Few-Shot Domain Adaptation for End-to-End Communication

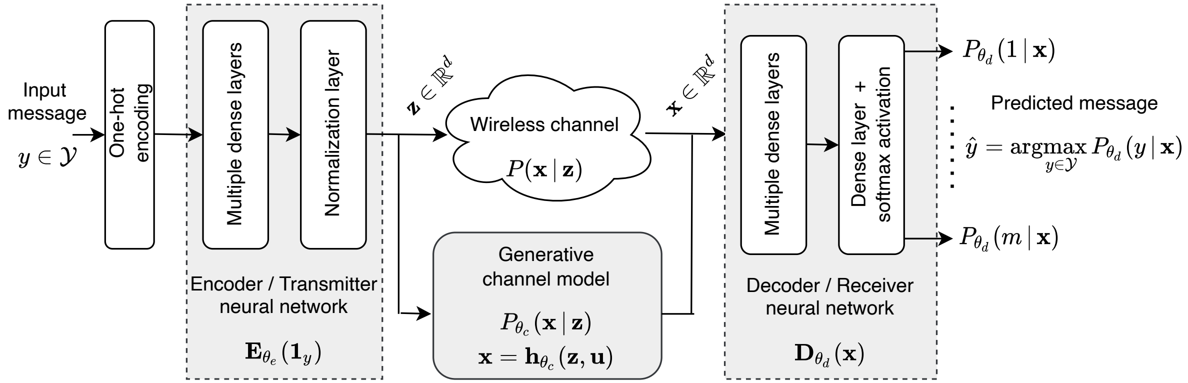
Jayaram Raghuram[1], Yijing Zeng[1], Dolores García[2], Rafael Ruiz[2], Somesh Jha[1,3], Joerg Widmer[2], Suman Banerjee[1]

[1]UW-Madison, [2]IMDEA Networks, [3]XaiPient

Spotlight

End-to-End Learning for Communication

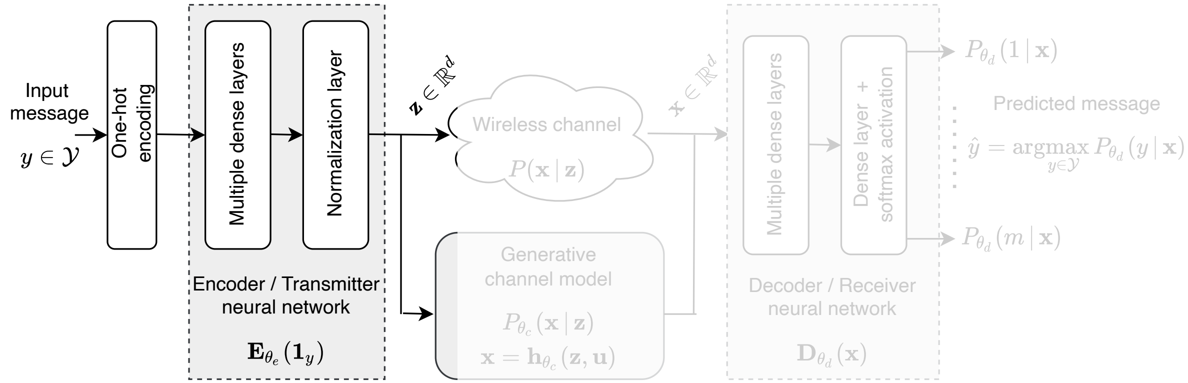
- We want to transmit one of m messages and decode it accurately at the receiver
- Encoder/Transmitter NN: learns a custom encoding (modulation) most suitable for the channel conditions
- Decoder/Receiver NN: learns to predict the transmitted message from the channel output
- Channel model: a generative model of the channel condition density

