
CONTACT INFORMATION	DIRO University of Montreal 2920 chemin de la Tour Montreal (QC), Canada	Email: see website Web: http://www.di.ens.fr/~slacoste Nationality: Canadian
CURRENT POSITION	Assistant Professor University of Montreal , Montreal, Canada Department of Computer Science and Operational Research (DIRO)	Sept 2016 – present
RESEARCH INTERESTS	Machine learning, optimization & statistics: structured prediction, graphical models. Applications: Computer vision, NLP, information retrieval & computational biology.	
EDUCATION	University of California, Berkeley , Berkeley, CA, USA Ph.D. in Computer Science <i>Designated Emphasis in Communication, Computation and Statistics</i> Dissertation title: Discriminative Machine Learning with Structure Committee: Michael I. Jordan (chair), Peter L. Bartlett, Peter J. Bickel, Dan Klein	2003 – 2009
	McGill University , Montréal, QC, Canada B.Sc. Triple Honours in Mathematics, Physics and Computer Science <i>First Class Honours in Math. & Physics, First Class Honours in Math. & CS</i> <i>Anne Molson Gold Medal</i> (best in math), <i>Dean's Honour List</i> , GPA 3.96/4.00	1999 – 2003
SELECTED AWARDS AND DISTINCTIONS	Google Focused Research Award Google Faculty Research Award MCMCSki IV Honorable Mention Poster Prize NIPS Outstanding Reviewer Award Research in Paris Fellowship, City of Paris Wolfson College Junior Research Fellowship, University of Cambridge UC Berkeley College of Engineering Graduate Student Prize <i>For outstanding scholarship and achievement at the College of Engineering</i> NSERC Postgraduate Scholarship (Canadian NSF) NATEQ Scholarship (Québec NSF) <i>Ranked first in the Mathematical Sciences competition</i> McGill Moyse Travelling Scholarship <i>Only one awarded in the Faculty of Science of McGill University</i> Berkeley, Cornell, MIT and McGill Graduate Fellowships	2016 2015 2014 2013, 2015 2011–2012 2009–2011 2008 2005–2008 2003–2005 2003 2003
ACADEMIC AND RESEARCH POSITIONS	École Normale Supérieure & INRIA , Paris, France Computer Science Department <i>Research Scientist</i> in SIERRA project-team	Sept 2013 – Aug 2016

École Normale Supérieure & INRIA, Paris, France Sept 2011 – Aug 2013
Computer Science Department
Research in Paris Fellow in SIERRA project-team

University of Cambridge, Cambridge, UK Nov 2008 – Aug 2011
Department of Engineering
Research Associate in Machine Learning Group

Yahoo! Research, Santa Clara, CA, USA May 2006 – August 2006
Research Intern

TEACHING

University of Montreal:
advanced structured prediction and optimization (IFT 6085) (grad class, 2017)
probabilistic graphical model (IFT 6269) (grad class, 2016 and 2017)

École Normale Supérieure:
statistical machine learning (1st year master, 2014 and 2015)
programming projects for machine learning (2nd year master, 2014 and 2015)

École Normale Supérieure de Cachan:
probabilistic graphical models (2nd year master, 2015)

University of California, Berkeley:
practical machine learning (grad class, 2006 and 2008)

PROFESSIONAL ACTIVITIES

Area chair for ICML 2015–2018, NIPS 2016–2017, AISTATS 2018, CVPR 2018.

Sponsorship chair for AISTATS 2018.

Workshop co-organizer for *The Generative and Discriminative Learning Interface* international workshop, NIPS 2009.

Session chair for the International Conference on Continuous Optimization, 2013.

Peer Reviewer for major journals and conferences in:

- machine learning (JMLR, TPAMI, Machine Learning, TKDE;
NIPS, ICML, COLT, UAI, AISTATS, IJCAI, EMNLP)
- statistics (Annals of Statistics, JRSS-B, Statistics and Computing)
- optimization (Mathematical Programming, SIOPT)
- computer vision (IJCV; CVPR, ECCV).

