MIMO Channel Estimation using Score-Based Generative Models

Asad Aali, Marius Arvinte, Jon Tamir– asad.aali@utexas.edu
Electrical and Computer Engineering, University of Texas at Austin

Introduction

- Channel Estimation is a critical task in multiple-input multiple-output (MIMO) communications
- Recovering accurate, high-dimensional channel state information (CSI) using reduced pilot (P) overhead has become a major open research problem
- Estimating accurate CSI with data-driven methods is important for future communication systems that integrate AI in physical layer processing

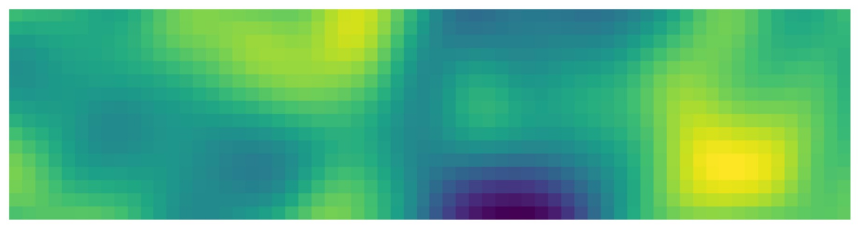
Goal
Develop robust, data-driven, deep learning-based MIMO channel estimation algorithms for high-dimensional communication scenarios

Wireless System Theory

- MIMO forward model: $\mathbf{Y} = \mathbf{HP} + \mathbf{N}.$

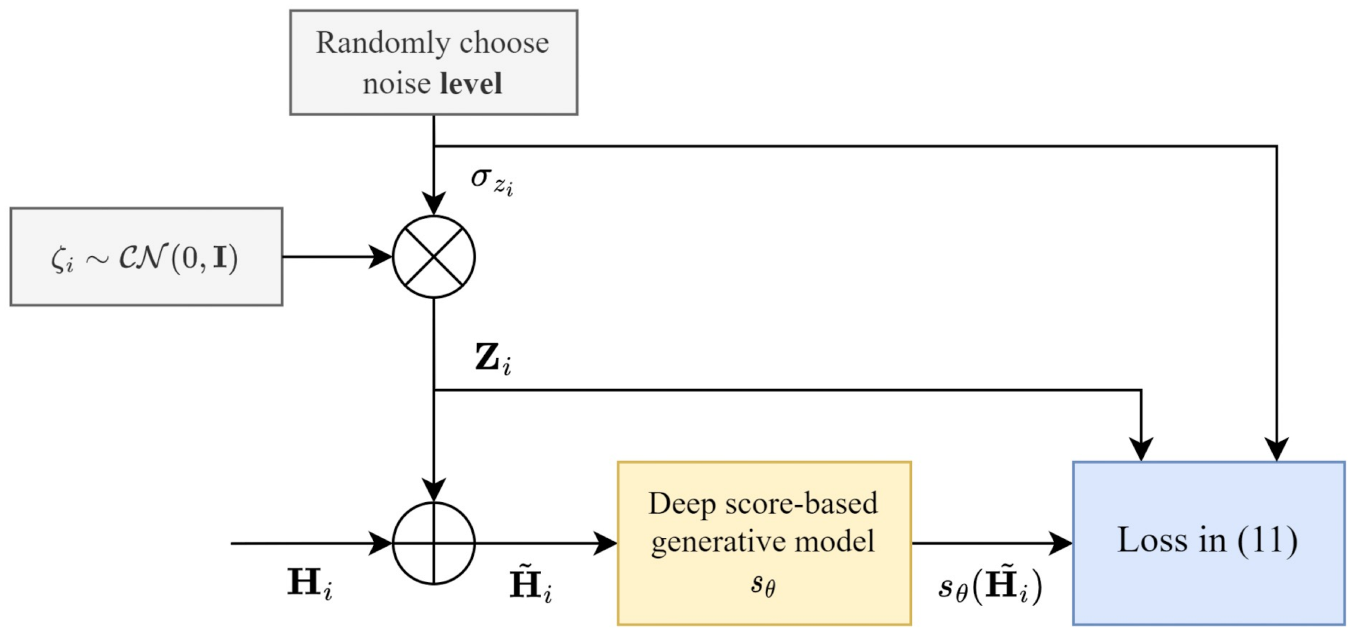
$\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$	Channel state information matrix
$\mathbf{p}_i \in \mathbb{C}^{N_t}$	Pilot symbol
$\sigma^2_{\text{pilot}}\mathbf{I}$	Complex Additive White Gaussian Noise

