Alexander Alemi

email: me@alexalemi.com
URL: http://alexalemi.com
papers: google scholar
github: alexalemi

Current position

2025-present *Member of Technical Staff*, Anthropic

Previous experience

2024 *Senior Research Scientist*, Google Deepmind
2018-2024 *Senior Research Scientist*, Google Research, Machine Perception
2016-2018 *Software Engineer*, Google, Mountain View, CA.
2015-2016 *Postodoc*, Disney Research, Boston
2009-2015 *Research Associate*, Laboratory of Atomic and Solid-State Physics (LASSP), Cornell
2015 *Research Intern*, Google, Mountain View, CA.
2014 *Research Intern*, Disney Reseach, Boston.

Areas of specialization

Neural Scaling Laws • Approximate Bayesian Inference • Information Theory & Deep Learning • Representation Learning • Physics • Machine Learning • Python Programming

Education

2015 PhD in Physics, Cornell with Jim Sethna
2013 MSc in Physics, Cornell
2009 BSc in Physics, California Institute of Technology (Caltech)

Publications

2024 K Everett, L Xiao, M Worstman, AA Alemi, R Novak, PJ Liu, I Gur, J Sohl-Dickstein, LP Kaelbling, J Pennington "Scaling Exponents Across Parameterizations and Optimizers" *ICML* arXiv:2407.05872

B Lester, J Lee, AA Alemi, J Pennington, A Roberts, J Sohl-Dickstein, N Constant "Training LLM over Neurally Compressed Text" arXiv:2404.03626

2023 PAGI "Beyond Human Data: Scaling Self-Training for Problem-Solving with Language Models" *TMLR* arXiv:2312.06585

PAGI "Frontier Language Models are not Robust to Adversarial Arithmetic, or "What do I need to say to you so you agree 2+2=5?"" arXiv:2311.07587

M Wortsman, PAGI "Small-scale proxies for large-scale Transformer training instabilities" *ICLR 2024 oral* arXiv:2309.14322

I Seroussi, AA Alemi, M Helias, Z Ringel "Speed Limits for Deep Learning" arXiv:2307.14653

AA Alemi, B Poole "Variational Prediction" Submitted to AABI

2022 Y Ruan, S Singh, WR Morningstar, AA Alemi, S Ioffe, I Fischer, JV Dillon "Weighted Ensemble Self-Supervised Learning" *ICLR 2023*, arXiv:2211.09981

Y Du, D Ho, AA Alemi, E Jang, M Khansari "Bayesian Imitation Learning for End-to-End Mobile Manipulation" *ICML 2022*, arXiv:2202.07600

2021 I Korshunova, D Stutz, AA Alemi, O Wiles, S Gowal "A Closer Look at the Adversarial Robustness of Information Bottleneck Models" *ICML 2021 AML Workshop Poster*, arXiv:2107.05712

S Stanton, P Izmailov, P Kirichenko, AA Alemi, AG Wilson "Does Knowledge Distillation Really Work?" *NeurIPS2021*, arXiv:2106.05945

2020 AA Alemi, WR Morningstar, B Poole, I Fischer, JV Dillon "VIB is Half Bayes", *AABI 2021 Oral*, arXiv:2011.08711

WR Morningstar, AA Alemi, JV Dillon "PAC^m-Bayes: Narrowing the Empirical Risk Gap in the Misspecified Bayesian Region" *AISTATS2022*, arXiv:2021.09629

WR Morningstar, C Ham, AG Gallagher, B Lakshminarayanan, AA Alemi, JV Dillon "Density of States Estimation for Out-of-Distribution Detection", *AISTATS 2022 Oral*, arXiv:2006.09273

DS Karls, M Bierbaum, AA Alemi, RS Elliot, JP Sethna, EB Tadmor "The OpenKIM Processing Pipeline: A Cloud-Based Automatic Materials Property Computation Engine", *ChemPhys* arXiv:2002.05380

I Fischer, AA Alemi "CEB Improves Model Robustness", *Entropy*, arXiv:2002.05380

2019 R Novak, L Xiao, J Hron, J Lee, AA Alemi, J Sohl-Dickstein, SS Schoenholz "Neural Tangents: Fast and Easy Infinite Neural Networks in Python", *ICLR 2020* arXiv:1912.02803

AA Alemi "Variational Predictive Information Bottleneck", arXiv:1910.10831

R Shwartz-Ziv, AA Alemi "Information in Infinite Ensembles of Infinitely-Wide Networks", *AABI* arXiv:1911.09189

Z Dong, D Oktay, B Poole, AA Alemi "On Predictive Information in RNNs", arXiv:1910.09578

T Conte, E DeBenedictis, N Ganesh, T Hylton, JP Strachan, RS Williams, AA Alemi, L Altenberg, G Crooks, J Crutchfield, L del Rio, J Deutsch, M DeWeese, K Douglas, M Esposito, M Frank, R Fry, P Harsha, M Hill, C Kello, J Krichmar, S Kumar, SC Liu, S Lloyd, M Marsili, I Nemenman, A Nugent, N Packard, D Randall, P Sadowski, N Santhanam, R Shaw, A Stieg, E Stopnitzky, C Teuscher, C Watkins, D Wolpert, J Yang, Y Yufik, "Thermodynamic Computing", *CCC* arXiv:1911.01968

