Generative vs. Predictive Models in mMIMO Channel Prediction

USC Viterbi School of Engineering
Ming Hsieh Department of Electrical and Computer Engineering

Ju-Hyung Lee, Joohan Lee, Andreas F. Molisch

Idea

Keyword: Generative model, massive MIMO, cross-antenna channel prediction, Diffusion models, DDPM, VQ-VAE.

- This work introduces a **Vector Quantization-based generative autoencoder (VQ-VAE)** for robust mMIMO cross-antenna channel prediction.
- **Generative models outperform Predictive models** for mMIMO cross-antenna channel prediction, particularly in noisy conditions.
- A **complexity analysis** of autoencoder-based models and diffusion models shows a trade-off between accuracy and computational efficiency.

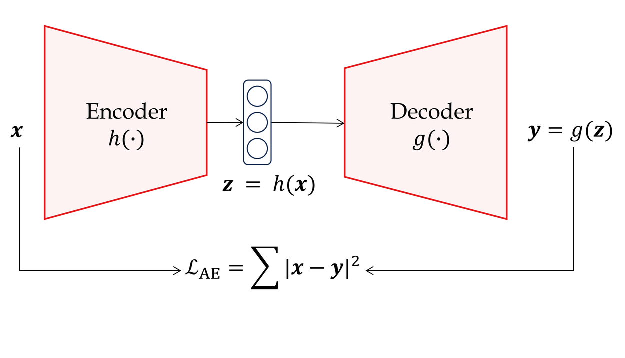
Contribution

Our key contributions are summarized as follows:

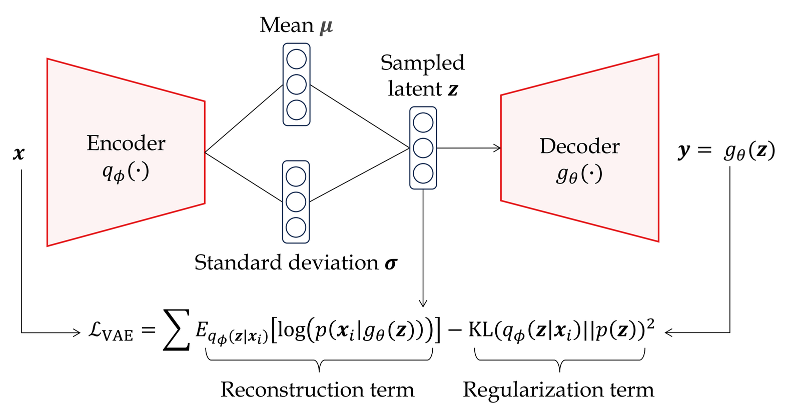
- We demonstrate that AE-based generative models outperform AE-based predictive models for mMIMO cross-antenna channel prediction, especially under noisy conditions.
- The proposed VQ-VAE shows robustness in noisy channel prediction, achieving up to 15 [dB] NMSE gains over standard AEs and about 9 [dB] over VAEs in noisy conditions.
- We highlight VQ-VAE's lower computational complexity compared to diffusion models.

Background

- Generative models capture underlying data patterns for generation, while predictive models focus on direct outcome prediction without understanding of the data distribution.
- AEs use a deterministic approach to compress and reconstruct data, lacking data generation capabilities. VAEs introduce probabilistic latent spaces, allowing them to generate novel outputs that share patterns with the training data.
- To address the challenges in mMIMO channel prediction, this work focuses on developing noise-robust models that maintain accuracy while balancing computational complexity.