- A narrowband, point-to-point MIMO communication scenario between a transmitter and receiver
- Channel estimation requires estimating the channel state information matrix H, using the received pilot matrix Y, while having knowledge of the transmitted pilot matrix P



CDL-D channel distribution p(H)

Score Model Training



- A database of known channels is used to train a score-based generative model in an unsupervised manner
- A denoising score-matching framework learns the score (the gradient of log-prior distribution)

$$\mathcal{L}_{\text{score}}(\theta) = \mathbb{E}_{\mathbf{H}_i \sim p_H, \mathbf{Z}_j \sim p_{Z_j}} \left[\sigma^2_{z_j} \left\| s_\theta(\mathbf{H}_i + \mathbf{Z}_j) + \frac{\mathbf{Z}_j}{\sigma^2_{z_j}} \right\|_2^2 \right].$$

Proposed Method

Score Models for channel generation – sample from prior H ~ p(H)

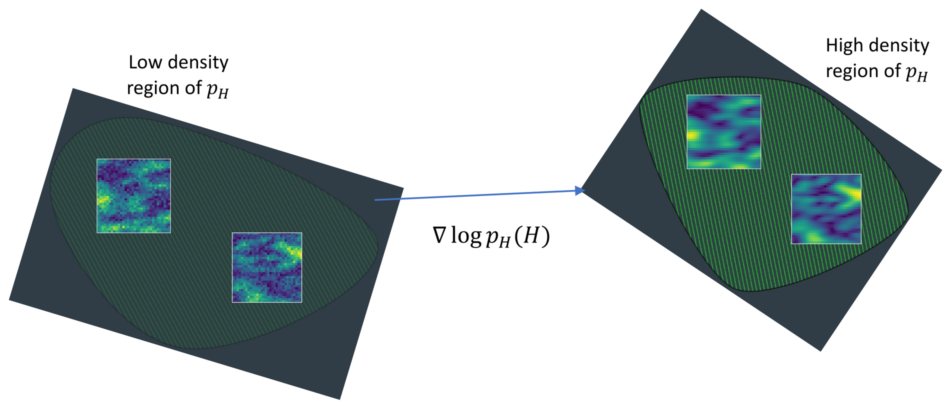
