

Flavio du Pin Calmon

33 Oxford Street, MD 347, Cambridge MA, 02138

☎ 617-768-7484 • ✉ flavio@seas.harvard.edu • 🌐 <https://www.seas.harvard.edu/~flavio>

Professional Experience

Harvard University <i>Assistant Professor of Electrical Engineering</i>	Cambridge, MA <i>July 2017–present</i>
IBM T.J. Watson Research Center <i>Inaugural Data Science for Social Good Post-Doctoral Fellow</i>	Yorktown Heights, NY <i>2015–2017</i>
Technicolor Research <i>Research Intern</i>	Palo Alto, CA <i>2012</i>

Education

Massachusetts Institute of Technology (MIT) <i>Ph.D. in Electrical Engineering and Computer Science</i>	Cambridge, MA <i>2015</i>
Universidade Estadual de Campinas (Unicamp) <i>M.Sc. in Electrical Engineering</i>	São Paulo, Brazil <i>2009</i>
Universidade de Brasília (UnB) <i>B.Sc. in Communications Engineering</i>	Distrito Federal, Brazil <i>2006</i>

Awards, Honors, and Fellowships

- 2019–2024: **NSF Faculty Early Career Development Program (CAREER) Award.**
- 2019: **Google Faculty Research Award.**
- 2018: **IBM Open Collaborative Research Award** for investigating fairness in machine learning.
- 2018–2019: **Lemann Brazil Research Award** for teaching Machine Learning in Latin America.
- 2018: **Nvidia GPU Grant.**
- 2017: **Dean’s Competitive Fund for Promising Scholarship**, Harvard University.
- 2016: **IBM inaugural Social Good Post-Doctoral Fellowship** for research in data science projects that promote social good.
- 2011: **Avery Alan Ashdown Leadership Award** for outstanding service and leadership to the MIT community.
- 2009: **Irwin Mark Jacobs and Joan Klein Jacobs Presidential Fellowship** for graduate studies at MIT.
- 2007-2009: **The State of São Paulo Research Foundation (FAPESP) scholarship** for M.Sc. research.
- 2006: **Honor of Merit Award from the Brazilian Council of Engineers** for best overall academic performance in the Communications Engineering graduating class of the University of Brasilia.
- 2004-2006: **Brazilian National Council for Scientific Development (CNPq) scholarship** for undergraduate research.

Sponsored Research

- 2019–2021: **NSF EAGER: AI-DCL: Collaborative Research: Understanding and Overcoming Biases in STEM Education using Machine Learning (Lead PI).** Co-PIs: Muriel Médard (MIT), Nilanjana Dasgupta (UMass Amherst).
- 2019–2021: **NSF CIF: Medium: Collaborative Research: Information-theoretic Guarantees on Privacy in the Age of Learning (co-PI).** Lead PI: Lalitha Sankar (ASU). Co-PI: Oliver Kosut (ASU).
- 2019–2024: **NSF CAREER: Information-Theoretic Foundations of Fairness in Machine Learning (Lead PI).**

Tutorials

- 2019: **International Symposium on Information Theory Tutorial: Privacy and Fairness in Data Science: An Information-theoretic Perspective.**
- 2019: **Simons Institute:** Speaker at the Symposium on Information-Theoretic Methods for Privacy.

Teaching

- **Harvard SEAS:**
 - Fall 2019: **Information Theory (ENG-SCI 250).**
 - Spring 2019: **Signals and Systems (ENG-SCI 156).** Instructor Rating: 4.7/Course rating: 4.4.

Spring 2018: Signals and Systems (ENG-SCI 156). Instructor Rating: 4.9/Course rating: 4.1.

Fall 2017: Information Theory (ENG-SCI 250). Instructor Rating: 4.8/Course rating: 4.6.

○ **Outreach:**

August 2019: Machine Learning for Engineers. Taught at the Faculdade de Engenharia Elétrica e da Computação (FEEC), Universidade de Campinas (Unicamp), Brazil.

○ **Harvard Business School:**

Spring and Fall 2019:: Harvard Business Analytics Program. Lecturer for on-campus immersion program.