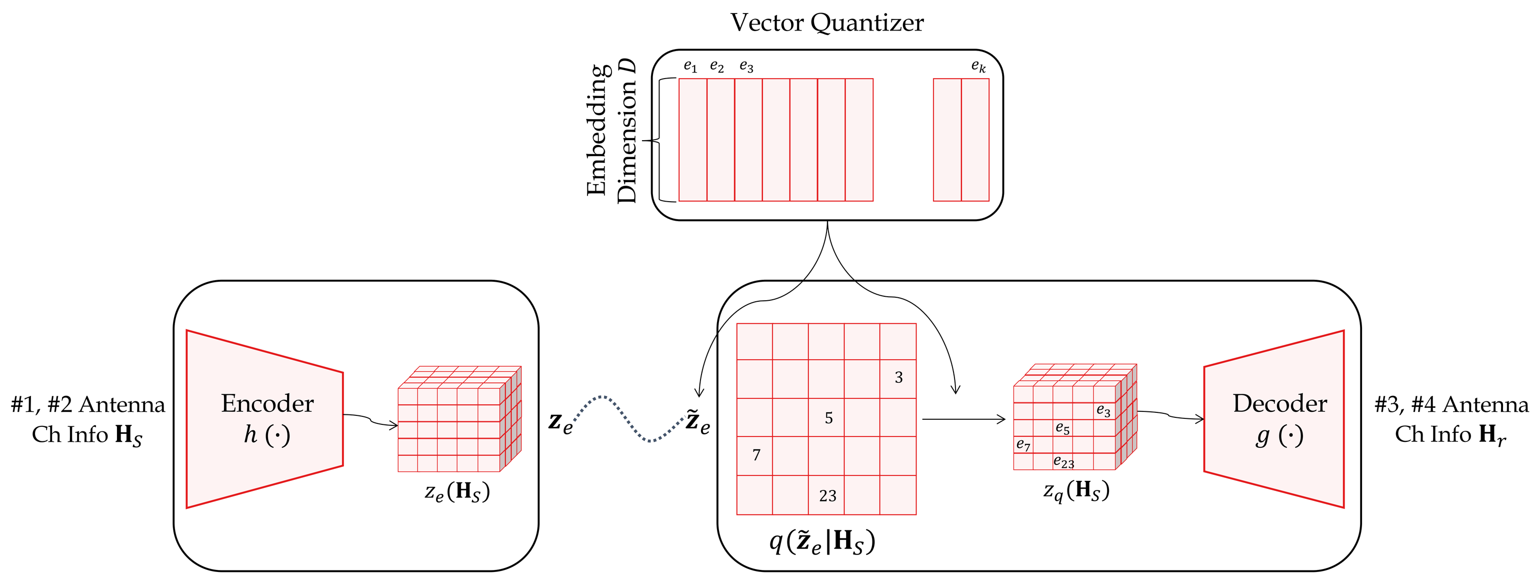


AE



VAE

VQ-VAE for mMIMO Channel Prediction



VQ-VAE based mMIMO channel prediction

Task & Approach

- Accurate estimation of the wireless channel $\mathbf{H} \in \mathbb{C}^{M \times N}$, where M and N represent the number of transmitting and receiving antennas, is essential for reliable communication. However, imperfect channel estimation can degrade signal quality and reduce data rates.
- In the uplink scenario, each UE transmits M pilot symbols, and the BS estimates the channel coefficients $\mathbf{H}_k \in \mathbb{C}^{M \times 1}$ for each antenna k. The total feedback overhead scales as $\mathcal{O}(M \times K)$, which becomes a significant challenge as M and K grow.
- To reduce this overhead, we propose a *Cross-Antenna* prediction strategy, where a neural network predicts the channel responses of some antennas using the information from a subset of others. Specifically, given measurements from M_s antennas, $\mathbf{H}_s$, the network predicts the channel coefficients for the remaining M_r antennas, $\mathbf{H}_r$.
- The encoder compresses channel information from a subset of antennas, $\mathbf{H}_s$, into a latent vector z_e using convolutional layers. During transmission, z_e may be affected by noise, producing $\tilde{z}_e$. The vector quantizer maps $\tilde{z}_e$ to the nearest codebook entry e_i, and the quantized vector z_q is passed to the decoder, which reconstructs the channel information. This approach reduces feedback overhead and improves noise robustness.

