Adil Salim

Machine Learning Research

asalim.math@gmail.com
adil-salim.github.io
adil-salim
AdilSlm
adil-salim

Employment

Feb 2024–... **Member of Technical Staff**, *Microsoft AI*, Redmond, USA.

Apr 2022–Jan 2024 **Senior Researcher**, *Microsoft Research*, Redmond, USA.
Machine Learning Foundations group.

Fall 2021 **Google Research Fellow**, *Simons Institute*, UC Berkeley, Berkeley, USA.
Program: Geometric Methods in Optimization and Sampling.

2019–2021 **Postdoctoral researcher**, *KAUST*, Visual Computing Center, Thuwal, KSA.
Host: Peter Richtárik.

2018 **Chief Technical Officer**, *Revna Sciences*, Paris, France, Consulting in R&D.

Fall 2015 **Freelance data scientist**, *ThePriceHub.com*, Paris, France, Priced industrial custom parts.

Summer 2015 **Freelance data scientist**, *Kwanko*, Bourg-la-Reine, France, Developed a Real Time Bidding algorithm.

Summer 2015 **Internship**, *CREST-ENSAE*, Malakoff, France, Free stochastic calculus and Large random matrices.
Supervisor: Philippe Biane

Summer 2014 **Internship**, *Insight-UCD*, Dublin, UK, Importance sampling, MCMC methods.
Supervisor: Florian Maire

Summer 2013 **Auditor**, *Deloitte*, Neuilly-sur-Seine, France, Audited banks and insurance companies.

Education

2015–2018 **Ph.D**, *Telecom Paris and Paris–Saclay University*, Paris, France, Research in optimization for machine learning.
Summa cum laude. Defended on November 26, 2018.
Title: Random monotone operators and application to stochastic optimization
Supervisors: Pascal Bianchi, Walid Hachem
Jury president: Antonin Chambolle
Referees: Jérôme Bolte, Bruno Gaujal
Examiner: Panayotis Mertikopoulos
Invited: Jérémie Jakubowicz, Eric Moulines

2014–2015 **M.Sc**, *Paris–Saclay University*, Orsay, France, Probability-Statistics.
Magna cum laude

2012–2015 **M.Sc**, *ENSAE Paris*, Malakoff, France, Data Science and Statistical Engineering.
Top 3 Master's thesis out of ∼200 students
Title: Free stochastic calculus: Processes with free increments
Supervisor: Philippe Biane
Jury: Alexandre Tsybakov

2009–2012 **Preparatory Classes for the "Grandes Ecoles"**, *Lycée Lakanal*, Sceaux, France, Mathematics/Physics.
Three-year undergraduate intensive course

2009 **Baccalauréat Scientifique**, *Lycée Mistral*, Fresnes, France, Spe. Mathematics.
Summa cum laude

Awards and distinctions

2024 **KAUST Rising star in AI**.
2023 **ICLR notable top 5% paper**.
2023 **ALT best student paper award for paper co-authored with MSR intern**.
2022 **SIAM Early Career Travel Award**.
2022 **ICML 2022 free registration**.
2022 **ICML 2022 Tutorial**.
2021 **Google Research Fellowship, Simons Institute, UC Berkeley**.
2020 **Top 33% ICML reviewer**.
2019 **NeurIPS spotlight paper**.
2019 **Top 50% NeurIPS reviewer**.
2018 **GDR ISIS Travel Grant for PhD mobility**, *EPFL, Lausanne, Switzerland*.
2018 **NeurIPS Workshop Black in AI Travel Grant**, *Montreal, Canada*.
2017 **NIPS Workshop Black in AI Travel Grant**, *Los Angeles, USA*.
2015 **Top 3 Master's thesis out of nearly 200 students**, *ENSAE Paris, Malakoff, France*.
2009–2012 **Merit Scholarship**, *Lycée Lakanal, Sceaux, France*.

Publications

[1] Victor Priser, Pascal Bianchi, Adil Salim, Long-time asymptotics of noisy SVGD outside the population limit *arXiv preprint arXiv:2406.11929*, June 2024.

