# Tejas Jayashankar

(630) 391-8118 | tejasj@mit.edu | https://www.linkedin.com/in/tkj97/ | github.com/tkj516

## Education

**Massachusetts Institute of Technology** — Cambridge, MA
*PhD Student in Electrical Engineering and Computer Science* — February 2022 – Present
***Signals, Information & Algorithms Lab*** — GPA: 5.00/5.00

**Massachusetts Institute of Technology** — Cambridge, MA
*Master of Science in Electrical Engineering and Computer Science* — August 2019 – February 2022

**University of Illinois at Urbana-Champaign** — Champaign, IL
*Bachelor of Science in Electrical Engineering with a Minor in Mathematics* — August 2015 – May 2019

## Research Interests and Objective

I am interested in leveraging generative AI, representation learning and information theory to develop novel methods for audio-visual signal synthesis, data compression and for solving inverse problems. I strive to combine my learnings from fundamental research conducted in an academic setting with my applied research skills developed through industry internships to build technologies centered on theoretical developments with practical applications.

## Academic Research

**Massachusetts Institute of Technology** — Cambridge, MA
*Graduate Research Assistant - PhD* — September 2022 – Present

- Developing new algorithms and architectures for generative AI, representation learning and density estimation with applications to neural compression, inverse problems and audio-visual synthesis.
- Implementing a new suite of *Bayesian-inspired algorithms* for solving *inverse problems* with *diffusion-based priors*. Results demonstrate that our method achieves state-of-the-art performance with reductions of up to 95% in terms of both MSE and bit-error rate (BER) over existing (traditional and score-based) methods when applied to the challenging class of digital communication signals.
- Developing end-to-end spectral representation learning techniques and novel architectures for learning *ordered latent spaces*. Initial results on large image datasets demonstrate that our method outperforms existing methods with further promising applications to variable rate neural compression, progressive signal generation and fast image samplers using diffusion models.
- Studying new research directions that involve in-context learning and fairness in generative AI (e.g., with image generation models or LLMs) and techniques for physics-inspired generative modeling (e.g., Schrödinger's bridge).

**Massachusetts Institute of Technology** — Cambridge, MA
*Graduate Research Assistant - Masters* — September 2019 – May 2022

- Developed customizable lossless and lossy data compression architectures by amalgamating techniques from deep generative modeling, graphical models and statistical inference.
- Trained generative image priors optimized for density estimation based on the deep generalized convolutional sum-product network (DGCSPN) architecture.
- Implemented a novel universal compression algorithm that performs *source-agnostic encoding* using a lightweight LDPC code and an adaptable decoder. The decoder performs belief-propagation between the DGSCPN network and a traditional LDPC graph by adapting to different source statistics by leveraging plug-and-play priors to compute the marginal distribution over *all* pixels in a *single pass* of the DGCSPN.
- Demonstrated that our algorithm operating in lossless model achieves a 1.7× gain in compression rate over JBIG2 on the binarized MNIST dataset, and a 1.4× gain in compression rate on FashionMNIST and grayscale CIFAR-10 as compared to widely used universal compressors.
- Lossy experiments demonstrate that our model outperforms JPEG with a 1.5× decrease in distortion at mid to high rates on the CIFAR-10 dataset and achieves competitive results with WebP.

**University of Illinois at Urbana-Champaign** — Champaign, IL
*Undergraduate Researcher* — May 2017 – May 2019

- Developed an efficient and novel scheme for video frame interpolation based on the Local All-Pass (LAP) algorithm for optical flow estimation.
- Optimized the algorithm to perform efficient and fast motion estimates for accurate temporal interpolation resulting in a runtime of under 4 seconds on a CPU.
- Performed qualitative and quantitative assessments on real-world video sequences. Results demonstrate that our algorithm is competitive in terms of MSE and produces perceptually pleasing interpolated videos despite having 100× fewer parameters than state-of-the-art neural networks.

## Work Experience

**Google Research** Cambridge, MA
*PhD Student Researcher* September 2022 – April 2023

- Developed new architectures for video compression using transformers to reduce computational complexity and improve rate-distortion (R-D) performance over existing methods.
- Implemented novel extensions to the state-of-the-art video compression transformer (VCT) by using frame-grouping, thereby achieving 2× improvements in inference speed and nearly 2× reduction in bitrate.
- Studied new techniques for training large video compression transformers such as efficient optimization using stop gradients, adaptable encoding using I/B/P-frame conditioning and flow-based entropy model conditioning.
- Performed ablations with various multi-frame encoding/decoding architectures such as convolutional architectures, spatio-temporal transformer architectures and frame aggregation techniques.
- Results show that a dual-feature concatenation architecture with a convolutional decoder achieves a 2× rate and speed improvement over the original VCT.

**Meta AI** Menlo Park, CA
*Research Intern* May 2022 – September 2022

- Implemented a new generative model for singing voice conversion that leverages pre-trained ASR fine-tuned self-supervised features and a parallel bank of transposed convolutions (PBTC) for $f_0$ modeling.
- Developed a technique for modeling the rich harmonic structure in singing voices by utilizing a parallel bank of strided and dilated transposed convolutions.
- Performed exhaustive ablation studies to demonstrate that singing voice synthesis and conversion is bolstered through the use of ASR fine-tuned Wav2Vec 2.0 features.
- Through various MOS studies, demonstrated that a PBTC-based architecture that uses ASR fine-tuned Wav2Vec 2.0 features along with $f_0$ shifting during inference leads to the best results.