SUPERVISION

Since taking a faculty position at INRIA in September 2013, I have supervised 7 PhD students (1 graduated), 1 Master intern and 1 post-doc:

- Rémi Le Priol, (PhD student), *Modern accelerated optimization methods for structured prediction*, since 2017.
- Aristide Baratin, (PhD student), *Generalization in Structured Prediction and Deep Learning*, since 2017.
- Gabriel Huang, (PhD student), *Meaningful unsupervised learning via the adversarial framework*, since 2017.
- Gauthier Gidel, (PhD student), *Frank-Wolfe Algorithms for Saddle Point Problems and Beyond*, since 2016.
- Rémi Leblond, (PhD student), *Distributed and Streaming Versions of Fast Incremental Gradient Methods for Big Structured Data*, since 2015.

- Jean-Baptiste Alayrac, (PhD student), *Structured Learning from Video and Natural Language*, (co-advised with Josef Sivic and Ivan Laptev) since 2014.
- Anastasia Podosinnikova, (PhD student), *On the method of moments for estimation in latent variable models*, (co-advised with Francis Bach), 2013–2016. (now a post-doc at MIT)
- Anton Osokin, (Postdoc), working on structured prediction and computer vision, 2014–2016. (now an Assistant Professor at CS HSE, Moscow, Russia)
- Isabella Lukasewitz, (master internship for Polytechnique), *Block-Coordinate Frank-Wolfe Optimization*, summer 2013. (now a consultant at McKinsey & Company)

INVITED TALKS

I gave more than 40 invited talks at universities around the world, at workshops and conferences. Here are some examples:

- “SAGA: a Fast Incremental Gradient Method”, invited talk in the Nonsmooth Optimization cluster, 22nd International Symposium on Mathematical Programming (ISMP), Pittsburgh, USA, 2015
- “Sequential Kernel Herding: Frank-Wolfe Optimization for Particle Filtering”, Reasoning and Learning Lab seminar, McGill University, Montreal, Canada, 2014
- “Frank-Wolfe Optimization Insights in Machine Learning”, Department of Computer Science, Tsinghua University, Beijing, China, 2014
- “Recent Advances in Frank-Wolfe Optimization”, invited talk, 4th IMA Conference on Numerical Linear Algebra and Optimisation, Birmingham, UK, 2014
- “Harnessing the Structure of Data for Discriminative Machine Learning”, Department of Statistics, University of Oxford, Oxford, UK, 2013
- “Block-Coordinate Frank-Wolfe Optimization for Structured SVMs”, invited talk, The Fourth International Conference on Continuous Optimization (ICCOPT), Lisbon, Portugal, 2013
- “Frank-Wolfe Optimization Insights in Machine Learning”, Artificial Intelligence Laboratory, Stanford University, Palo Alto, USA, 2012
- “Approximate Inference for the Loss-Calibrated Bayesian”, Statistical Laboratory Seminar, University of Cambridge, Cambridge, UK, 2011

PEER-REVIEWED PUBLICATIONS

(Available at: <http://scholar.google.com/citations?user=oejm5IUAAAAJ>)
26 publications in top peer-reviewed international conferences and journals; about 1800 citations; h-index: 17 (according to Google Scholar).

1. A. Osokin, F. Bach and **S. Lacoste-Julien**. On Structured Prediction Theory with Calibrated Convex Surrogate Losses. In *Advances in Neural Information Processing Systems 30 (NIPS)*, 2017. (oral) [1 citation]
2. F. Pedregosa, R. Leblond and **S. Lacoste-Julien**. Breaking the Nonsmooth Barrier: a Scalable Parallel Method for Composite Optimization. In *Advances in Neural Information Processing Systems 30 (NIPS)*, 2017. (spotlight)
3. J.-B. Alayrac, J. Sivic, I. Laptev and **S. Lacoste-Julien**. Joint Discovery of Object States and Manipulating Actions. In *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, 2017. [2 citations]
4. D. Arpit, S. Jastrzebski, N. Ballas, D. Krueger, E. Bengio, M. S. Kanwal, T. Maharaj, A. Fischer, A. Courville, Y. Bengio and **S. Lacoste-Julien**. A Closer Look at Memorization in Deep Networks,. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 2017. [4 citations]
5. G. Gidel, T. Jebara and **S. Lacoste-Julien**. Frank-Wolfe Algorithms for Saddle Point Problem. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017. [1 citation]

