Abbas Mehrabian

MONTRÉAL QC CANADA http://cs.mcgill.ca/~amehra13/

APPOINTMENT
- **McGill University**: September 2017 – present.
  CRM-ISM Postdoctoral Fellow (supervised by Luc Devroye and Louigi Addario-Berry).
- **University of California, Berkeley**: January 2017 – May 2017.
  Simons-Berkeley Research Fellow at the Simons Institute for the Theory of Computing.
- University of British Columbia and Simon Fraser University: June 2015–December 2016.
  PIMS-NSERC Postdoctoral Fellow (supervised by Nick Harvey in UBC and Petra Berenbrink in SFU).

EDUCATION
- **University of Waterloo**, Canada: May 2011 – April 2015.
  PhD in Combinatorics and Optimization (supervised by Nick Wormald and Joseph Cheriyan).
  During the PhD, visited Monash University in Australia for nine months and worked with Nick Wormald.
  Won the Governor General's Gold Medal (best PhD thesis of the year at UWaterloo).
- University of Waterloo, Canada: September 2009 – April 2011.
  Master of Mathematics in Combinatorics and Optimization under supervision of Nick Wormald.
- Sharif University of Technology, Tehran, Iran: September 2004 – August 2009.
  BSc in Computer Engineering and in Mathematics (two majors).

AWARDS
- **Governor General's Gold Medal for the PhD thesis, 2015.**
  Awarded to the student who achieves the highest academic standing at the doctoral level. Only one UWaterloo PhD recipient received this medal in 2015.
- **Simons-Berkeley Research Fellowship**, 2017
  *Awarded by the Simons Institute for the Theory of Computing at UC Berkeley, where I participated in the programs **Foundations of Machine Learning**, and Pseudorandomness (Spring 2017).*
- **CRM-ISM Postdoctoral Fellowship**, 2017–2019
- **NSERC Postdoctoral Fellowship**, 2015–2017
- **Vanier Canada Graduate Scholarship**, 2013–2015
- PIMS Postdoctoral Fellowship, 2015
- Ontario Graduate Scholarship, 2012–2013
- University of Waterloo President's Graduate Scholarship, 2012–2013
- Gold Medal in Iranian National Mathematics Olympiad, 2003.

RESEARCH INTERESTS
- Machine Learning Theory
- Mathematical Statistics
- Randomized algorithms and probabilistic analysis of algorithms
- Random graphs and stochastic processes on graphs
- In the past I have worked in diverse areas such as probability theory, random matrices, random graphs, and graph theory; my focus has been on machine learning theory for the past two years. For a detailed description of problems I have worked on, see http://cs.mcgill.ca/~amehra13/research.htm. For a list of publications, see below.

SELECTED PUBLICATIONS

◇ **Nearly-tight VC-dimension bounds for piecewise linear neural networks**, with Nick Harvey and Chris Liaw, in Conference on Learning Theory (COLT'17).
Classification is one of the learning tasks in which deep neural networks have been particularly successful, e.g., for image recognition. A natural foundational question that arises is: what are the theoretical limits on the classification power of these networks? The established way to formalize this question is by considering the VC-dimension, as it is well known that this asymptotically determines the sample complexity of PAC learning with such classifiers. We proved new upper and lower bounds on the VC-dimension of deep neural networks with the ReLU activation function. These bounds are tight for almost the entire range of parameters. Letting W be the number of parameters and L be the number of layers, we proved that the VC-dimension is $O(WL\log(W))$ and $\Omega(WL\log(W/L))$. This improves both the previously known upper bounds and lower bounds. These bounds generalize to arbitrary piecewise linear activation functions, and also hold for the pseudodimensions of these networks.

◇ **Sample-Efficient Learning of Mixtures**, with Hassan Ashtiani and Shai Ben-David, accepted for presentation in the AAAI Conference on Artificial Intelligence (AAAI'18), preprint available on arXiv.
In PAC learning of probability distributions (also known as density estimation), we are given an i.i.d. sample generated from an unknown target distribution, and want to output a distribution that is close to the target in total variation distance. Let $\mathcal{F}$ be an arbitrary class of probability distributions, and let $\mathcal{F}^k$ denote the class of k-mixtures of elements of $\mathcal{F}$. Assuming the existence of a method for learning $\mathcal{F}$ with sample complexity $m_{\mathcal{F}}(\epsilon)$, we provided a method for learning $\mathcal{F}^k$ with sample complexity $O(k\log k\cdot m_{\mathcal{F}}(\varepsilon)/\varepsilon^2)$. Our mixture learning algorithm has the property that, if the $\mathcal{F}$-learner is proper and agnostic, then the $\mathcal{F}^k$-learner would be proper and agnostic as well. This general result enables us to improve the best known sample complexity upper bounds for the class of mixtures of k Gaussians in $\mathbb{R}^d$, from $\min\left\{\widetilde{O}(k^3d^2/\varepsilon^4), \widetilde{O}(k^4d^4/\varepsilon^2)\right\}$ to $\widetilde{O}(kd^2/\varepsilon^4)$.