B Seybold, E Fertig, A Alemi, I Fischer, "Dueling Decoders: Regularizing Variational Autoencoder Latent Spaces", arXiv:1905.07478

I Fischer, A Alemi, JV Dillon, TFP Team, "Variational Autoencoders with Tensorflow Probability Layers", *Tensorflow Blog*

B Poole, S Ozair, A van den Oord, AA Alemi, G Tucker, "On Variational Bounds of Mutual Information", *ICML 2019* arXiv:1905.06922

CB Clement, M Bierbaum, KP O'Keeffe, AA Alemi, "On the Use of ArXiv as a Dataset", *ICLR 2019 workshop RLGM* arXiv:1905.00075

2018 AA Alemi, JV Dillon, I Fischer, "Uncertainty in the Variational Information Bottleneck", *UAI 2018 workshop on Uncertainty in Deep Learning, contributed oral* arXiv:1807.00906

AA Alemi, I Fischer, "TherML: Thermodynamics of Machine Learning", *ICML 2018 Workshop, contributed oral.*, arXiv:1807.04162

AA Alemi, I Fischer, "GILBO: One Metric to Measure Them All", *NIPS 2018 Spotlight* arXiv:1802.04874

JV Dillon, I Langmore, D Tran, E Brevdo, S Vasudevan, D Moore, B Patton, A Alemi, M Hoffman, RA Saurous, "TensorFlow Distributions", *arXiv:1711.10604*

S Abu-El-Haija, B Perozzi, R Al-Rfou, A Alemi, "Watch your step: Learning graph embeddings through attention", *NIPS 2018 Poster* arXiv:1710.09599

AA Alemi, B Poole, I Fischer, JV Dillon, RA Saurous, K Murphy, "Fixing a Broken ELBO", *Oral ICML 2018* arXiv:1711.00464

2017 LX Hayden, AA Alemi, PH Ginparg, JP Sethna, "Jeffrey's prior sampling of deep sigmoidal networks", *arXiv: 1705.10589*

K Fragkiadaki, J Huang, A Alemi, S Vijayanarasimhan, S Ricco, R Sukthankar, "Motion Prediction Under Multimodality with Conditional Stochastic Networks" *arXiv: 1705.02082*

M Bierbaum, BD Leahy, AA Alemi, I Cohen, JP Sethna, "Light Microscopy at Maximal Prediction", *PRX 7 (4), 041007*, arXiv: 1702.07336

C Szegedy, S Ioffe, V Vanhoucke, AA Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning." *AAAI, 4278-4284*, arXiv: 1602.07261

AA Alemi, I Fischer, JV Dillon, KP Murphy, "Deep Variational Information Bottleneck", *ICLR 2017*, arXiv: 1612.00410

B Poole, AA Alemi, J Sohl-Dickstein, A Angelova, "Improved generator objectives for GANS" *NIPS GAN Workshop 2016*, arXiv: 1612.02780

2016 R Shin, AA Alemi, G Irving, O Vinyals, "Tree-Structured Variational Autoencoder" *ICLR 2017 Submission*

G Irving, C Szegedy, AA Alemi, N EEn, F Chollet, J Urban, "Deepmath-deep sequence modesl for premise selection". *NIPS 2016*, arXiv: 1606.04442

2015 CJM Mathy, F Gonda, D Schmidt, N Derbinsky, AA Alemi, J Bento, FM Delle Fave, JS Yedidia, "SPARTA: Fast global planning of collision-avoiding robot trajectories". Learning, Inference and Control of Multi-Agent Systems Workshop, *NIPS 2015*

AA Alemi, "Zombies Reading Segmented Graphic Articles on the ArXiv". PhD Thesis, Cornell University

JM Cashore, X Zhao, AA Alemi, Y Liu, PI Frazier, "Clustering via Content-Augmented Stochastic Blockmodels". arXiv: 1505.06538

LX Hayden, R Chachra, AA Alemi, PH Ginsparg, JP Sethna, "Canonical Sectors and Evolution of Firms in the US Stock Markets". *Quantitative Finance*, arXiv: 1503.06205

AA Alemi, P Ginsparg, "Text Segmentation Based on Semantic Word Embeddings", arXiv: 1503.05543

AA Alemi, M Bierbaum, CR Myers, JP Sethna, "You Can Run, You Can Hide: The Epidemiology and Statistical Mechanics of Zombies", arXiv: 1503.01104. Phys Rev E 92. 052801

2014 A Taloni, AA Alemi, E Ciusani, JP Sethna, S Zapperi, CAM La Porta, "Mechanical Properties of Growing Melanocytic Nevi and the Progression to Melanoma", *PloS one*, 9 (4), e94229

