Juan Felipe Gomez

juangomez@g.harvard.edu | http://felipe-gomez.com/ | Felipe-Gomez

Experience/Education

Harvard University *Expected 2026*
Ph.D. in Physics, Advised by Professor Flavio du Pin Calmon Cambridge, MA

Princeton University *Summer 2023-2025*
Summer Research Intern at the Center for Statistics and Machine Learning Princeton, NJ

California Institute of Technology *Sep. 2016 - Jun. 2020*
B.S. in Physics Pasadena, CA
Concentration in Condensed-Matter Physics, GPA 4.0/4.0

Awards and Honors

DOE CSGF Fellow *April 2020 - August 2024*
I was one of four Harvard graduate students in 2020 who were awarded the Department of Energy Computational Science Graduate Fellowship (DOE CSGF).

Goldwater Scholar *April 2019*
I was one of three Caltech students who won the Goldwater Scholarship in 2019.

Selected Research Projects

Unifying Re-Identification, Attribute Inference, and Data Reconstruction Risks in DP *2025*

- Showed that various common notions of privacy risk (see title of paper) can be bounded under a single unified framework; **NeurIPS 2025**.
- Showed that this framework improves upon existing bounds, which leads to more accurate downstream models.
- We provide an open-source library for practitioners to train Pytorch models with interpretable privacy guarantees.

Attack-aware Noise Calibration for Differential Privacy *2024*

- Showed that calibrating noise to an interpretable privacy risk (such as accuracy of inference attacks) instead of privacy parameter ε increases utility in language and vision models by over 10%; **NeurIPS 2024**.
- Similar to the work above, this work shows how to train models with interpretable privacy guarantees, and provides an open-source Pytorch implementation for ease of adoption.

Algorithmic Arbitrariness in Content Moderation *2024*

- Empirically observed that LLMs finetuned for content moderation exhibit predictive multiplicity, where equally accurate models assign conflicting predictions to the same content; **FAccT 2024**.
- Made concrete policy suggestions for content moderation and intermediary liability laws, which led to a policy brief for G20 2024.
- This work showed that predictive multiplicity is a practical problem and provides a road map for how to empirically measure and mitigate it.

The Saddle-Point Method in Differential Privacy *2023*

- Derived a constant-time closed-form formula for computing the privacy parameter ε under high composition; **ICML 2023**.
- Led implementation and experiments; matched/beat numerical baselines; see open-source code here.
- Our approach removes numerical difficulties from computing ε, making DP algorithms more reliable.

Schrödinger Mechanisms: Optimal Differential Privacy Mechanisms for Small Sensitivity *2023*

- Derived scalar optimal DP mechanisms in the practical small-sensitivity and high composition regime; **ISIT 2023**.
- Demonstrated a connection between quantum physics and differential privacy.
- Relevance to Human-Centered AI: Clarified when and why Gaussian noise is optimal; which increases confidence in its widespread use in AI.

Publications

I publish and collaborate across areas. Each field has its own conventions for ordering authors. Below, starred publications highlight publications where I am first author or co-first author.

[1] B. Kulynych*, **J. F. Gomez***, G. Kaissis, J. Hayes, B. Balle, F. P. Calmon, and J.-L. Raisaro, “Unifying re-identification, attribute inference, and data reconstruction risks in differential privacy”, arXiv preprint arXiv:2507.06969, 2025, (to appear in NeurIPS 2025).

[2] A. Gilani, **J. F. Gomez**, S. Asoodeh, F.P. Calmon, O. Kosut, and L. Sankar. ”Optimizing Noise Distributions for Differential Privacy”. In Forty-second International Conference on Machine Learning.

[3] B. Kulynych*, **J. F. Gomez***, G. Kaissis, F. P. Calmon, and C. Troncoso, “Attack-aware noise calibration for differential privacy,” in Advances in Neural Information Processing Systems, vol. 37, 2024, pp. 134868–134901.

[4] **J. F. Gomez***, C. Machado*, L. M. Paes, and F. P. Calmon, “Algorithmic arbitrariness in content moderation,” in Proc. 2024 ACM Conf. on Fairness, Accountability, and Transparency (FAccT ’24), 2024, pp. 2234–2253.

[5] W. Alghamdi, **J. F. Gomez**, S. Asoodeh, F. P. Calmon, O. Kosut, and L. Sankar, “The saddle-point method in differential privacy,” in Proc. 40th Int. Conf. on Machine Learning (ICML), 2023, pp. 508–528.

[6] W. Alghamdi, S. Asoodeh, F. P. Calmon, **J. F. Gomez**, O. Kosut, and L. Sankar, “Schrödinger mechanisms: Optimal differential privacy mechanisms for small sensitivity,” in Proc. IEEE Int. Symp. on Information Theory (ISIT), 2023, pp. 2201–2206.

[7] W. Alghamdi, S. Asoodeh, F. P. Calmon, **J. F. Gomez**, O. Kosut, and L. Sankar, “Optimal multidimensional differentially private mechanisms in the large-composition regime,” in Proc. IEEE Int. Symp. on Information Theory (ISIT), 2023, pp. 2195–2200.

[8] Y. Wang, P. A. Lee, D. M. Silevitch, **J.F. Gomez** et al., “Antisymmetric linear magnetoresistance and the planar Hall effect,” Nat. Commun., vol. 11, Art. no. 216, 2020.

Invited and Contributed Talks

Algorithmic Arbitrariness in Content Moderation *Jan 2026*

Contributed Talk at AAAI Model Uncertainty Workshop (forthcoming)

Towards Practical Differentially Private Algorithms *July 2025*

DOE CSGF Outgoing Fellow Talk

Attack Aware Noise Calibration for Differential Privacy *Feb 2025*

Invited Talk at OpenAI

Schrödinger mechanisms and Optimal Multidimensional Mechanisms *Jul 2023*

Two Contributed Talks at IEEE Int. Symp. on Information Theory

Service

Peer Reviewing

I have served as a reviewer for ICML 2023-2025, Neurips 2023-2025, FAcct 2024-2025, and ISIT 2025

Teaching Fellow, Harvard University

Note: the DOE CSGF prevented me from teaching from Spring 2021 - Spring 2024

Electromagnetism from an Analytic, Numerical and Experimental Perspective (Dr. Gregorio Ponti, Dr. Anna Wang-Holtzen), Fall 2025

Advanced Scientific Computing: Stochastic Methods for Data Analysis, Inference, and Optimization (Professor Weiwei Pan), Fall 2021