Evaluation & Training

- Reconstruction accuracy is measured using normalized mean squared error (NMSE).
- AE, VAE, and VQ-VAE are trained with MSE loss to minimize reconstruction error.
- VQ-VAE additionally applies a VQ loss, aligning the encoder's latent output $\boldsymbol{z}_e$ to its closest codebook vector z_q to ensure consistent reconstruction. The final VQ-VAE loss combines MSE and VQ losses:

$$\mathcal{L}_{\text{VQVAE}} = \mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{VQLoss}} \qquad (1)$$

Dataset

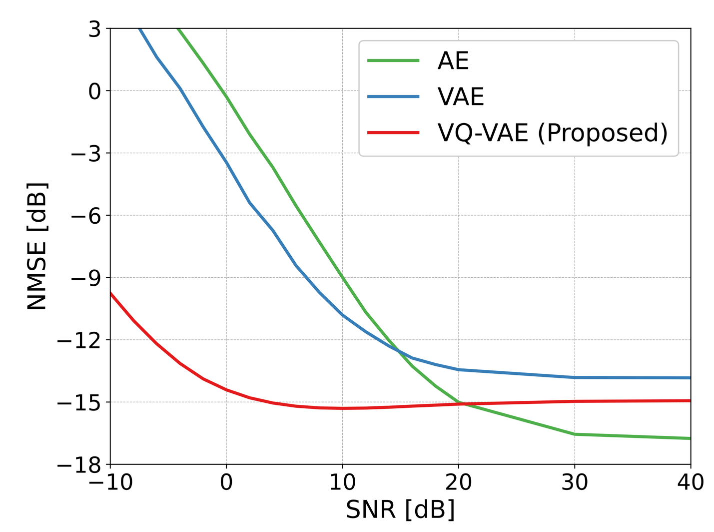
- **3GPP CDL-Family Channels**: We train on 9000 samples from the CDL-C model and evaluate on 1000 samples from CDL-A to D channels, following 3GPP CDL-family models (TR 38.901).

Experiments

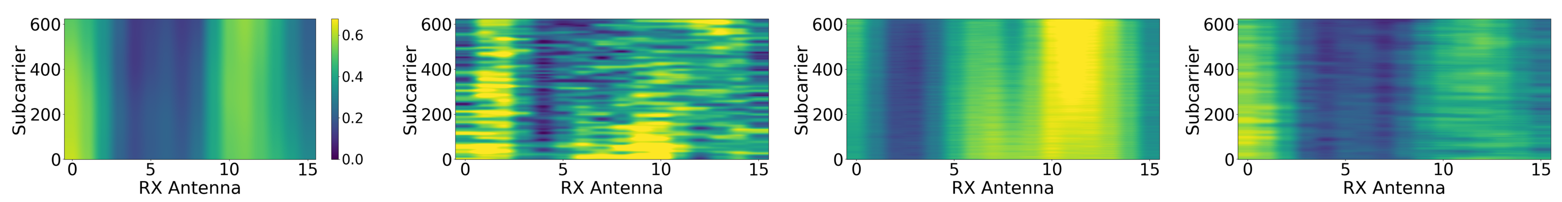
Baseline

- **AE**: Predictive model using Conv layers for latent space compression and reconstruction.
- **VAE**: Generative model with latent space regularization via mean and log-variance layers.
- **Diffusion (DDPM)**: Generative model with a UNet architecture, adding and removing Gaussian noise iteratively for data generation.
- **VQ-VAE (Proposed)**: Generative model using vector quantization to enhance reconstruction by mapping latent vectors to codebook entries.

Comparison



Comparison



(a) Ground-Truth (b) AE ($\gamma = 0$ [dB]) (c) VAE ($\gamma = 0$ [dB]) (d) VQ-VAE ($\gamma = 0$ [dB])

Complexity

	Inference time [ms]	Memory [MB]	Training time [s]	Memory [MB]
AE	3.374	62.67	514	216
VAE	4.442	108.59	840	469
VQ-VAE	5.340	175.20	1576	1099
Diffusion	122.624	1385.41	30491	36135

Conclusions

- The proposed VQ-VAE model achieved up to 15 [dB] NMSE gains over standard AEs and 9 [dB] over VAEs, demonstrating the model's effectiveness in noisy wireless environments.
- Future research could explore the application of generative models for other tasks, such as feedback compression in mMIMO systems, focusing on optimizing architectures to balance prediction accuracy and computational efficiency.