Annealed Langevin Dynamics

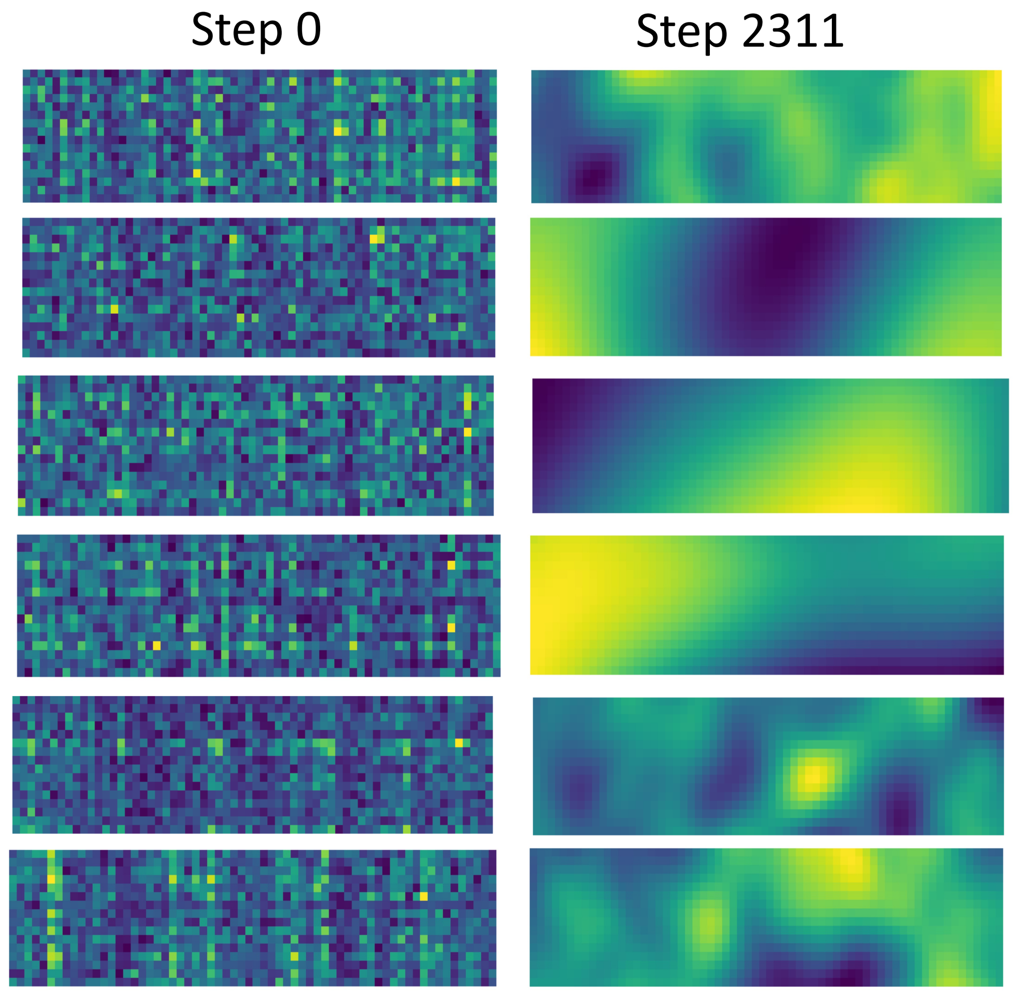
$$\mathbf{H}_{t+1} \leftarrow \mathbf{H}_t + \alpha_t \cdot \nabla \log p_H(\mathbf{H}_t) + \beta_t \cdot \zeta_t.$$

- $\alpha_t \cdot \psi_H(\mathbf{H}_t)$ increases the likelihood of the current sample.
- $\beta_t \cdot \zeta_t$ represents a perturbation to the above process.

Let p_H denote the distribution of MIMO (CDL-D) channels for a stochastic environment. The score of p_H at H is defined as:

$$\psi_H(\mathbf{H}) = \nabla \log p_H(\mathbf{H}),$$





Score Models for channel reconstruction – sample from posterior H ~ p(H|Y)

- Use the learned score-based model in conjunction with the received pilots to iteratively update the channel estimate and perform posterior sampling

Posterior Sampling Update

$$\mathbf{H} \leftarrow \mathbf{H} + \alpha \cdot \psi_{H|Y}(\mathbf{H}|\mathbf{Y}) + \beta \cdot \zeta,$$