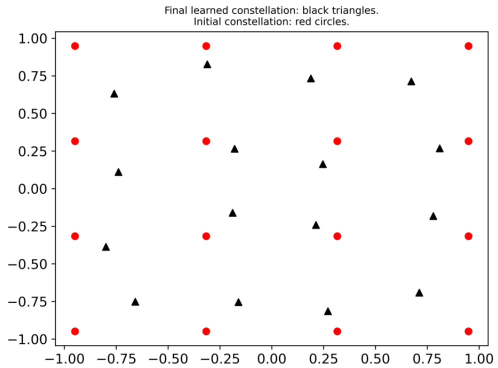


- Message y is equivalent to a class label
- Decoder is a classifier
- Channel output x is the feature vector

Encoder/Transmitter NN

- Input is a one-hot coded message: all zeros except 1 at the message index
- Example: For m = 4 messages, the message 2 is encoded as [0, 1, 0, 0]
- Output $\boldsymbol{z}$ is a custom encoding or modulation, a vector of dimension $\boldsymbol{d}$
- Usually, $d = 2$. Corresponds to the In-phase and Quadrature-phase (IQ) components of a transmitted RF signal

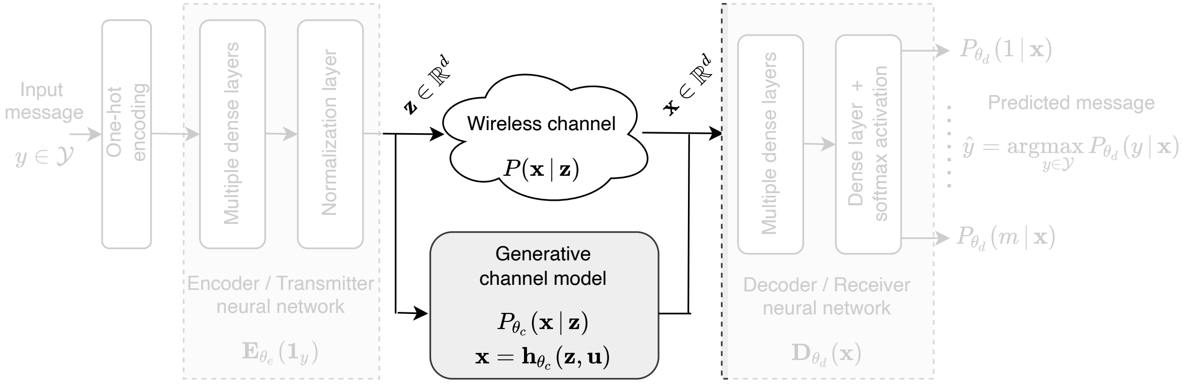


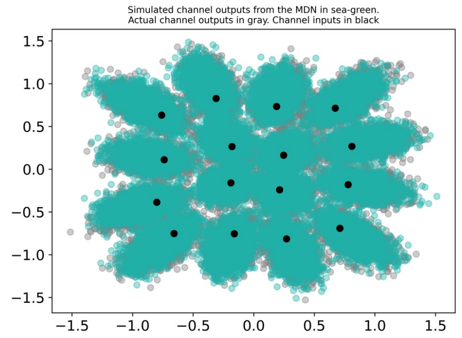


Red circles: standard 16-QAM
Black triangles: autoencoder learned

Generative Channel Model

- Learns the conditional probability density $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{z})$ of channel output given channel input
- We model the channel using a Gaussian Mixture density network (MDN)
- $P(\boldsymbol{x} \mid \boldsymbol{z})$ is a Gaussian mixture for each unique $\boldsymbol{z}$
- There are m Gaussian mixtures, one per message or corresponding symbol $\boldsymbol{z}$
- MDN is a neural network that predicts the parameters of the m Gaussian mixtures