○ **MIT:**

Spring 2014: MIT Teaching Certificate Program for graduate students

Professional Activities and Service

2019–2022: Member of the Harvard Brazil Studies Program Faculty Steering Committee.

2019: Publicity Chair for the IEEE North American School of Information Theory (Boston, July 2019).

2019: Technical Program Committee Member for the IEEE International Symposium on Information Theory (ISIT).

2018–Present: NSF Panel Member for CISE CCF.

2018: Program Committee Member for the ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*).

2018: Member of the Harvard Brazil Studies Program Faculty Advisory Committee.

2018–Present: Member of the Harvard SEAS EE graduate admissions committee.

2016–Present: Session chair for Information Theory and Applications Workshop (2016,2017,2018), Allerton Conference on Communication, Control, and Computing (2018), and IEEE ISIT (2018)

2017: Data Jam organizer at the KDD 2017 Broadening Participation in Data Mining Workshop for underrepresented students.

2014-2015: Mentor for the MIT undergraduate research opportunities program (UROP).

2008–Present: Reviewer for the IEEE Trans. on Info. Theory, IEEE Trans. on Wireless Communications, IEEE Trans. on Info. Forensics and Security, IEEE International Symposium on Info. Theory, IEEE Info. Theory Workshop, IEEE International Conference on Communications, IEEE Vehicular Technology Conference

Students, Post-Docs, and Visiting Scholars

○ **Graduate Student Advisees:**

- **Hao Wang (G3, Applied Math):** Information-theoretic foundations of privacy and fairness in machine learning.

- **Hsiang Hsu (G2, Computer Science):** Information-theoretic foundations of representation learning.

- **Wael Alghamdi (G2, Applied Math):** Estimating distribution functionals from data.

- **Madeleine Barowsky (G1, Computer Science):** An information-theoretic view of interpretable machine learning.

○ **Master Student Advisees:**

- **Filip Michalsky (ME in Computational Science & Engineering):** Learning discriminatory patterns in middle school science education from data.

- **Claire (Zheng) Yang (ME in Computational Science & Engineering, jointly advised with Prof. Demba Ba):**

○ **Undergraduate Research Advisees (Harvard College):**

- **Marguerite Basta:** Image processing for medical applications.

- **Lisa Vo:** Estimation-theoretic view of privacy.

- **David Xu:** Computational creativity using machine learning.

○ **Post-doc Advisees:**

- **Berk Ustun (CRCS Post-doc):** Fairness using counterfactual distributions

- **Shahab Asoodeh:** Joined Feb 25th, 2019.

- **Javier Zazo (CRCS Post-doc, jointly sponsored with Prof. Demba Ba):** Scalable algorithms for optimal transport.

○ **Visitors Hosted:**

- **Prof. José Cândido Siveira Santos Filho (Unicamp/Brazil):** Creating a summer course on machine learning for Brazilian electrical engineering students (June 2018–June 2019)

Invited Seminars and Talks

○ **2019**

- Universidade Estadual de Campinas (Unicamp), Brazil

- Invited Session on Fairness and Privacy, IEEE International Symp. on Info. Theory

- Hamilton Institute Seminar, NUI Maynooth, Ireland
- Brown University Data Science Colloquium
- Anheuser-Busch InBev Board Annual Retreat
- Microsoft Research Seattle
- Google Research Seattle
- Simons Institute Symposium on Information-Theoretic Methods for Privacy
- Northeastern University SPIRAL Seminar Series
- Boston University CISE Seminar
- Information Theory and Applications Workshop (ITA), San Diego
- Harvard IACS Computefest
- **2018**
 - Worcester Polytechnic Institute, Electrical Engineering Seminar
 - Harvard Statistics Seminar
 - Harvard IACS Seminar
 - Stanford ISL Information Theory Forum
 - Mitsubishi Electric Research Laboratories
 - UMass Amherst Information Theoretic Privacy Workshop
 - Information Theory and Applications Workshop (ITA), San Diego
- **2017**
 - Machine Learning for Creativity Workshop at SIGKDD'17, Halifax, CA
 - MIT (Prof. Yury Polyanskiy's Group Meeting)
 - NSF Workshop, University of Delaware, 2017
- **Pre-2017**
 - 50th Conference on Information Sciences and Systems (CISS), Princeton NJ, 2016
 - New Jersey Institute of Technology, 2016
 - IBM T.J. Watson Research Center, 2015
 - Harvard University, EE Seminar Series, 2015
 - University of Southern California, 2015
 - Princeton University, 2014
 - University of Illinois at Urbana-Champaign, 2014
 - Interdisciplinary Workshop on Data Privacy, Hamilton Institute, Ireland, 2014
 - Hamilton Institute, Ireland, 2013
 - Polytech Annecy-Chambéry, Annecy, France, 2013
 - UC Berkeley, California, 2012
 - Technicolor Research, Palo Alto, California 2012

