A Learning Approach to Cooperative Communication System Design

Yuxin Lu[⋆], Peng Cheng[†], Zhuo Chen[*], Wai Ho Mow[⋆] and Yonghui Li[†]

[⋆]The Hong Kong University of Science and Technology.
[†]The University of Sydney.
[*]Data 61, CSIRO.

The 45th IEEE ICASSP, 2020

Supported by the HK RGC (GRF 16233816)

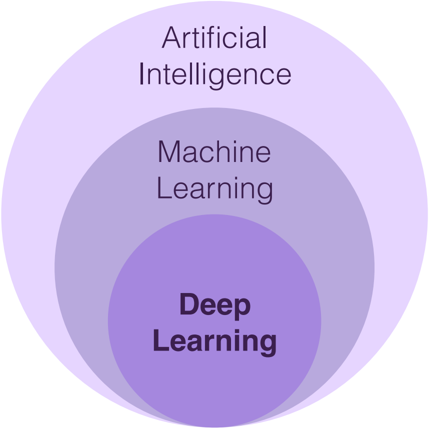
Outline

1. Background and Motivation
2. Relay-Assisted Cooperative Communication System
3. Learning the Cooperative System
4. Simulation Results
5. Conclusion

Outline

1. Background and Motivation
2. Relay-Assisted Cooperative Communication System
3. Learning the Cooperative System
4. Simulation Results
5. Conclusion

Deep Learning



- Deep learning (DL) is a branch of machine learning ⇒ Learn to make own decisions
- Structures algorithms in **layers** ⇒ Create an "artificial neural network"

Deep Learning in Communication

- Conventional communication system is optimized in a **block-wise** manner:
  source/channel coding, modulation, demodulation, source/channel decoding, equalization

Deep Learning in Communication

- Conventional communication system is optimized in a **block-wise** manner:
  source/channel coding, modulation, demodulation, source/channel decoding, equalization
- Deep learning techniques have been applied to replace certain blocks: channel coding/estimation

Deep Learning in Communication

- Conventional communication system is optimized in a **block-wise** manner:
  source/channel coding, modulation, demodulation, source/channel decoding, equalization
- Deep learning techniques have been applied to replace certain blocks: channel coding/estimation

Individualized component-wise approach might not optimize the overall system function!

Deep Learning in Communication

Can we optimize the communication system in a holistic manner?

Deep Learning in Communication

Can we optimize the communication system in a holistic manner?

- Joint design of the transmitter and receiver over the channel
- Expand the optimization space
- ...

Deep Learning in Communication

Can we optimize the communication system in a holistic manner?

- Joint design of the transmitter and receiver over the channel
- Expand the optimization space
- ...

Yes. Communication **Autoencoder**!

Deep Learning in Communication

Can we optimize the communication system in a holistic manner?

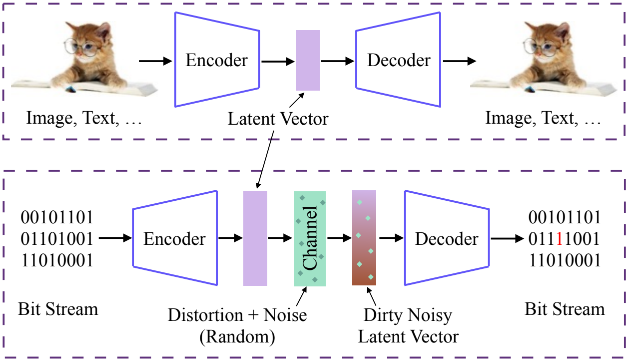
- Joint design of the transmitter and receiver over the channel
- Expand the optimization space
- ...

Yes. Communication **Autoencoder**!