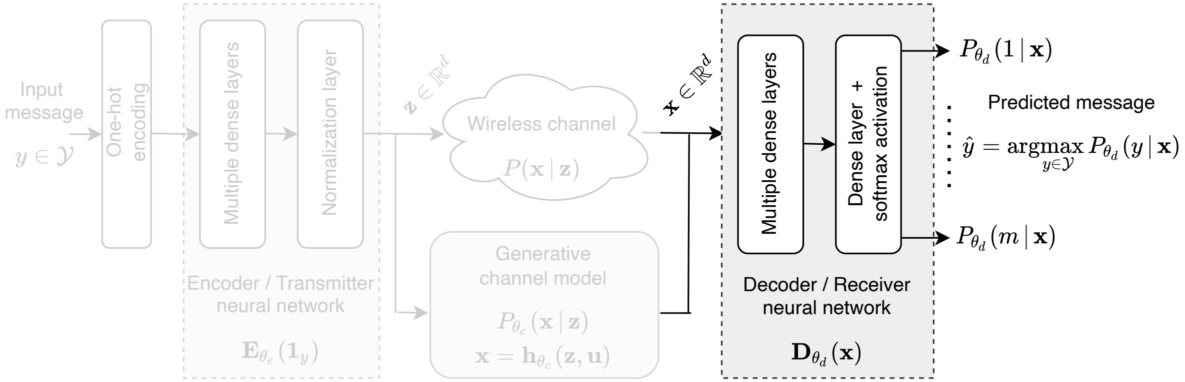
Decoder/Receiver NN

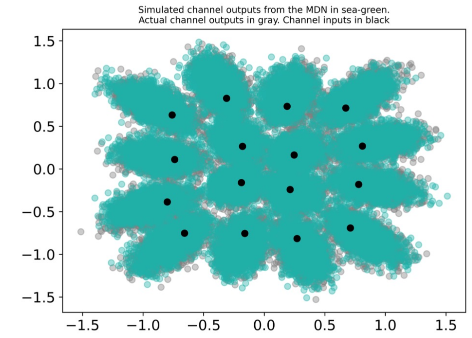
- Classifier network that takes $\boldsymbol{x}$ as input and predicts the probability of the m messages
- Decoder output: predicted probability of each message class $P_{\boldsymbol{\theta}_d}(y \mid \boldsymbol{x})$
- Message with the highest probability is the decoded message

$$\hat{y}(\mathbf{x}) = \operatorname{argmax}_{y \in \mathcal{Y}} P_{\boldsymbol{\theta}_d}(y \mid \mathbf{x})$$

- Symbol error rate (SER) is the expected error in the decoder's prediction

$$\mathbb{E}_{(\mathbf{x},y)}[\mathbb{1}(\hat{y}(\mathbf{x}) \neq y)]$$





Need for Few-Shot Domain Adaptation

- Channel is dynamic: distribution can change frequently (e.g. a wireless link)
- Autoencoder performance can degrade under a changing channel distribution

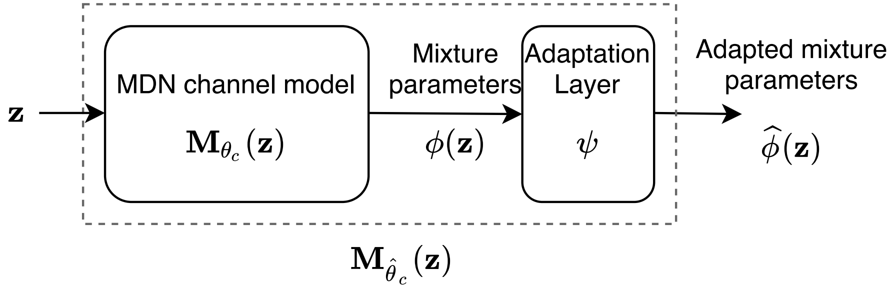
- Retraining the channel model and autoencoder frequently is not practical:
  - ❑ Frequent training data collection lowers the throughput
  - ❑ Time consuming and often hard to update the transmitter/encoder side frequently

Can we adapt the channel model and decoder using only a small set of samples from the target channel distribution?