Publications

Journal Publications.....

[J1] M. Diaz, H. Wang, F. P. Calmon, and L. Sankar, "On the robustness of information-theoretic privacy measures and mechanisms," *IEEE Trans. Inf. Theory (accepted)*, 2019.

[J2] H. Wang, L. Vo, F. P. Calmon, M. Médard, K. R. Duffy, and M. Varia, "Privacy with estimation guarantees," *IEEE Trans. Inf. Theory (accepted)*, 2019.

[J3] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Trans. Inf. Theory (accepted)*, 2019.

[J4] S. Majumdar, B. Han, F. P. Calmon, B. Glicksberg, R. Horesh, A. Kumar, A. Perer, E. V. Marschall, D. Wei, A. Mojsilović, and K. Varshney, "Confronting data sparsity to identify potential sources of zika virus spillover infection among primates," *Epidemics*, vol. (in print), 2019.

[J5] F. P. Calmon, D. Wei, B. Vinzamuri, K. N. Ramamurthy, and K. Varshney, "Data pre-processing for discrimination prevention: Information-theoretic optimization and analysis," *IEEE J. Sel. Topics Signal Proces*, vol. 12, no. 5, pp. 1106–1119, Oct. 2018.

[J6] F. P. Calmon, Y. Polyanskiy, and Y. Wu, "Strong data processing inequalities for input constrained additive noise channels," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1879–1892, March 2018.

[J7] J. Liao, L. Sankar, V. Y. F. Tan, and F. P. Calmon, "Hypothesis testing under mutual information privacy constraints in the high privacy regime," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 4, pp. 1058–1071, 2018.

[J8] F. P. Calmon, Á. A. M. de Medeiros, and M. D. Yacoub, "Mutual outage probability," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3138–3150, 2017.

- [J9] F. P. Calmon, A. Makhdoumi, M. Médard, M. Varia, M. Christiansen, and K. R. Duffy, "Principal inertia components and applications," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5011–5038, 2017.
- [J10] M. M. Christiansen, K. R. Duffy, F. P. Calmon, and M. Médard, "Multi-user guesswork and brute force security," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6876 – 6886, Dec 2015.
- [J11] S. Salamatian, A. Zhang, F. Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "Managing your private and public data: Bringing down inference attacks against your privacy," *IEEE J. Sel. Topics Signal Proces*, vol. 9, no. 7, pp. 1240–1255, 2015.
- [J12] A. Rezaee, F. P. Calmon, L. M. Zeger, and M. Médard, "Speeding multicast by acknowledgment reduction technique (SMART) enabling robustness of QoE to the number of users," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 7, pp. 1270 –1280, Aug. 2012.
- [J13] F. P. Calmon and M. D. Yacoub, "MRCS – selecting maximal ratio combined signals: a practical hybrid diversity combining scheme," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3425–3429, Jul. 2009.

Selected Peer-Reviewed Conference Proceedings.....