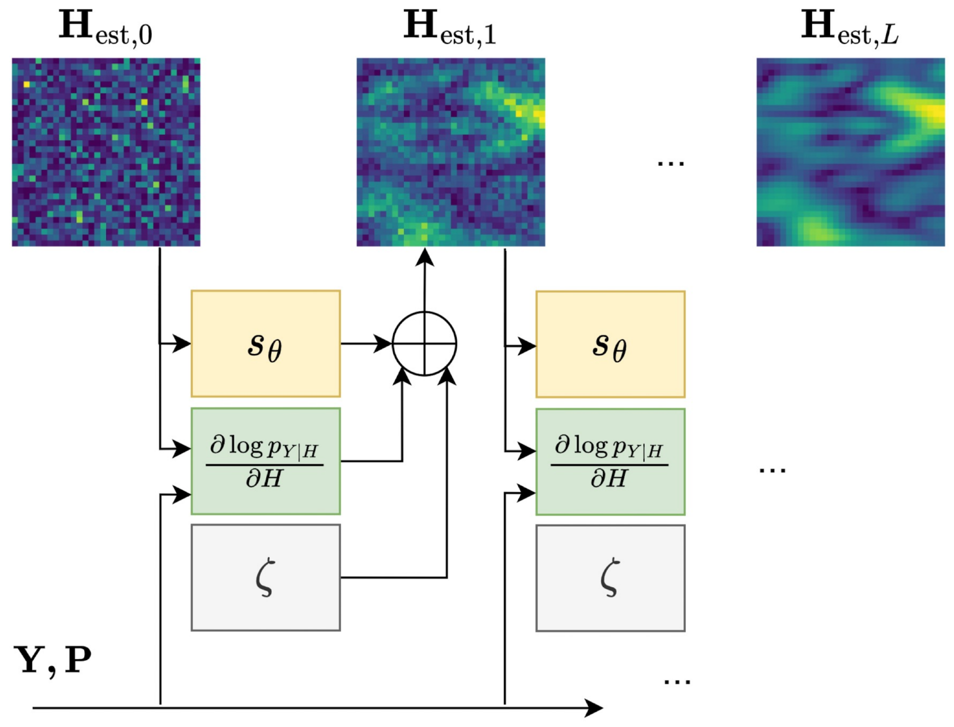
Using Bayes Rule for $p(H|Y)$ $\frac{p_{Y|H}(\mathbf{Y}|\mathbf{H}) \cdot p_H(\mathbf{H})}{p_Y(\mathbf{Y})}$

$$\log p_{H|Y}(\mathbf{H}|\mathbf{Y}) = \log p_{Y|H}(\mathbf{Y}|\mathbf{H}) + \log p_H(\mathbf{H}) - \log p_Y(\mathbf{Y}).$$

Taking Gradient on both sides

$$\mathbf{H} \leftarrow \mathbf{H} + \alpha \cdot \psi_{Y|H}(\mathbf{Y}|\mathbf{H}) + \alpha \cdot \psi_H(\mathbf{H}) + \beta \cdot \zeta.$$

$$\nabla \log p_{Y|H}(\mathbf{Y}|\mathbf{H}_{\text{est},i}) = \frac{(\mathbf{H}_{\text{est},i}\mathbf{P} - \mathbf{Y})\mathbf{P}^{\text{H}}}{\sigma^2_{\text{pilot}} + 2\beta \cdot \alpha_i \cdot \sigma^2_{z_i}},$$



$$\mathbf{H}_{\text{est},i+1} = \mathbf{H}_{\text{est},i} + \alpha_i \cdot (\nabla \log p_{Y|H}(\mathbf{Y}|\mathbf{H}_{\text{est},i}) + \nabla \log p_H(\mathbf{H}_{\text{est},i})) + \sqrt{2\beta \cdot \alpha_i} \cdot \sigma_{z_i} \cdot \zeta,$$

Algorithm 1 MIMO Channel Estimation with Score-Based Generative Models.

Inputs: Pilot matrix $\mathbf{P}$, received pilots $\mathbf{Y}$, pretrained score-based model s_θ, received noise power σ^2_{pilot}, inference noise levels $\sigma^2_{z_i}$ (same as what s_θ was trained with), hyper-parameters α_0, β and $r < 1$.

Generate random initial estimate: $\mathbf{H}_{\text{est}} \sim \mathcal{CN}(0, \mathbf{I})$

for $i = 1, \ldots, L$

Set annealed noise level $\sigma \leftarrow \sigma_{z_i}$.

