Vincent Zaballa — Ph.D. Candidate in Biomedical Engineering

Irvine, CA
vzaballa@uci.edu • vz415.github.io/bio_ml_blog • vz415

Education

University of California, Irvine — **Irvine, CA**
Ph.D., Biomedical Engineering — *2019–2025 (expected)*
Thesis: "Machine Learning Methods for Posterior Approximation, Mutual Information Optimization, and Integrating Structural and Systems Biology in Drug Discovery."

Imperial College London — **London, UK**
M.Res., Bioengineering — *2015–2016*

Texas A&M University — **College Station, TX**
B.S. & M.Eng., Biomedical Engineering, Minor in Electrical Engineering, cum laude — *2010–2015*

Research Interests

My research focuses on unraveling biological mysteries through the development of novel and scalable machine learning methods. I am particularly interested in advancing probabilistic approaches that align with the native structure of biological data. Generative models, such as diffusion processes and flow matching, represent a promising area for Bayesian optimal experimental design (BOED). By incorporating techniques like calibration and conformal prediction, these models have the potential to improve active learning and experimental efficiency, with applications in biology. I am also deeply committed to open-source research; I developed LFIAX and plan to release additional tools to promote reproducibility and transparency in the field.

Experience

University of California, Irvine — **Irvine, CA**
Graduate Research Assistant — *Sept. 2019–Present*

- Proved new method to simultaneously optimize an intractable likelihood function and experimental designs, resulting in 52% better predictive accuracy over competing methods (*coded in LFIAX*). (In Review)
- Developed new method to perform diffusion guidance applied to small molecule generation in systems biology and structure-based drug design (*NeurIPS MLSB, 2024*).
- Demonstrated a new approach to integrating structural and systems biology information, improving predictions by 1.34% without collecting additional experimental data (*ICML ML4LMS, 2024*).
- Modeled the intractable likelihood of a high-dimensional dataset of the Bone Morphogenetic Protein (BMP) signaling pathway using normalizing flows (*NeurIPS GenBio, 2023*).
- Created a novel approach for optimizing Mutual Information & experimental designs in Simulation-based Inference, improving feasibility in high-dimensional design optimization (*ICML DAE, 2023*).
- Created a new method for automated biological model selection with 166–298% better accuracy over baseline methods (*ICML CompBio, 2022*).
- Improved average experimental design efficiency by 225% using surrogate probabilistic modeling of physics-based biological models and Bayesian optimization (*NeurIPS LMRL, 2021*).

Fannin Innovation Studio — **Houston, TX**
Project Manager — *June 2018–Aug. 2019*

- Managed commercialization and R&D of a pioneering fetal surgery device, resulting in a $2 million SBIR Phase II Grant.
- Oversaw engineering efforts (two hardware engineers and one software engineer), introduced new quality control computer vision software in OpenCV.

Drip — **Colorado Springs, CO**
Data Scientist — *April 2017–Oct. 2017*

- Developed analytics and machine learning algorithms to draw insights from Colorado water data records within a SaaS application.
- Member of Go Code Colorado 2017 Hackathon winning team.

Cambiando Paradigmas — **La Paz, Bolivia**
Data Analyst — *Jan. 2017–June 2018*

- Improved returns from renewable energy investments by 4.5x over previous portfolio, using random forests and optimization.

Imperial College London — **London, UK**
Whitaker Fellow — *Oct. 2015–Oct. 2016*
- Designed a novel orthopedic implant that reduces typical bone strain by 90% (validated via finite element analysis).

Covidien, Ltd. — **Boulder, CO**
Electrosurgical R&D Intern — *May 2014–Aug. 2014*
- Designed, created, and tested a surgical instrument for wound healing (*US Patent US10278685B2*).

Biomedical Device Lab, Texas A&M — **College Station, TX**
Research Assistant — *Jan. 2013–May 2015*
- **Undergraduate Research Scholar Thesis:** "Characterization Of Automatically Reticulated Shape Memory Polymer Foams."
- Created an automated system for processing Shape Memory Polymer (SMP) foams, reducing manufacturing time by 4.5 hours.
- Prototyped and tested new *in vitro* blast wound model using new SMP foam designs.

Publications

2024: **V. Zaballa**, E. Hui, "Optimizing Likelihoods via Mutual Information: Bridging Simulation-Based Inference and Bayesian Optimal Experimental Design." *Under Review.*

2024: **V. Zaballa**, E. Hui, "Systems-Structure-Based Drug Design." *NeurIPS MLSB Workshop, 2024.*

2024: **V. Zaballa**, E. Hui, "Reducing Uncertainty Through Mutual Information in Structural and Systems Biology." *ICML ML4LMS Workshop, 2024.*

2023: **V. Zaballa**, E. Hui, "Approximation of Intractable Likelihood Functions in Systems Biology via Normalizing Flows." *NeurIPS GenBio Workshop, 2023.*

2023: **V. Zaballa**, E. Hui, "Stochastic Gradient Bayesian Optimal Experimental Designs for Simulation-based Inference." *ICML Differentiable Almost Everything Workshop, 2023.*

2022: **V. Zaballa**, E. Hui, "An Optimal Likelihood Free Method for Biological Model Selection." *ICML Workshop on Computational Biology, 2022.*

2015: **V. Zaballa**, D. Friedrichs, "A Novel Tissue Welding Device using Low-Voltage Far-Field Coaxial Electrospinning." *ASME Design of Medical Devices Conference Journal, Jan. 2015.*

Invited Talks & Presentations

Apr. 2022: "Optimal Design of Experiments for Simulation-Based Inference of Mechanistic Acyclic Biological Networks." *Lightning Talk at SoCalSysBio 2022, Los Angeles, CA.*

Funding

2022: NIH F31 Ruth R. L. Kirschstein Research Service Award Fellowship: $125,000

Teaching Experience

Biomechanics III — **University of California, Irvine**
Teaching Assistant — *Spring 2025 (anticipated)*
- Led recitation sections and addressed questions in a junior-level biomechanics course focusing on solving (partial) differential equations in biomechanics.

Biology 50A: Cell and Molecular Engineering — **University of California, Irvine**
Teaching Assistant — *Winter 2025*
- Led recitation sessions and addressed questions in a sophomore-level biology and synthetic biology class.

Cancer Systems Biology Short Course — **University of California, Irvine**
Teaching Assistant — *June 2022*
- Taught students how to run end-to-end Python gene expression processing notebooks.

Maths 2 — **Imperial College London**
Teaching Assistant — *Oct.–Dec. 2015*
- Led recitation sessions in multivariate calculus and differential equations for second-year students.

Academic Service

2021–2024: Reviewer, AISTATS (2022, 2023, 2025)

2024: Reviewer, ICLR 2025

2022: Contributor, *Probabilistic Machine Learning: Advanced Topics* by Kevin Murphy
– Created normalizing flow JAX examples in Google Colab used in the book.

Software

Contributor, **JAX Optax** optimization library (2023–2024)
Creator, **LFIAX**: JAX conditional normalizing flows (11 stars) (2022)

Awards & Fellowships

2022: ARCS Scholar Award ($\$10,000$)
2021: Google Computer Science Research Mentorship Program
2021: Summer Inclusive Excellence Award – for promoting minority access to undergraduate research ($\$5,000$)
2017: Go Code Colorado Hackathon Winning Team ($\$25,000$)
2015–2016: Whitaker International Fellowship
2015: Imperial College London Postgraduate Award for Outstanding Students ($\$10,000$)