- [C1] H. Hsu, S. Asoodeh, and F. P. Calmon, "Discovering information-leaking samples and features," in *NeurIPS Workshop on Privacy and Machine Learning*, 2019.
- [C2] C. Marx, F. P. Calmon, and B. Ustun, "On the multiplicity of predictions in classification," in *NeurIPS Workshop on Human-Centric Machine Learning*, 2019.
- [C3] H. Wang, B. Ustun, and F. P. Calmon, "Repairing without retraining: Avoiding disparate impact with counterfactual distributions," in *Proc. International Conference on Machine Learning (ICML)*, 2019.
- [C4] H. Hsu, S. Salamatian, and F. P. Calmon, "Correspondence analysis using neural networks," in *Int. Conf. on Artificial Intelligence and Statistics (AISTATS)*, 2019.
- [C5] H. Wang, M. D. Torres, J. C. S. S. Filho, and F. P. Calmon, "Generalization bounds via wasserstein distance," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2019.
- [C6] H. Hsu, S. Asoodeh, and F. P. Calmon, "Information-theoretic privacy watchdogs," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2019.
- [C7] W. Algahmdi and F. P. Calmon, "Mutual information as a function of moments," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2019.
- [C8] J. Liao, L. Sankar, O. Kosut, and F. P. Calmon, "Robustness of maximal α -leakage to side information," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2019.
- [C9] H. Wang, B. Ustun, and F. P. Calmon, "Avoiding disparate impact with counterfactual distributions," in *NeurIPS Workshop on Ethical, Social and Governance Issues in AI*, 2018.
- [C10] H. Hsu, F. P. Calmon, J. C. S. Santos Filho, A. P. Calmon, and S. Salamatian, "Correspondence analysis of government expenditure patterns," in *NeurIPS Workshop on Machine Learning for the Developing World (ML4D)*, 2018.
- [C11] H. Hsu, S. Asoodeh, S. Salamatian, and F. P. Calmon, "Generalizing Bottleneck Problems," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 531–535, 2018.
- [C12] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "A tunable measure for information leakage," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 701–705, 2018.
- [C13] H. Wang, B. Ustun, and F. P. Calmon, "On the direction of discrimination: An information-theoretic analysis of disparate impact in machine learning," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 1216–1220, 2018.
- [C14] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "Privacy under hard distortion constraints," *IEEE Information Theory Workshop*, 2018.
- [C15] H. Wang, M. Diaz, F. P. Calmon, and L. Sankar, "The utility cost of robust privacy guarantees," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 706–710, 2018.
- [C16] H. Wang and F. P. Calmon, "An estimation-theoretic view of privacy," in *Proc. 55th Annual Allerton Conference on Communication, Control, and Computing*, 2017.