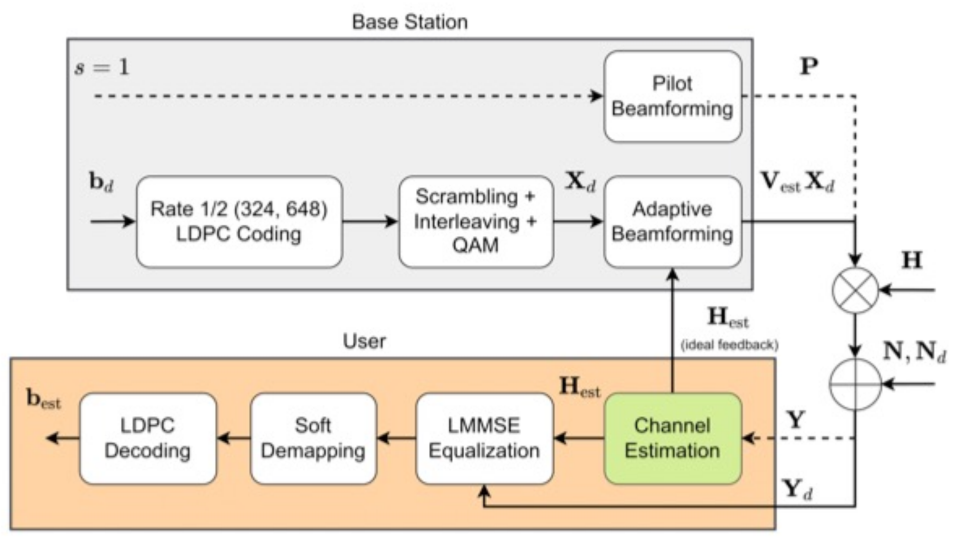
for $m = 1, \ldots, M$

Generate annealing noise $\zeta \sim \mathcal{CN}(0, \mathbf{I})$.

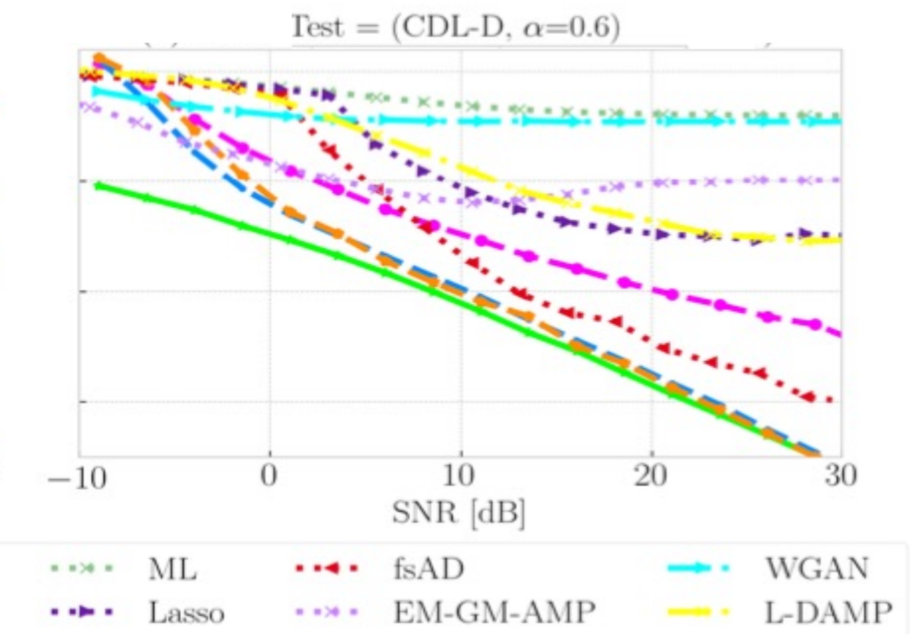
$\mathbf{H}_{\text{est}} \leftarrow \mathbf{H}_{\text{est}} + \alpha_0 \cdot r^i \cdot \frac{(\mathbf{H}_{\text{est}}\mathbf{P} - \mathbf{Y})\mathbf{P}^{\text{H}}}{\sigma^2_{\text{pilot}} + \sigma^2} + \alpha_0 \cdot r^i \cdot s_\theta(\mathbf{H}_{\text{est}}) + \sqrt{2\beta \cdot \alpha_0 \cdot r^i} \cdot \sigma \cdot \zeta.$