Proposed Method

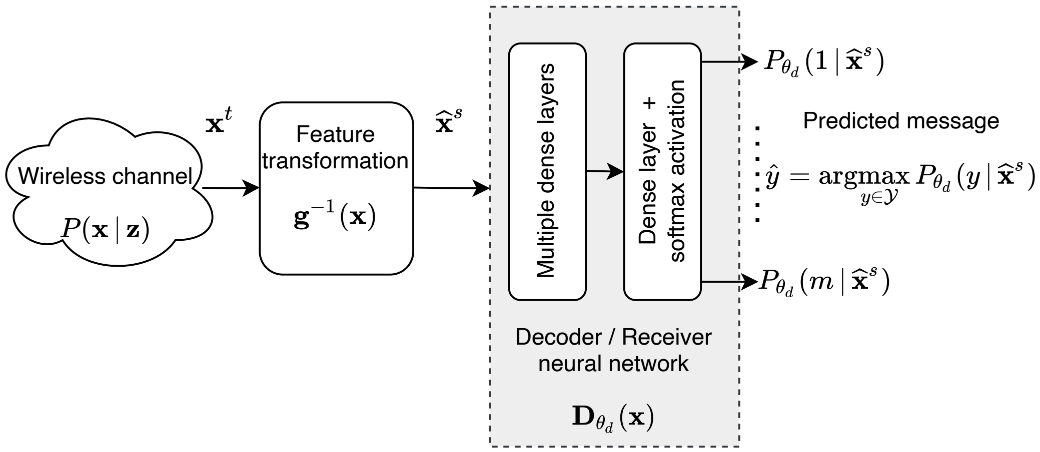
MDN (channel) Adaptation

- An adaptation layer with much smaller number of parameters compared to the MDN is used
- A closed-form Kullback-Leibler divergence is used as regularization in the small-sample setting



Decoder Adaptation

- A feature transformation that approximately maps decoder inputs from target distribution back to the source distribution
- Decoder and encoder networks remain unchanged