- [C17] J. Liao, L. Sankar, F. P. Calmon, and V. Y. Tan, "Hypothesis testing under maximal leakage privacy constraints," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2017, pp. 779–783.
- [C18] F. P. Calmon, D. Wei, B. Vinzamuri, K. N. Ramamurthy, and K. R. Varshney, "Optimized pre-processing for discrimination prevention," in *Advances in Neural Information Processing Systems (NIPS)*, 2017, pp. 3995–4004.
- [C19] M. Riemer, A. Vempaty, F. P. Calmon, F. Heath, R. Hull, and E. Khabiri, "Correcting forecasts with multifactor neural attention," in *Proc. International Conference on Machine Learning (ICML)*, June 2016.
- [C20] J. Liao, L. Sankar, V. Y. F. Tan, and F. P. Calmon, "Hypothesis testing in the high privacy limit," in *Proc. 54th Annual Allerton Conference on Communication, Control, and Computing*, 2016.
- [C21] A. Makhdoumi, F. P. Calmon, and M. Médard, "Forgot your password: Correlation dilution," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 2944–2948, 2015.
- [C22] F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 1796–1800, 2015.
- [C23] F. P. Calmon, Y. Polyanskiy, and Y. Wu, "Strong. data processing inequalities in power-constrained Gaussian channels," *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 2558–2562, 2015.
- [C24] F. P. Calmon, M. Varia, and M. Médard, "An exploration of the role of principal inertia components in information theory," in *Proc. IEEE Inf. Theory Workshop*, Nov. 2014.
- [C25] —, "On information-theoretic metrics for symmetric-key encryption and privacy," in *Proc. 52nd Annual Allerton Conference on Communication, Control, and Computing*, 2014.
- [C26] F. P. Calmon, M. Varia, M. Médard, M. M. Christiansen, K. R. Duffy, and S. Tessaro, "Bounds on inference," in *Proc. 51st Annual Allerton Conference on Communication, Control, and Computing*, 2013.
- [C27] M. M. Christiansen, K. R. Duffy, F. P. Calmon, and M. Médard, "Brute force searching, the typical set and guesswork," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2013, pp. 1257–1261.
- [C28] —, "Guessing a password over a wireless channel (on the effect of noise non-uniformity)," in *Proc. Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 51–55.
- [C29] S. Salamatian, A. Zhang, F. P. Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data," in *Proc. IEEE GlobalSIP*, 2013.
- [C30] J. Cloud, F. P. Calmon, W. Zeng, G. Pau, L. M. Zeger, and M. Médard, "Multi-path TCP with network coding for mobile devices in heterogeneous networks," in *Proc. IEEE 78th Vehicular Technology Conference*, 2013, pp. 1–5.
- [C31] F. P. Calmon and N. Fawaz, "A framework for privacy against statistical inference," in *Proc. 50th Annual Allerton Conference on Communication, Control, and Computing*, 2012.
- [C32] F. P. Calmon, M. Médard, L. Zeger, J. Barros, M. M. Christiansen, and K. R. Duffy, "Lists that are smaller than their parts: A coding approach to tunable secrecy," in *Proc. 50th Annual Allerton Conference on Communication, Control, and Computing*, 2012.
- [C33] F. P. Calmon, M. Médard, and M. Effros, "Equivalent models for multi-terminal channels," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2011.
- [C34] F. P. Calmon and M. Yacoub, "A general exact formulation for the outage probability in interference-limited systems," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2008.

Patents (partial list of granted and filed).....

- [P1] F. P. Calmon, R. Kokku, and A. Vempaty, "Automatic learning curriculum generation," Aug. 2019, US Patent 10,373,511.
- [P2] F. P. Calmon and K. R. Varshney, "Accelerating data-driven scientific discovery," Aug. 2019, US Patent 10,388,039.
- [P3] F. P. Calmon, M. Medard, L. M. Zeger, M. M. Christiansen, and K. R. Duffy, "Method and apparatus for secure communication," Jun. 2019, US Patent 10,311,243.
- [P4] F. P. Calmon, J. M. Cloud, M. Medard, and W. Zeng, "Multi-path data transfer using network coding," Jun. 2018, US Patent 10,009,259.
- [P5] F. P. Calmon, R. T. Goodwin, A. Jagmohan, K. C. Ratakonda, and A. Vempaty, "Data-driven models for improving products," Mar. 2018, US Patent App. 15/291,685.

[P6] F. P. Calmon, J. M. Cloud, M. Médard, and W. Zeng, "Multi-path data transfer using network coding," Jan. 2017, US Patent 9,537,759.

[P7] M. Médard, F. P. Calmon, and W. Zeng, "Method and apparatus for implementing distributed content caching in a content delivery network," Jun. 2016, US Patent 9,369,541.

Pre-Prints

[PRP1] H. Hsu, S. Salamatian, and F. P. Calmon, "Deep Orthogonal Representations: Fundamental Properties and Applications," *arXiv:1806.08449 [cs, math, stat]*, 2018.

[PRP2] J. Zazo, F. P. Calmon, A. P. Calmon, and A. Vempaty, "Ground metric estimation for discrete optimal transport," (*in preparation*), 2019.

[PRP3] D. Wei, K. N. Ramamurthy, and F. d. P. Calmon, "Optimized score transformation for fair classification," *arXiv:1906.00066 [cs]*, 2019.

[PRP4] F. P. Calmon, M. Médard, M. Varia, K. R. Duffy, M. M. Christiansen, and L. M. Zeger, "Hiding symbols and functions: New metrics and constructions for information-theoretic security," *arXiv:1503.08513 [cs]*.

Citizenship

Citizen of the United States of America and Citizen of Brazil (dual citizenship).