- Transmitter and receiver are represented by neural networks (NNs)
- Promising results have been obtained

Autoencoder

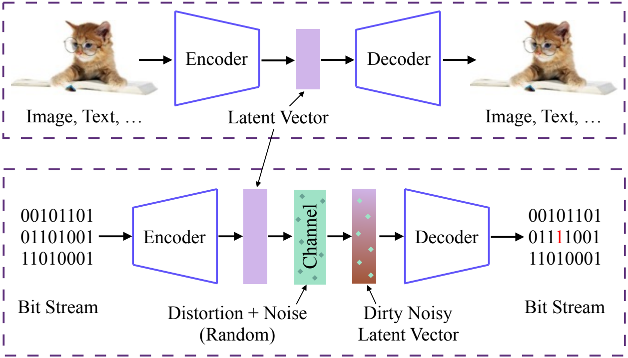
- General autoencoder (AE) learns data structure to compress (top)



General autoencoder (top) v.s. Communication autoencoder. Figure Credit: Zhao, Vuran, Guo and Scott

Autoencoder

- General autoencoder (AE) learns data structure to compress (top)



General autoencoder (top) v.s. Communication autoencoder. Figure Credit: Zhao, Vuran, Guo and Scott

- Communication AE learns the channel behavior to improve transmission accuracy

Autoencoder

Most existing applications are for point-to-point communications

More Complicated Scenarios

Can we design an AE to optimize more complicated communication scenarios?

More Complicated Scenarios

Can we design an AE to optimize more complicated communication scenarios?

Our focus: **Relay-assisted cooperative communication system**

Existing Works for AE+Relay

Constellation design for two-way relay networks[1]

⇒ Focused on constellation optimization. **No detection algorithm** was addressed

[1] T.Matsumine, T.Koike-Akino, and Y.Wang, "Deep learning-based constellation optimization for physical network coding in two-way relay networks," *arXiv preprint arXiv:1903.03713*, Mar. 2019.

Existing Works for AE+Relay

Constellation design for two-way relay networks[1]

⇒ Focused on constellation optimization. **No detection algorithm** was addressed

Our focus: **Joint optimization of the constellation and detection algorithm**

Start with a one-way relay network

[1] T.Matsumine, T.Koike-Akino, and Y.Wang, "Deep learning-based constellation optimization for physical network coding in two-way relay networks," *arXiv preprint arXiv:1903.03713*, Mar. 2019.

Outline

1 Background and Motivation

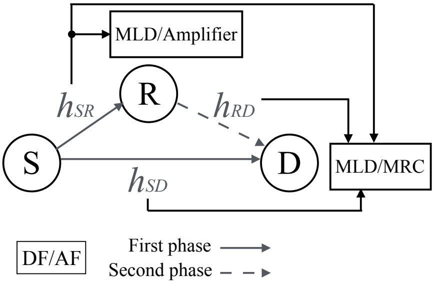
2 Relay-Assisted Cooperative Communication System

3 Learning the Cooperative System

4 Simulation Results

5 Conclusion

System Model



System model of a 3-node relay network.
Source (S), Relay (R), Destination (D)

- R: half-duplex
- Source message: $m_S \in \{1, 2, \cdots, 2^k\}$, encoded as $\mathbf{x}_S$ of length n
- k/n bits/independent channel uses

System Model

- First Phase:

$$\mathbf{y}_{SJ} = \sqrt{E_S}\mathbf{h}_{SJ}\mathbf{x}_S + \mathbf{n}_{SJ},\ J \in \{R, D\}, \tag{1}$$

E_S: average source transmit energy
$\mathbf{h}_{SJ}$: channel coefficient
$\mathbf{n}_{SJ}$: Gaussian noise vector $\mathcal{CN}(0, 2\sigma^2_{SJ}\mathbf{I})$

System Model

- First Phase:

$$\mathbf{y}_{SJ} = \sqrt{E_S}\mathbf{h}_{SJ}\mathbf{x}_S + \mathbf{n}_{SJ},\ J \in \{R, D\}, \tag{1}$$

E_S: average source transmit energy
$\mathbf{h}_{SJ}$: channel coefficient
$\mathbf{n}_{SJ}$: Gaussian noise vector $\mathcal{CN}(0, 2\sigma^2_{SJ}\mathbf{I})$