6. R. Leblond, F. Pedregosa and **S. Lacoste-Julien**. ASAGA: Asynchronous Parallel Saga. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017. [15 citations]
7. P. Germain, F. Bach and **S. Lacoste-Julien**. PAC-Bayesian Theory Meets Bayesian Inference. In *Advances in Neural Information Processing Systems 29 (NIPS)*, 2016. [2 citations]
8. J.-B. Alayrac, P. Bojanowski, N. Agrawal, J. Sivic, I. Laptev and **S. Lacoste-Julien**. Unsupervised Learning from Narrated Instruction Videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. (oral) [25 citations]
9. A. Osokin, J.-B. Alayrac, I. Lukasewitz, P. Dokania and **S. Lacoste-Julien**. Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVMs. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 2016. [15 citations]
10. A. Podosinnikova, F. Bach and **S. Lacoste-Julien**. Beyond CCA: Moment Matching for Multi-View Models. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 2016. [6 citations]
11. **S. Lacoste-Julien** and M. Jaggi. On the Global Linear Convergence of Frank-Wolfe Optimization Variants. In *Advances in Neural Information Processing Systems 28 (NIPS)*, 2015. [70 citations]
12. R. Krishnan, **S. Lacoste-Julien** and D. Sontag. Barrier Frank-Wolfe for Marginal Inference. In *Advances in Neural Information Processing Systems 28 (NIPS)*, 2015. [12 citation]
13. T. Hofmann, A. Lucchi, **S. Lacoste-Julien**, and Brian McWilliams. Variance Reduced Stochastic Gradient Descent with Neighbors. In *Advances in Neural Information Processing Systems 28 (NIPS)*, 2015. [16 citation]
14. A. Podosinnikova, F. Bach and **S. Lacoste-Julien**. Rethinking LDA: Moment Matching for Discrete ICA. In *Advances in Neural Information Processing Systems 28 (NIPS)*, 2015. [7 citation]
15. **S. Lacoste-Julien**, F. Lindsten and F. Bach. Sequential Kernel Herding: Frank-Wolfe Optimization for Particle Filtering. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015. [18 citat.]
16. V. Chari, **S. Lacoste-Julien**, I. Laptev and J. Sivic. On Pairwise Costs for Network Flow Multi-Object Tracking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. [45 citations]
17. A. Defazio, F. Bach and **S. Lacoste-Julien**. SAGA: A Fast Incremental Gradient Method with Support for Non-Strongly Convex Composite Objectives. In *Advances in Neural Information Processing Systems 27 (NIPS)*, 2014. [301 citations]
18. **S. Lacoste-Julien**, M. Jaggi, M. Schmidt and P. Pletscher. Block-Coordinate Frank-Wolfe Optimization for Structural SVMs. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, 2013. [193 citations]
19. **S. Lacoste-Julien**, K. Palla, A. Davies, G. Kasneci, T. Graepel and Z. Ghahramani. SiGMa: Simple Greedy Matching for Aligning Large Knowledge Bases. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD)*, 2013. [52 citations]
20. F. Bach, **S. Lacoste-Julien** and G. Obozinski. On the Equivalence between Herding and Conditional Gradient Algorithms. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 2012. [54 citations]

21. **S. Lacoste-Julien**, F. Huszár and Z. Ghahramani. Approximate Inference for the Loss-Calibrated Bayesian. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011. [17 citations]
22. **S. Lacoste-Julien**, F. Sha and M. Jordan. DiscLDA: Discriminative Learning for Dimensionality Reduction and Classification. In *Advances in Neural Information Processing Systems 21 (NIPS)*, 2008. [348 citations]
23. **S. Lacoste-Julien**, B. Taskar, D. Klein and M. Jordan. Word Alignment via Quadratic Assignment. In *Proceedings of the North American Chapter of the Association for Computational Linguistics Annual Meeting (HLT-NAACL)*, 2006. [85 citations]
24. B. Taskar, **S. Lacoste-Julien** and M. Jordan. Structured Prediction, Dual Extragradient and Bregman Projections. *Journal of Machine Learning Research (JMLR)*, 7:1627-1653, 2006. [120 citations]
25. B. Taskar, **S. Lacoste-Julien** and M. Jordan. Structured Prediction via the Extragradient Method. In *Advances in Neural Information Processing Systems 18 (NIPS)*, 2005. [61 citations]
26. B. Taskar, **S. Lacoste-Julien** and D. Klein. A Discriminative Matching Approach to Word Alignment. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2005. [217 citations]