◇ **Agnostic Distribution Learning via Compression**, with Hassan Ashtiani and Shai Ben-David, submitted, preprint available on arXiv.
We introduced a novel method for distribution learning via a form of compression. Having a large enough sample from a target distribution, can one compress that sample set, by picking only a few instances from it, in a way that allows recovery of (an approximation to) the target distribution from the compressed set? We prove that if this is the case for all members of a class of distributions, then there is a sample-efficient way of distribution learning for this class. As an application of our approach, we provide a sample-efficient method for agnostic distribution learning with respect to the class of mixtures of k axis-aligned Gaussian distributions over $\mathbb{R}^d$. This method uses only $\widetilde{O}(kd/\varepsilon^2)$ samples (to guarantee with high probability an error of at most ε). This improves the previously known bounds, and is tight up to logarithmic factors.

JOURNAL PUBLICATIONS

◇ Christopher Liaw, Abbas Mehrabian, Yaniv Plan, and Roman Vershynin. A simple tool for bounding the deviation of random matrices on geometric sets. In *Geometric Aspects of Functional Analysis*, volume 2169 of *Lecture Notes in Math.*, pages 277–299. Springer, Cham, 2017.

◇ Jeannette Janssen and Abbas Mehrabian. Rumors spread slowly in a small-world spatial network. *SIAM Journal on Discrete Mathematics*, 31(4):2414–2428, 2017. Extended abstract in WAW 2015.

◇ Abbas Mehrabian. Justifying the small-world phenomenon via random recursive trees. *Random Structures & Algorithms*, 50(2):201–224, 2017.

◇ Hüseyin Acan, Andrea Collevecchio, Abbas Mehrabian, and Nick Wormald. On the push&pull protocol for rumor spreading. *SIAM J. Discrete Math.*, 31(2):647–668, 2017. Extended abstract in PODC'15.

◇ Andrea Collevecchio, Abbas Mehrabian, and Nick Wormald. Longest paths in random Apollonian networks and largest r-ary subtrees of random d-ary recursive trees. *J. Appl. Probab.*, 53(3):846–856, 2016.

◇ Abbas Mehrabian and N. Wormald. It's a small world for random surfers. *Algorithmica*, 76(2):344–380, 2016. Extended abstract in APPROX/RANDOM 2014.

◇ A. Mehrabian and A. Pourmiri. Randomized rumor spreading in poorly connected small-world networks. *Random Structures & Algorithms*, 49(1):185–208, 2016. (Conference version in DISC 2014).

- N. Alon and A. Mehrabian. Chasing a fast robber on planar graphs and random graphs. *Journal of Graph Theory*, 78(2):81–96, 2015.
- S. Ehsani, S. ShokatFadaee, M. Fazli, A. Mehrabian, S. Sadeghabad, M. Safari, and M. Saghafian. A bounded budget network creation game. *ACM Transactions on Algorithms*, 11(4):34:1–34:25, 2015. (Conference version in SPAA 2011).
- A. Mehrabian. The fast robber on interval and chordal graphs. *Discrete Applied Mathematics*, 180:188 – 193, 2015.
- S. Akbari, A. Daemi, O. Hatami, A. Javanmard, and A. Mehrabian. Nowhere-zero unoriented flows in Hamiltonian graphs. *Ars Combinatoria*, CXX:51–63, 2015.
- E. Ebrahimzadeh, L. Farczadi, P. Gao, A. Mehrabian, C. M. Sato, N. Wormald, and J. Zung. On longest paths and diameter in random Apollonian networks. *Random Structures & Algorithms*, 45(4):703–725, 2014.
- A. Mehrabian, D. Mitsche, and P. Prałat. On the maximum density of graphs with unique-path labelings. *SIAM Journal on Discrete Mathematics*, 27(3):1228–1233, 2013.
- A. Mehrabian and N. Wormald. On the stretch factor of randomly embedded random graphs. *Discrete & Computational Geometry*, 49(3):647–658, 2013.
- A. Mehrabian. Cops and robber game with a fast robber on expander graphs and random graphs. *Annals of Combinatorics*, 16(4):829–846, 2012.
- A. Mehrabian. On the density of nearly regular graphs with a good edge-labeling. *SIAM Journal on Discrete Mathematics*, 26(3):1265–1268, 2012.
- N. Alon and A. Mehrabian. On a generalization of Meyniel's conjecture on the Cops and Robbers game. *Electronic Journal of Combinatorics*, 18(1):Paper 19, 7 pages, 2011.
- A. Mehrabian. Lower bounds for the cop number when the robber is fast. *Combinatorics, Probability and Computing*, 20(4):617–621, 2011.
- A. Mehrabian. The capture time of grids. *Discrete Mathematics*, 311(1):102–105, 2011.
- S. Akbari, A. Daemi, O. Hatami, A. Javanmard, and A. Mehrabian. Zero-sum flows in regular graphs. *Graphs and Combinatorics*, 26(5):603–615, 2010.