[2] Marah Abdin *et al.*, Phi-3 Technical report: A Highly Capable Language Model Locally on Your Phone *arXiv preprint arXiv:2404.14219*, April 2024.

[3] Adil Salim. A Strong Law of Large Numbers for Random Monotone Operators. *Set-Valued and Variational Analysis*, 2023.

[4] Suriya Gunasekar, Yi Zhang *et al.*, Textbooks Are All You Need. *arXiv preprint arXiv:2306.11644*, June 2023.

[5] Krishnakumar Balasubramanian, Larry Goldstein, Nathan Ross, Adil Salim, Gaussian random field approximation via Stein's method with applications to wide random neural networks. *arXiv preprint arXiv:2306.16308*, June 2023.

[6] Sitan Chen, Sinho Chewi, Holden Lee, Yuanzhi Li, Jianfeng Lu, Adil Salim, The probability flow ODE is provably fast. *NeurIPS 2023*.

[7] Michael Ziyang Diao, Krishna Balasubramanian, Sinho Chewi, Adil Salim, Forward-backward Gaussian variational inference via JKO in the Bures-Wasserstein Space. *ICML 2023*.

[8] Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, Anru R. Zhang, Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. *ICLR 2023,* ***Top 5% paper***.

[9] Sinho Chewi, Sébastien Bubeck, Adil Salim, On the complexity of finding stationary points of smooth functions in one dimension. *ALT 2023,* ***Best student paper award***.

[10] Lukang Sun, Adil Salim, Peter Richtárik, Federated Sampling with Langevin Algorithm under Isoperimetry. *TMLR 2023*.

[11] Krishnakumar Balasubramanian, Sinho Chewi, Murat A. Erdogdu, Adil Salim, Matthew Zhang, Towards a Theory of Non-Log-Concave Sampling: First-Order Stationarity Guarantees for Langevin Monte Carlo. *COLT 2022*.

[12] Yongxin Chen, Sinho Chewi, Adil Salim, Andre Wibisono, Improved analysis for a proximal algorithm for sampling. *COLT 2022*.

[13] Adil Salim, Lukang Sun, Peter Richtárik. Complexity Analysis of Stein Variational Gradient Descent Under Talagrand's Inequality T1. *ICML 2022*.

[14] Adil Salim, Laurent Condat, Dmitry Kovalev and Peter Richtárik. An Optimal Algorithm for Strongly Convex Minimization under Affine Constraints. *AISTATS 2022*.

[15] Adil Salim, Laurent Condat, Konstantin Mishchenko and Peter Richtárik. Dualize, Split, Randomize: Fast Nonsmooth Optimization Algorithms. *Journal of Optimization Theory and Applications*, 2022

[16] Adil Salim and Peter Richtárik. Primal Dual Interpretation of the Proximal Stochastic Gradient Langevin Algorithm. *NeurIPS 2020*.

[17] Adil Salim, Anna Korba and Giulia Luise. The Wasserstein Proximal Gradient Algorithm. *NeurIPS 2020*.

[18] Dmitry Kovalev, Adil Salim and Peter Richtárik. Optimal and Practical Algorithms for Smooth and Strongly Convex Decentralized Optimization. *NeurIPS 2020*.

[19] Anna Korba, Adil Salim, Michael Arbel, Giulia Luise and Arthur Gretton. A Non-Asymptotic Analysis for Stein Variational Gradient Descent. *NeurIPS 2020*.

[20] Sélim Chraibi, Ahmed Khaled, Dmitry Kovalev, Adil Salim, Peter Richtárik and Martin Takáč. Distributed Fixed Point Methods with Compressed Iterates. *arXiv preprint arXiv:1912.09925*, December 2019.

[21] Sélim Chraibi, Adil Salim, Samuel Horváth, Filip Hanzely and Peter Richtárik. Learning To Optimize Via Dual Space Preconditioning. *Technical Report*, September 2019.

[22] Adil Salim, Dmitry Kovalev and Peter Richtárik. Stochastic Proximal Langevin Algorithm: Potential Splitting and Nonasymptotic Rates. *NeurIPS 2019,* ***Spotlight***.