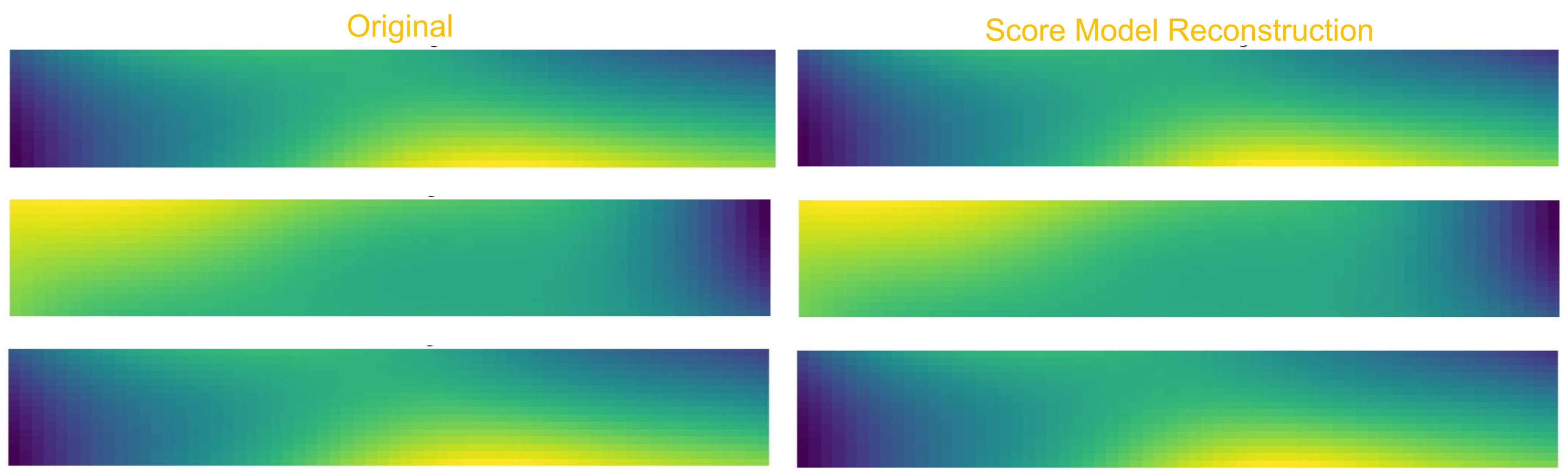
Output: Estimated channel matrix $\mathbf{H}_{\text{est}}$.

$$\mathbf{Y} = \mathbf{HP} + \mathbf{N}.$$



Results





Discussion and Conclusion

- Introduced an unsupervised, probabilistic approach for MIMO channel estimation using a reduced number of pilot symbols
- Our results on simulated channels show that performance is favorable in-distribution, as well as in out-of-distribution settings

Next Steps:
1. Generic score-model training/inference code
2. Learning score models from noisy data
3. Accelerating inference with score models
4. Training and testing score models with real measurements
5. Score models beyond 2D channels

References

1. Arvinte, M., & Tamir, J. I. (2022). MIMO Channel Estimation using Score-Based Generative.
2. Y. Song and S. Ermon, Generative Modeling by Estimating Gradients of the Data Distribution.
3. Jalal, A, et al. Robust Compressed Sensing MRI with Deep Generative Priors.