Zachary Bezemek

AI Research Scientist — Generative Models (Training & Inference)

zwb@duke.edu • (248) 978-5905 • https://bezemekz.github.io/home/

SUMMARY

AI Research Scientist specializing in diffusion and language models, designing training objectives and inference algorithms grounded in stochastic analysis. First-author ICLR Oral with substantial benchmark gains across text, code, and scientific domains. Currently leading an independent generative AI research program at Duke University.

EXPERIENCE

Phillip Griffith Assistant Research Professor of Mathematics — Duke University | Aug 2023–Present

- Lead an independent generative AI research program grounded in stochastic processes.
- Designed planner-aware diffusion training objectives and novel sampling strategies adopted in open-source implementations.
- First-author ICLR Oral. Introduced theory-informed masked diffusion LM training objective requiring minimal modification to standard MDM pipelines. At matched architecture and parameter scale, achieved +40% relative protein foldability (DPLM-150M, ESMFold), 4× MAUVE for unconditional text generation (110M MDLM, OpenWebText), and +23% relative code infilling pass@10 (Open-dLLM 0.5B, HumanEval) over the same models trained with standard losses.
- Developed novel masked diffusion language model sampling scheme. On fixed pretrained models, achieved relative improvements of +22% protein foldability (DPLM-150M, ESMFold, vs RDM sampling), +8% pLDDT for RNA generation (150M custom MDM, AlphaFold3, vs ancestral), and +65% code infilling pass@1 (7B DiffuLLaMA, HumanEval, vs entropy-based confidence) over the strongest published sampling baselines.
- Mentoring ML researchers executing experiments, large-scale evaluations, and open-source releases.

EDUCATION

PhD, Mathematics — Boston University | *Sep 2018 – May 2023*

- Advisor: Prof. Konstantinos Spiliopoulos
- Thesis: Interacting particle systems in multiscale environments: asymptotic analysis.
- Used tools from stochastic analysis, stochastic control, PDEs, and optimal transport to study rare events in systems of many interacting particles with applications to biology, ecology, social sciences, economics, molecular dynamics, and the study of spatially homogeneous granular media.

Bachelor of Science, Advanced Mathematics — Michigan State University | *Sep 2014 – May 2018*

- Additional Major in Japanese, Honors College.
- Undergraduate research projects in speech processing, modeling of crowd dynamics, and stochastic models of cancer cell response to chemotherapy.

SELECTED NON-ACADEMIC PROJECT

RivusVox Editor — Near-Live Zero-Shot Adaptive Speech Editing | 2024

- Built diffusion-based 320M parameter speech editing system designed for live-stream scenarios.

- Reduced inference time from several minutes to ~6 seconds on a single GTX 1080 Ti GPU via systematic inference pipeline optimization (phoneme alignment model replacement, improved silence detection and padding, eliminating redundant transfers between CPU and GPU).
- Developed necessary features to support live-stream audio editing (live transcription, transcript/audio/video stream alignment, dynamic rubberbanding).
- Implemented parameter-efficient fine-tuning (AdaSpeech) for accent adaptation and quality improvements.
- Maintained open-source codebase with live demos.

SELECTED PUBLICATIONS

- Planner Aware Path Learning in Diffusion Language Models Training — ICLR 2026 Oral. arXiv:2509.23405
- Path Planning for Masked Diffusion Model Sampling — ICLR Workshop Outstanding Paper Award 2025. arXiv:2502.03540
- Importance Sampling for the Empirical Measure of Weakly Interacting Diffusions — Applied Mathematics & Optimization, 2024. arXiv:2210.08591
- Moderate Deviations for Fully Coupled Multiscale Weakly Interacting Particle Systems — SPDE Analysis & Computations, 2023. arXiv:2202.08403
- Large Deviations for Interacting Multiscale Particle Systems — Stochastic Processes and their Applications, 2023. arXiv:2011.03032
- Rate of Homogenization for Fully-Coupled McKean-Vlasov SDEs — Stochastics and Dynamics, 2023. arXiv:2202.07753

AWARDS & DISTINCTIONS

- ICLR 2026 Oral Presentation (First Author).
- Outstanding Paper Award — ICLR Deep Generative Model in Machine Learning: Theory, Principle and Efficacy Workshop 2025.
- First Place — Deep Learning Boot Camp (Erdős Institute, 2024).
- Top 5 Project — Data Science Boot Camp (Erdős Institute, 2024).
- Lewis Blake Award for Excellence in Teaching — Duke University.

COMMUNICATION & LEADERSHIP

- Invited speaker at ICLR, ICIAM, Brin MRC, INFORMS APS, CRM, KTH, CMU, USC, TU Berlin, Seminar on Stochastic Processes, and other international research venues.
- Award-winning educator teaching probability, stochastic processes, and analysis at Duke University.
- Organizer of graduate student workshops and research programs and mentor to undergraduate students in mathematics and machine learning.

TECHNICAL SKILLS

Python • PyTorch • HuggingFace • NumPy/SciPy • Git/GitHub • Linux • LaTeX