**Meta AI** Menlo Park, CA
*Research Intern* June 2021 – November 2021

- Developed a variable bitrate neural speech codec with a hierarchical VQ-VAE (HVQVAE) backbone that operates at bitrates between 3.2 and 12.8 kbps with adjustable computation complexity. The HVQVAE is trained with a efficient masking loss to accommodate packet losses of up to 120 ms.
- In low-complexity mode, cepstral, periodicity and $f_0$ features are encoded into a bitstream using a VQ-VAE and then decoding the same using a parametric Griffin-Lim vocoder. In high-complexity mode, binned-spectrogram and $f_0$ features are encoded and decoding is done with a powerful WaveRNN decoder.
- Performed MOS studies on the DNS dataset to compare our codec against commonly used codecs like Opus and Codec2, and against neural codecs such as Lyra.
- Results demonstrated that the high-complexity decoder outperformed Opus by 0.2 in MOS while operating at 3.2 kbps, almost half of Opus' bitrate while the low-complexity decoder is competitive with Opus.
- Consistently outperformed the commonly used Codec2 encoder for speech compression, and the high-complexity decoder operating at 3.2 kbps outperformed Lyra in terms of MOS.

**Mitsubishi Electric Research Laboratories** Cambridge, MA
*Research Intern* April 2019 – May 2019

- Investigated uncertainty characterization techniques for adversarial attack detection on ASR by devising robust audio attacks on the Mozilla DeepSpeech ASR engine in Tensorflow.
- Created adversarial audio attacks robust to noise reduction and auditory masking by implementing a differentiable spectral subtraction algorithm that leverages straight-through gradient estimates.
- Leveraged dropout to estimate the uncertainty distribution of a transcription from the "mean" transcription by using the string edit distance as a metric.
- Experiments against various baselines demonstrated that an SVM-based classifier using the first four moments of the estimated uncertainty distribution as features resulted in the highest detection accuracy with algorithm runtimes under 1 second.
- The robustness of the algorithm is exemplified by detection accuracies of 89%, 92% and 94% on noise robust, auditory masked and urban sound adversarial attacks respectively.

## Publications and Patents

**(Accepted at NeurIPS 2023)** T. Jayashankar, G. C. Lee, A. Lancho, A. Weiss, Y. Polyanskiy, and G. W. Wornell, "Score-based Source Separation with Applications to Digital Communication Signals," *arXiv preprint arXiv:2306.14411*, 2023

T. Jayashankar, J. Wu, L. Sari, D. Kant, V. Manohar, and Q. He, "Self-supervised Representations for Singing Voice Conversion," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5

T. Jayashankar, "Image Compression using Sum-Product Networks," Master's thesis, Massachusetts Institute of Technology, 2022

T. Jayashankar, T. Koehler, K. Kalgaonkar, Z. Xiu, J. Wu, J. Lin, P. Agrawal, and Q. He, "Architecture for Variable Bitrate Neural Speech Codec with Configurable Computation Complexity," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 861–865

J. Le Roux, T. Jayashankar, and P. Moulin, "System and Method for Detecting Adversarial Attacks," Oct. 14 2021, US Patent App. 16/844,826

T. Jayashankar, J. Le Roux, and P. Moulin, "Detecting Audio Attacks on ASR Systems with Dropout Uncertainty," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, vol. 2020-October, pp. 4671–4675, 2020

T. Jayashankar, P. Moulin, T. Blu, and C. Gilliam, "LAP-Based Video Frame Interpolation," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 4195–4199

## Teaching Experience

**6.438: Algorithms for Inference**
*Graduate Teaching Assistant, Fall 2021*

- Full time teaching assistant for the graduate level course on statistical inference algorithms. Topics include belief-propagation, approximate inference, variational inference and Markov chain Monte-Carlo (MCMC).

**ECE 210 Honors**
*Undergraduate Teaching Assistant, Spring 2018 – Fall 2018*

- Taught students introductory programming skills in MATLAB and Python. Specifically developed homework and projects geared to help students understand the computational aspects of signal processing.

## Leadership & Activities

**MIT EECS Graduate Student Association**
*President 2021 – 2022, VP of Publicity 2020 – 2021*

**Tau Beta Pi, Illinois Alpha Chapter**
*President 2018 – 2019, Corresponding Secretary 2017 – 2018*

**IEEE Eta Kappu Nu (HKN)**
*Member*

## Technical Skills

Python, PyTorch, Tensorflow, Jax, Docker, Git, C++, C, MATLAB, Linux

## Fellowships and Awards

- Fano Fellowship (MIT, Spring 2020)
- Hewlett Packard 2 Fellowship (MIT, Fall 2019)
- University Honors: Bronze Tablet (UIUC, May 2019)
- Edward C. Jordan Award (UIUC, 2019)
- Jules D. Falzer Scholarship (UIUC, 2017 – 2018)