REVIEWING

I serve as reviewer at some of the most prestigious journals (**TPAMI**, **RA-L**) and conferences (**ICLR**, **ICML**, **NeurIPS**, **AISTATS**, **ICRA**, **IROS**, **HUMANOIDS**) in our field.

INVITED SEMINAR TALKS

February 2022	RobotLearn Lab , INRIA, Grenoble, France.
February 2021	AMLab , University of Amsterdam, Amsterdam, The Netherlands.
December 2020	BioRob Lab , EPFL, Lausanne, Switzerland.

PUBLICATION LIST

JOURNALS

- J1** **Tosatto, S.**; Carvalho, J.; Peters, J. (2021). Batch Reinforcement Learning with a Nonparametric Off-Policy Policy Gradient, *Transaction of Pattern Analysis and Machine Intelligence (TPAMI)* [IEEE Xplore](#), [arXiv](#)
- J2** **Tosatto, S.**; Akrou, R.; Peters, J. (2021). An Upper Bound on Nadaraya-Watson Kernel Regression under Lipschitz Assumptions *MDPI Stats* [MDPI](#), [arXiv](#)

CONFERENCES

- C1** Lan, Q.; **Tosatto, S.**; Farrahi, H.; Mahmood AR. (2022) Model-Free Policy Learning with Reward Gradients *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)* [arXiv](#)
- C2** Garg, S.; **Tosatto, S.**; Pan, Y.; White, M.; Mahmood AR. (2022) An Alternate Policy Gradient Estimator for Softmax Policies *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)* [arXiv](#)
- C3** **Tosatto, S.**; Chalvatzaki, G.; Peters, J. (2021) Contextual Latent-Movements Off-Policy Optimization for Robotic Manipulation Skills *Proceeding of the International Conference on Robotics and Automation (ICRA)* [IEEE Xplore](#), [arXiv](#)
- C4** **Tosatto, S.**; Carvalho, J.; Abdulsamad, H.; Peters, J. (2020). A Nonparametric Off-Policy Policy Gradient, *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)* [mlr](#), [arXiv](#)
- C5** **Tosatto, S.**; D'Eramo, C.; Pajarinen, J.; Restelli, M.; Peters, J. (2019). Exploration Driven By an Optimistic Bellman Equation, *Proceedings of the International Joint Conference on Neural Networks (IJCNN)* [IEEE Xplore](#), [tu-darmstadt](#)
- C6** Rueckert, E.; Nakatenus, M.; **Tosatto, S.**; Peters, J. (2017). Learning Inverse Dynamics Models in O(n) time with LSTM networks, *Proceedings of the IEEE International Conference on Humanoid Robots (HUMANOIDS)* [IEEE Xplore](#), [tu-darmstadt](#)
- C7** **Tosatto, S.**; Pirotta, M.; D'Eramo, C; Restelli, M. (2017). Boosted Fitted Q-Iteration, *Proceedings of the International Conference of Machine Learning (ICML)* [mlr](#), [tu-darmstadt](#)

WORKSHOPS

- W1** **Tosatto, S.**; Patterson, A.; White, M.; Mahmood, R. (2022) A Gradient Critic for Policy Gradient Estimation, Workshop on Reinforcement Learning and Decision Making (**RLDM**).
- W2** Lang, Q.; **Tosatto, S.**; Homayoon, F.; Mahmood, R. (2022) Model-free Policy Learning with Reward Gradients, Workshop on Reinforcement Learning and Decision Making (**RLDM**).
- W3** Lan, Q.; **Tosatto, S.**; Yangchen, P.; White, M.; Mahmood, R. (2022) Making Policy Gradient Estimators for Softmax Policies Robust to Non-stationarities, Workshop on Reinforcement Learning and Decision Making (**RLDM**).
- W4** **Tosatto, S.**, João Carvalho, Jan Peters (2019) A Nonparametric Off-Policy Policy Gradient, *Workshop on Reinforcement Learning and Decision Making (RLDM)*.

UNDER REVISION OR IN PREPARATION

- C8 **Tosatto, S.**; Patterson, A.; White, M.; Mahmood, R. (2022) A Temporal-Difference Approach to Policy Gradient Estimation. *International Conference of Machine Learning (ICML)*.
- C9 Przystupa, M.; **Tosatto, S.**; Jagersand, M. (2022) Deep Probabilistic Movement Primitives *International Conference on Robot Learning (CoRL)*.
- C10 Karimi, A.; Jin, J.; Luo, J.; Mahmood, R.; **Tosatto, S.** (2022) Continuous-Time Continuous-Option Policy Gradient for Agile Robotic Manipulation. *International Conference on Robot Learning (CoRL)*.

THESES

- T1 **Tosatto, S.** (2020) Towards Off-Policy Reinforcement Learning for Robotics, *Ph.D. Thesis*
- T2 **Tosatto, S.** (2017) Bootstrapped Fitted Q-Iteration, *M.Sc. Thesis*

SUMMARY

I am a very passionate researcher, with publication at top tier venues in machine learning and robotics. Thanks to **my working attitude** and **team-work ability**, I will produce **high-quality research**, collaborating both inside your department, and outside with industries and other universities. I supervised students at all levels: bachelor, master and Ph.D., guiding them to successful thesis submission and defense. **I believe that I will be a great addition to your department.**

REFERENCES

Prof. Jan Peters, (Ph.D. Advisor)

TU Darmstadt and Max Planck Institute for Machine Intelligence, mail@jan-peters.net, +49 6151-16-25374

Prof. Martha White, (Coauthor)

University of Alberta and PI at Amii, mwhite@ualberta.ca

Prof. Georgia Chalvatzaki, (Coauthor)

TU Darmstadt, georgia.chalvatzaki@tu-darmstadt.de

Prof. Elmar Rückert, (Coauthor)

Chair of Cyber-Physical-Systems at Montanuniversität Leoben rueckert@unileoben.ac.at, +43 3842 402 1901

Prof. Joni Pajarinen, (Coauthor)

Aalto University, joni.pajarinen@aalto.fi, +35 850 3094771

Dr. Riad Akrou, (Coauthor)

Scool at INRIA, riad.akrou@inria.fr