- Second Phase:

$$\mathbf{y}_{RD} = \sqrt{E_R}\mathbf{h}_{RD}\mathbf{x}_R + \mathbf{n}_{RD}, \tag{2}$$

E_R: average relay transmit energy
$\mathbf{h}_{RD}$: channel coefficient
$\mathbf{n}_{RD}$: Gaussian noise vector $\mathcal{CN}(0, 2\sigma^2_{RD}\mathbf{I})$

AF Relaying

- AF relay node:
  - Symbol-wise amplifying operation $x_R = \frac{y_{SR}}{\sqrt{P_S |h_{SR}|^2 + 2\sigma_{SR}^2}}$, $x_R \in \mathbf{x}_R$, $y_{SR} \in \mathbf{y}_{SR}$, $h_{SR} \in \mathbf{h}_{SR}$
  - Drawback: noise amplification $\Leftarrow \mathbf{y}_{SR} = \sqrt{E_S}\mathbf{h}_{SR}\mathbf{x}_S + \mathbf{n}_{SR}$

AF Relaying

- AF relay node:
  - Symbol-wise amplifying operation $x_R = \frac{y_{SR}}{\sqrt{P_S|h_{SR}|^2+2\sigma_{SR}^2}}$, $x_R \in \mathbf{x}_R$, $y_{SR} \in \mathbf{y}_{SR}$, $h_{SR} \in \mathbf{h}_{SR}$
  - Drawback: noise amplification $\Leftarrow \mathbf{y}_{SR} = \sqrt{E_S}\mathbf{h}_{SR}\mathbf{x}_S + \mathbf{n}_{SR}$
- Destination:
  - Maximal-ratio combining (MRC)
  - Optimal in the context of AF
  - High complexity: $\mathcal{O}(n \cdot 2^k)$ per block

DF Relaying

- DF relay node:
  - Maximum-likelihood decoding (MLD) $\mathbf{x}_R = \arg\min_{\mathbf{x}\in\mathcal{C}} \|\mathbf{y}_{SR} - \mathbf{h}_{SR}\sqrt{E_S}\mathbf{x}\|^2$, where $\mathcal{C}$ is code book, $|\mathcal{C}| = 2^k$.
  - Drawback: hard decision $\Rightarrow$ information loss

DF Relaying

- DF relay node:
  - Maximum-likelihood decoding (MLD)
    $\mathbf{x}_R = \arg\min_{\mathbf{x}\in\mathcal{C}} \|\mathbf{y}_{SR} - \mathbf{h}_{SR}\sqrt{E_S}\mathbf{x}\|^2$, where $\mathcal{C}$ is code book, $|\mathcal{C}| = 2^k$.
  - Drawback: hard decision $\Rightarrow$ information loss
- Destination:
  - Near-optimal decoder (NOD)
    $\arg\max_{\mathbf{x}_S\in\mathcal{C}} \Pr(\mathbf{y}_{SD}|\mathbf{x}_S) \sum_{\mathbf{x}_R\in\mathcal{C}} \Pr(\mathbf{x}_S \to \mathbf{x}_R) \Pr(\mathbf{y}_{RD}|\mathbf{x}_R)$
  - Near-optimal in the context of DF
  - High complexity: $\mathcal{O}(n \cdot 2^k \cdot 2^k)$ per block

Outline

1 Background and Motivation

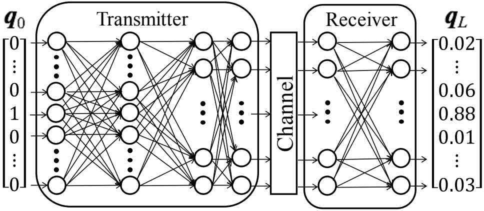
2 Relay-Assisted Cooperative Communication System