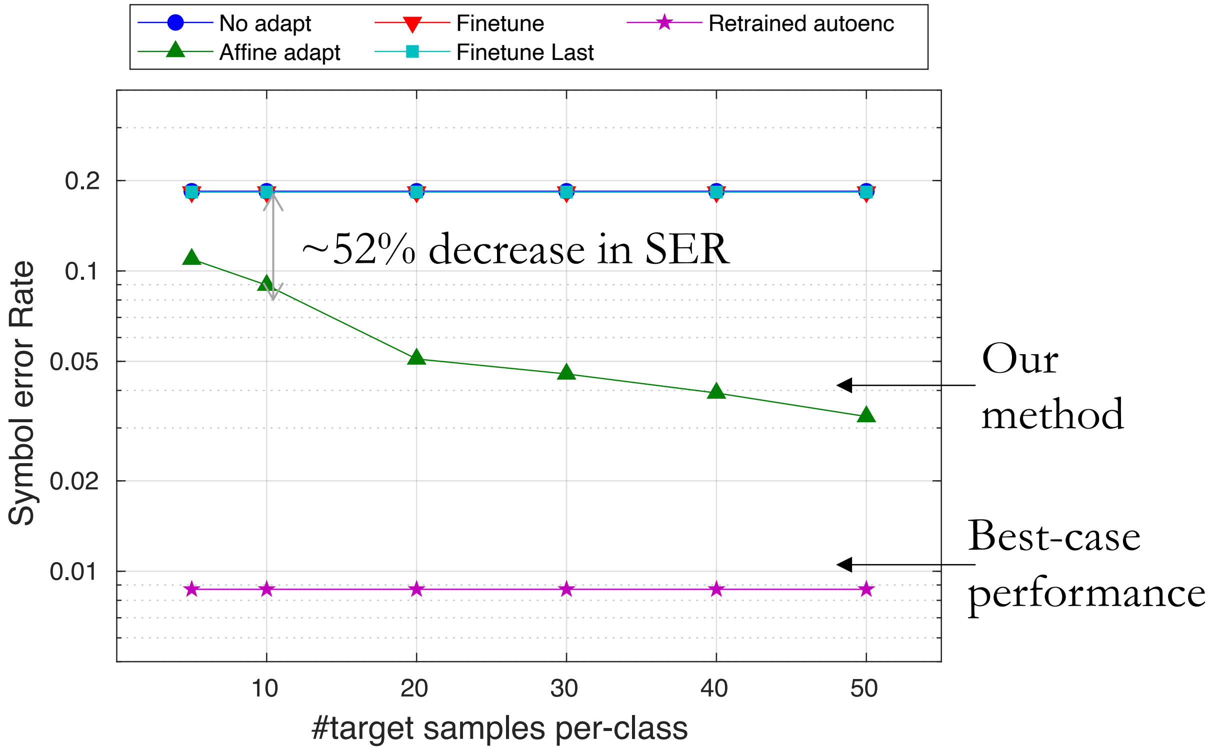
Experiments

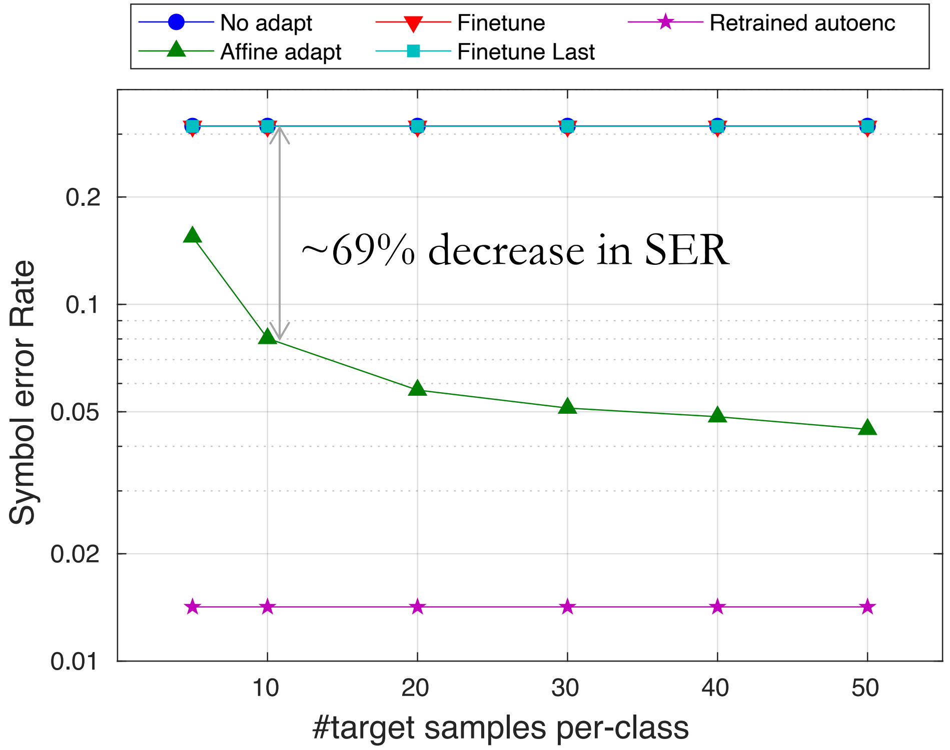
- We evaluate on several simulated but realistic distribution changes
  - ❑ AWGN → Ricean fading (different signal-to-noise ratios)
  - ❑ AWGN → Uniform fading
  - ❑ Ricean fading → Uniform fading
  - ❑ Uniform fading → Ricean fading
  - ❑ Random Gaussian mixture → Random Gaussian mixtures (50 datasets)

- Real experiments on a mmWave FPGA platform
  - ❑ Ultra-wide-band mm-wave transceiver, 60 GHz RF front-end antennas
  - ❑ Distribution changes: IQ (in-phase, quadrature-phase) imbalance-based distortion to the symbol constellation

Results - FPGA experiments

- Distribution changes introduced by IQ imbalance-based symbol distortion
- We varied the level of IQ imbalance. Results below for 25% and 30% distortion





Summary

We proposed a few-shot domain adaptation method for autoencoder-based e2e communication:

1) A sample- and parameter-efficient adaptation of the Gaussian MDN channel

2) An optimal feature transformation at the decoder that approximately maps the target-domain inputs to the source domain

- Paper: https://openreview.net/forum?id=4F1gvduDeL
- Code repo: https://github.com/jayaram-r/domain-adaptation-autoencoder

Thank you!