[23] Michael Arbel, Anna Korba, Adil Salim and Arthur Gretton. Maximum Mean Discrepancy Gradient Flow. *NeurIPS 2019*.

[24] Anna Korba, Adil Salim, Michael Arbel and Arthur Gretton. Yet another look at Stein Variational Gradient Descent. *ICML 2019 Workshop on Stein's Method*.

[25] Pascal Bianchi, Walid Hachem and Adil Salim. A Fully Stochastic Primal-Dual Algorithm. *Optimization Letters*, June 2020.

[26] Adil Salim and Walid Hachem. On the Performance of the Stochastic FISTA. *Technical Report*, March 2019.

[27] Adil Salim, Pascal Bianchi and Walid Hachem. Snake: a Stochastic Proximal Gradient Algorithm for Regularized Problems over Large Graphs. *IEEE Transaction on Automatic Control*, May 2019.

[28] Pascal Bianchi, Walid Hachem and Adil Salim. A constant step Forward-Backward algorithm involving random maximal monotone operators. *Journal of Convex Analysis*, May 2019.

[29] Pascal Bianchi, Walid Hachem and Adil Salim. Constant step stochastic approximations involving differential inclusions: Stability, long-run convergence and applications. *Stochastics*, 2018.

[30] Sholom Schechtman, Adil Salim and Pascal Bianchi. Passty Langevin. *CAp 2019*.

[31] Adil Salim, Pascal Bianchi and Walid Hachem. A Constant Step Stochastic Douglas-Rachford Algorithm with Application to Non Separable Regularization. *IEEE ICASSP 2018*.

[32] Adil Salim, Pascal Bianchi and Walid Hachem. A Stochastic Proximal Point Algorithm for Total Variation Regularization over Large Scale Graphs. *IEEE CDC 2016*.

[33] Adil Salim, Pascal Bianchi and Walid Hachem. Snake: a Stochastic Proximal Gradient Algorithm for Regularized Problems over Large Graphs. *NIPS 2017 Workshop Black in AI*.

[34] Rahul Mourya, Pascal Bianchi, Adil Salim and Cédric Richard. An adaptive Distributed Asynchronous Algorithm with Application to Target Localization. *IEEE CAMSAP 2017*.

[35] Pascal Bianchi, Walid Hachem and Adil Salim. Convergence d'un algorithme du gradient proximal stochastique à pas constant et généralisation aux opérateurs monotones aléatoires. *GRETSI 2017*.

[36] Adil Salim, Pascal Bianchi and Walid Hachem. Snake: a Stochastic Proximal Gradient Algorithm for Regularized Problems over Large Graphs. *CAp 2017*.

Professional activities

Reviewing

2019–... **Area chair**, *Black in AI Neurips workshop 2019, 2020, 2021, 2022*.

2017–... **Reviewer**, *Conferences: NeurIPS 2019, 2020, 2021, ICML 2020, 2021, 2022, ICLR 2021, 2022, 2023 ACM SIGKDD 2021, COLT 2022, Journals: Nature, Journal of Machine Learning Research, Transactions on Machine Learning Research, Annals of statistics, Journal of the Royal Statistical Society: Series B, Mathematical Programming, Set-Valued and Variational Analysis, Applied Mathematics and Optimization, IEEE Transactions on Information Theory, SIAM Journal on Imaging Sciences, IEEE Transactions on Signal and Information Processing over Networks, IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, Automatica, Numerical Algorithms, Journal of Mathematical Analysis and Applications, Journal of Scientific Computing, Rendiconti del Circolo Matematico di Palermo, Mathematics of Operations Research, Journal of Optimization Theory and Applications, IMA Journal of Numerical Analysis, Foundations of Computational Mathematics, Black in AI Neurips workshop 2018, 2023, BlackAIR 2021*.

Organization

2024 **Workshop chair**, *NeurIPS*, Vancouver, Canada.

2023 **Co-organizer of a mini symposium**, *SIAM Conference on Optimization 2023*, Seattle, USA, Title: Wasserstein Gradient Flows and Applications (three sessions).

