

Department of Statistics Seminar Series



Sujay Sanghavi
Department of Electrical and
Computer Engineering
The University of Texas at Austin

Speeding Up and Guiding Diffusion Language Models

Wednesday, May 6th, 2026

12:10-1:30 PM

Hill Center - 552

Zoom Meeting: Meeting ID: 926 1608 1995

Password: 153525

<https://rutgers.zoom.us/j/92616081995?pwd=TqPVh6Y1O6aLUt8i2uUBCRhrgH6Ayx.1>

Light refreshments will be served in Hill 452 11:40am

Abstract:

Diffusion language models (DLMs) start from a set of noise/mask tokens which they iteratively refine into text in parallel. In this talk we focus on two unique capabilities of DLMs:

(A) Parallel Generation: In DLMs one can generate multiple tokens in each forward pass; however this results in drastic drops in accuracy. We identify the reason for this: parallel generation results in sampling tokens from the product marginals instead of the true joint distribution. We develop a new lightweight approximate sampling architecture that significantly fixes this issue.

(B) Test-time adaptation via reward guidance: diffusion models open the possibility of tilting the distribution from which outputs are generated at test time, with no training, based on gradient feedback from a downstream reward model. This has been shown to great efficacy for images; DLMs however have to contend with the discreteness (and non-differentiability) of tokens. We develop a new way to effectively reward-guide DLMs.

We demonstrate our results on open-weights diffusion LMs at the 7B scale. No prior knowledge of DLMs is assumed.

Bio:

Sujay Sanghavi is the Bettie Margret Smith Professor of ECE and CS at the University of Texas at Austin, where he conducts research on machine learning. At UT Austin, Sujay is the director of two NSF TRIPODS Institutes, is on the scientific board of its Center for Generative AI, and has founded the Amazon Science Hub. In industry, Sujay has been a founding quant of a mid-frequency stat-arb algorithmic trading team at Engineers Gate, and until recently a Principal Research Scientist and Amazon Scholar at Amazon. He is currently a Visiting Scientist at Google.