Refereed Conference Proceedings

- Hassan Ashtiani, Shai Ben-David, and Abbas Mehrabian. Sample-efficient learning of mixtures. *arXiv preprint arXiv:1706.01596*, 2017. Accepted for presentation in the AAAI Conference on Artificial Intelligence (AAAI'18).
- Nick Harvey, Christopher Liaw, and Abbas Mehrabian. Nearly-tight VC-dimension bounds for piecewise linear neural networks. In Satyen Kale and Ohad Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory (COLT)*, volume 65 of *Proceedings of Machine Learning Research*, pages 1064–1068, Amsterdam, Netherlands, 07–10 Jul 2017. PMLR.
- Omer Angel, Abbas Mehrabian, and Yuval Peres. The String of Diamonds Is Tight for Rumor Spreading. In Klaus Jansen, José D. P. Rolim, David Williamson, and Santosh S. Vempala, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*, volume 81 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:9, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Petra Berenbrink, Peter Kling, Christopher Liaw, and Abbas Mehrabian. Tight load balancing via randomized local search. In *Parallel and Distributed Processing Symposium (IPDPS), 2017 IEEE International*, pages 192–201. IEEE, 2017.
- J. Janssen and A. Mehrabian. Rumours spread slowly in a small world spatial network. In *Algorithms and models for the web graph (WAW '15)*, volume 9479 of *Lecture Notes in Comput. Sci.*, pages 107–118. Springer, Cham, 2015.

- H. Acan, A. Collevecchio, A. Mehrabian, and N. Wormald. On the push&pull protocol for rumour spreading. In *Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing*, PODC '15, pages 405–412, New York, NY, USA, 2015. ACM.
- A. Mehrabian and N. Wormald. It's a small world for random surfers. In K. Jansen, J. D. P. Rolim, N. R. Devanur, and C. Moore, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM '14)*, volume 28 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 857–871. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014.
- A. Mehrabian and A. Pourmiri. Randomized rumor spreading in poorly connected small-world networks. In F. Kuhn, editor, *Distributed Computing (DISC '14)*, volume 8784 of *Lecture Notes in Computer Science*, pages 346–360. Springer Berlin Heidelberg, 2014.
- S. Ehsani, M. Fazli, A. Mehrabian, S. Sadeghian Sadeghabad, M. Safari, M. Saghafian, and S. Shokat-Fadaee. On a bounded budget network creation game. In *Proceedings of the 23rd ACM symposium on Parallelism in algorithms and architectures*, SPAA '11, pages 207–214, New York, NY, USA, 2011. ACM.
- S. Alamdari and A. Mehrabian. On a DAG partitioning problem. In A. Bonato and J. Janssen, editors, *Algorithms and Models for the Web Graph (WAW '12)*, volume 7323 of *Lecture Notes in Computer Science*, pages 17–28. Springer Berlin Heidelberg, 2012.
- A. Mehrabian. A randomly embedded random graph is not a spanner. In *Proceedings of the 23rd Canadian Conference on Computational Geometry*, CCCG '11, pages 373–374, 2011.

Submitted papers

- Peter L. Bartlett, Nick Harvey, Chris Liaw, and Abbas Mehrabian. Nearly-tight VC-dimension and pseudodimension bounds for piecewise linear neural networks. Submitted (extended abstract in COLT'17, full version available in https://arxiv.org/abs/1703.02930), 2017.
- Hassan Ashtiani, Shai Ben-David, and Abbas Mehrabian. Agnostic distribution learning via compression. *arXiv preprint arXiv:1710.05209*. Submitted, 2017.
- Luc Devroye, Abbas Mehrabian, and Pat Morin. First-passage percolation time on hypercubes. Submitted, 2017.
- Luc Devroye, Vida Dujmović, Alan Frieze, Abbas Mehrabian, Pat Morin, and Bruce Reed. Notes on growing a tree in a graph. *arXiv preprint arXiv:1707.00083*. Submitted, 2017.
- Omer Angel, Abbas Mehrabian, and Yuval Peres. The string of diamonds is tight for rumor spreading. *arXiv preprint arXiv:1704.00874*, 2017. Submitted, 2017 (extended abstract in APPROX/RANDOM'17).

Teaching Experience

- Currently teaching Math 139: Calculus 1 with precalculus, McGill University, Fall 2017.
- Two guest lectures at Monash University and two at University of Waterloo.
- Teaching assistantship at University of Waterloo for six courses.

Service

- Organized the probability lab seminars (Fall 2017) at McGill University, machine learning theory reading group (Spring – Fall 2016) at the University of British Columbia, expander graphs reading group (Spring – Fall 2012), concentration of random variables reading group (Fall 2011 – Winter 2012), and graph theory seminars (Spring 2011) at the University of Waterloo.
- Refereed articles for the following journals/conferences: ACM Transactions on Algorithms, Journal of Graph Theory, Random Structures & Algorithms, Internet Mathematics, Discrete Mathematics, Discrete Applied Mathematics, Theoretical Computer Science, and Journal of Combinatorial Optimization/COLT 2017, FOCS 2017, STOC 2017, SODA 2017, 2016, 2013, 2012; ICALP 2016, DISC 2014, APPROX 2014, and FUN 2012.
- Reviewed several articles and books for Zentralblatt MATH.