2020 **Co-organizer of a mini symposium [Cancelled due to Covid-19 outbreak]**, *SIAM Conference on Optimization 2020*, Hong-Kong, Title: Gradient flows and interactions between optimization and simulation algorithms (two sessions).

2020 **Co-organizer of a mini symposium [Cancelled due to Covid-19 outbreak]**, *SIAM Conference on Optimization 2020*, Hong-Kong, Title: Recent advances in primal-dual splitting for convex optimization.

Selected talks

September 2023 **Optimal methods for decentralized optimization**, *Seminar*, John Hopkins University, Baltimore, USA.

May 2023 **Theory for diffusion models**, *ICLR*, Kigali, Rwanda.

April 2023 **A proximal method for sampling**, *Seminar*, UC Santa Barbara, USA.

March 2023 **Tutorial on optimization and sampling**, *Seminar*, University of Ottawa, Ottawa. Canada

February 2023 **Lower bounds for finding stationary points**, *ALT*, NUS, Singapore.

January 2023 **A proximal method for sampling**, *Geometric methods in optimization and sampling reunion*, Simons Institute, UC Berkeley, USA.

September 2022 **A proximal method for sampling**, *3rd Symposium on Machine Learning and Dynamical Systems*, Fields Institute, Toronto, Canada.

September 2022 **The Restricted Gaussian Oracle as a proximity operator for sampling**, *SIAM Conference on Mathematics of Data Science (MDS22)*, San Diego, USA.

July 2022 **Convergence theory for SVGD**, *ICML*, Baltimore, USA.

July 2022 **Sampling as First-Order Optimization over a space of probability measures (with Anna Korba)**, *ICML Tutorial*, Baltimore, USA.

June 2022 **Wasserstein Gradient Descent algorithms**, *Stein's Method: The Golden Anniversary*, IMS Auditorium, Singapore.

May 2022 **Stein Variational Gradient Descent**, *Seminar*, MSR New England, USA.

April 2022 **Tutorial on Stein Variational Gradient Descent**, *Advances in Stein's method and its applications in Machine Learning and Optimization*, Banff International Research Station, Canada.

March 2022 **Sampling as an optimization task**, *Seminar*, Google, Mountain View, USA.

October 2021 **Primal–Dual Davis–Yin algorithm(s)**, *INFORMS annual meeting*, Los Angeles, USA.

September 2021 **Primal–Dual interpretation of the proximal gradient Langevin algorithm**, *Workshop on Sampling Algorithms and Geometries on Probability Distributions*, Simons Institute, UC Berkeley, USA.

July 2021 **An Optimal Algorithm for Strongly Convex Minimization under Affine Constraints**, *EurOPT 2021*, Toulouse (Online), France.

June 2021 **An Optimal Algorithm for Strongly Convex Minimization under Affine Constraints**, *Research Seminar*, University of Genova (Online), Italy.

March 2021 **Primal–dual optimization and application to decentralized optimization**, *Machine Learning and Data Analytics Symposium (MLDAS) 2021*, Seattle (Online), USA.

March 2021 **Sampling as an Optimization task**, *Research Seminar*, Qatar Computing Research Institute (Online), Qatar.

March 2021 **Sampling as an Optimization task**, *Research Seminar*, MBZUAI, Abu Dhabi (Online), UAE.

March 2021 **From Dynamics to Complexity of Machine Learning algorithms**, *Research Seminar*, KAUST (Online), KSA.

February 2021 **Primal–dual optimization and applications to decentralized optimization and sampling**, *Research Seminar*, Microsoft Research, Redmond (Online), USA.

September 2020 **Primal Dual Interpretation of the Proximal Stochastic Gradient Langevin Algorithm**, *Second Symposium on Machine Learning and Dynamical Systems*, Fields Institute, Toronto (Online), Canada.

December 2019 **Stochastic Proximal Langevin Algorithm**, *NeurIPS*, Vancouver, Canada.

November 2019 **Langevin as an Optimization algorithm**, *Computer Science Graduate Seminar*, KAUST, KSA.

August 2019 **On Stochastic Primal–Dual Algorithms**, *International Conference on Continuous Optimization (ICCOPT)*, TU Berlin, Germany.