2013 MM Baraldi, AA Alemi, JP Sethna, S Caracciolo, C La Porta, S Zapperi, "Growth and Form of Melanoma Cell Colonies", *Journal of Statistical Mechanics: Theory and Experiment*, 2013, 02, p02032

PY Huang, S Kurasch, JS Alden, A Shekhawat, AA Alemi, PL McEuen, JP Sethna, UTE Kaiser, DA Muller, "Imaging Atomic Rearrangements in Two-Dimensional Silica Glass: Watching Silica's Dance", *Science* 342, 6155 p224-227

2011 RS Ottens, V Quetschke, S Wise, AA Alemi, R Lundoc, G Mueller, DH Reitze, DB Tanner, BF Whiting, "Near-field radiative heat transfer between macroscopic planar surfaces", *Physical Review Letters* 107 (1), 014301

Talks

2024 *How to Think about AI*, Osceola Neovates

2023 *Information Theory for Representation Learning*, InfoCog Workshop @ NeurIPS 2023
What's Missing? A Speculative Sketch of the Future of Machine Learning and Science ML and the Physical Sciences Workshop @ NeurIPS 2023
Variational Prediction, AABI2023
A Tale of Two Worlds: The Variational Approach to Machine Learning, UCF CRCV
Inferential Engines - Theoretical Physics for Machine Learning, Aspen

2022 *PAC^m Bayes* - Your Model is Wrong Workshop, NeurIPS 2021

2021 *Machine Learning and Thermodynamcis* - Scientific Machine Learning Mini-Course (SciML) @ CMU
VIB is Half Bayes - Advances in Approximate Bayesian Inference Symposium

2020 *Machine Learning and Thermodynamics* - Informal Statistical Physics Seminar, University of Maryland
TherML - American Physical Society Topical Group on Data Science
Variational Predictive Information Bottleneck - Information Theory and Applications Workshop

2019 *A Case for Compression* - NeurIPS Workshop on Information Theory and Machine Learning
TherML - Machine Learning and Physics, Aspen

2018 *Focusing on the Representation* - Cornell AI Seminar
Thermodynamics and Machine Learning - Machine Learning and Statistical Physics Workshop at CUNY
Thermodynamics and Machine Leanring - Chez Pierre Seminar at MIT
Thermodynamics and Machine Learning - Physics Seminar at Cornell
Panelist at Advances in Approximate Bayesian Inference Symposium, colocated at NeurIPS.
Fixing a BrokenELBO - ICML
Uncertainty in VIB - UAI UDL Workshop

2015 *The Statistical Mechanics of Zombies* - APS March Meeting

2014 *Finding Structure in the ArXiv* - APS March Meeting

2012 *A Group Theoretic Approach to Nonlinear and Gradient Elastic Terms for Graphene and Carbon Nanotubes* - APS March Meeting

2006 *Why Venus has No Moon* - AAS Meeting

Community Contributions

2021-2024 Action Editor at TMLR

2021 Area Chair, NeurIPS

2020 Co-organizer for NeurIPS Workshop: Deep Learning through Information Geometry
Program Committee for Uncertainty in Deep Learning Workshop at ICML

2017-2023 Regular reviewer for ICML, ICLR, NeurIPS, UAI, AABI, TMLR, JMLR

Grants, honors & awards

2023 Expert Reviewer for TMLR

2021 Top Reviewer for UAI
Expert Reviewer for ICML

2019 Best Reviewer, ICML

2010 Stephen and Margery Russel Distinguished Teaching Award

2009 Upperclass Merit Award

2008 Upperclass Merit Award

2007 Green Memorial Prize
2005 National Merit Scholar - Siemen's Scholar

Teaching & Outreach

2024-present Mentor for Take Stock in Children
2023 Brainiversity - Introduction to Statistics through Randomization
2022 Brainiversity - Order of Magnitude Physics
2021 Brainiversity - Information Theory
2019 Guest Lecturer for CS 294-131: Trustworthy Deep Learning (Special Topics in Deep Learning), Berkeley
2018 Life outside Academia Talk to Cornell Graduate Students
2015 GRASSHOPR - Computing with Bins and Beans
2014 Expanding Your Horizons - The Physics of Bubbles
2013 Expanding Your Horizons - The Physics of Bubbles
2012 Expanding Your Horizons - The Physics of Bubbles
TA - Physics 2218 - Physics III: Thermodynamics, Statistical Mechanics and Wave Phenomenon
2011 Physics Department Teaching Assistant Training Coordinator
Expanding Your Horizons - The Physics of Bubbles
GRASSHOPR - ... And Physics for All
2010 Physics Department Teaching Assistant Training Trainer
grader - Physics 3317 - Applications of Quantum Mechanics
TA - Physics 2217 - Physics II: Electricity and Magnetism
2009 TA - Physics 2213 - Physics II: Heat/Electromagnitism

Last updated: October 2, 2025 • Typeset in XeLaTeX
http://cv.alexalemi.com