May 2019 **Exponential Convergence Time of Gradient Descent for One-Dimensional Deep Linear Neural Networks**, *Mathematics of Deep Learning seminar*, KAUST, KSA.

April 2019 **Stochastic Chambolle-Pock**, *Visual Computing Center showcase*, KAUST, KSA.

March 2019 **Sampling as Convex Optimization**, *Guest Lecture on Optimization for Machine Learning*, KAUST, KSA.

July 2018 **A Splitting Algorithm for Minimization under Stochastic Linear Constraints**, *International Symposium on Mathematical Programming (ISMP)*, Bordeaux, France.

March 2018 **A stochastic Forward Backward algorithm with application to large graphs regularization**, *Machine Learning and Optimization seminar*, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

February 2018 **Snake: a Stochastic Proximal Gradient Algorithm for Regularized Problems over Large Graphs**, *GdR Information, Signal, Image et ViSion (ISIS) meeting*, Telecom Paris, France.

November 2017 **Distributed Douglas Rachford algorithm**, *ANR ODISSEE meeting*, Nice Sophia Antipolis University, France.

September 2017 **Stochastic Proximal Gradient algorithm**, *France / Japan Machine Learning Workshop*, Ecole Normale Supérieure Paris, France.

June 2017 **Snake**, *Conférence sur l'Apprentissage automatique (CAp)*, IMAG Grenoble, France.

December 2016 **A Stochastic Proximal Point Algorithm for Total Variation Regularization over Large Scale Graphs**, *Conference on Decision and Control (CDC)*, Las Vegas, USA.

Teaching and Supervision

2024 **Supervision of Khashayar Gatmiry's (PhD student at MIT with Jonathan Kelner) internship**, *Microsoft Research*, Redmond, USA.

2024 **Co-supervision of Michał Grudzień master's thesis**, *Oxford University*, Oxford, UK.

2023 **Co-supervision of Jaume De Dios Pont's (PhD student at UCLA with Terence Tao) internship**, *Microsoft Research*, Redmond, USA.

2022 **Co-supervision of Sinho Chewi's (PhD student at MIT with Philippe Rigollet) internship**, *Microsoft Research*, Redmond, USA.

2019 **Teaching assistant**, *KAUST*, Thuwal, KSA, 16h.
Optimization for Machine Learning

2019 **Co-supervision of Sélim Chraibi's master's thesis**, *KAUST*, Thuwal, Saudi Arabia.
Title: First order optimization algorithms and compressed optimization algorithms for Federated Learning

2018 **Co-supervision of Sholom Schechtman's master's thesis**, *Telecom Paris*, Paris, France.
Title: Passty Langevin algorithm

2017–2018 **Teaching assistant**, *Telecom Paris*, Paris, France, 32h.
Optimization for Machine Learning. Supervised an Image processing project for industrial master's students

2015–2017 **Teaching assistant**, *ENSAE Paris*, Malakoff, France, 128h.
Algebra, Measure Theory, Introduction to Stochastic Processes, Supervision of Applied statistics project: French presidential elections, Twitter graph and Markov chains

2013–2016 **Volunteer Professor and Member of the Teaching Pole**, *Various nonprofit organizations*, Paris, France, Coached ambitious students from disadvantaged backgrounds to prepare the competitive exams for prestigious French universities. $\sim$ 4 hours per week.
Organizations : Connex'cités, Gics

2005–2008 **Private teacher**, *Private lessons in mathematics and physics*, Paris, France, $\sim$ 1 hour per week.

References

Jérôme Bolte, *Professor*, Toulouse 1 Capitole University, France, jerome.bolte@tse-fr.eu.

Antonin Chambolle, *Research director*, CNRS and Paris Dauphine University, France, chambolle@ceremade.dauphine.fr.

Laurent Condat, *Research scientist*, KAUST, Saudi Arabia, laurent.condat@kaust.edu.sa.

Walid Hachem, *Research director*, CNRS and Gustave Eiffel University, France, walid.hachem@univ-eiffel.fr.

Peter Richtárik, *Professor*, KAUST, Saudi Arabia, peter.richtarik@kaust.edu.sa.