3 Learning the Cooperative System

4 Simulation Results

5 Conclusion

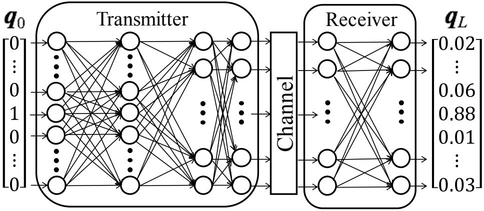
A typical AE



A typical AE for a point-to-point communication system

- Input: one-hot encoding, e.g., $\{00, 01, 11, 10\} \mapsto \{1000, 0100, 0010, 0001\}$

A typical AE



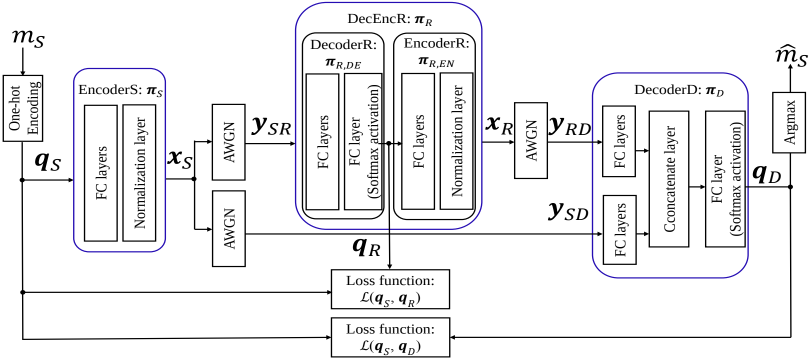
A typical AE for a point-to-point communication system

- Input: one-hot encoding, e.g., $\{00, 01, 11, 10\} \mapsto \{1000, 0100, 0010, 0001\}$
- Output: softmax, i.e., $\phi(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}$, $i = 1, 2, \ldots, k$ and $\mathbf{z} = [z_1, z_2, \ldots, z_k] \in \mathbb{R}^k$

Proposed AE Structure

$\mathbf{q}_R$: soft probability

Advantage: eliminate noise amplification and hard decision



Block diagram of the proposed AE for the cooperative communication system

End-to-End Loss

- Expected loss (a large number of data sets)

$$L_{SD}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D) = \mathbb{E}_{\mathbf{q}_S}[\mathcal{L}(\mathbf{q}_S, \mathbf{q}_D)] \tag{3}$$

End-to-End Loss

- Expected loss (a large number of data sets)

$$L_{SD}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D) = \mathbb{E}_{\mathbf{q}_S}[\mathcal{L}(\mathbf{q}_S, \mathbf{q}_D)] \tag{3}$$

- Estimated through sampling

$$L_{SD}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D) \triangleq \frac{1}{B}\sum_{i=1}^{B} \mathcal{L}(\mathbf{q}_{S,i}, \mathbf{q}_{D,i}) \tag{4}$$

B: batch size
$\{\mathbf{q}_{S,i}, \mathbf{q}_{D,i}\}$: the i-th input output pair of training sample

Objective

$$\textbf{(P1)} \quad \min_{\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D} \quad L_{SD}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D)$$

Proposed AE

How to design the training algorithm?

Proposed AE

How to design the training algorithm?

A desirable way: directly train the whole model to minimize L_{SD}

Proposed AE

How to design the training algorithm?

A desirable way: directly train the whole model to minimize L_{SD}

Experimental results **Do Not** demonstrate a favorable performance

Proposed AE

How to design the training algorithm?

A desirable way: directly train the whole model to minimize L_{SD}

Experimental results **Do Not** demonstrate a favorable performance

A novel training algorithm is required!

A Two-Stage Training Scheme

$$\{\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D\} \Rightarrow \{\boldsymbol{\pi}_S, \boldsymbol{\pi}_{R,DE}, \boldsymbol{\pi}_{R,EN}, \boldsymbol{\pi}_D\} \tag{5}$$

A Two-Stage Training Scheme

$$\{\boldsymbol{\pi}_S, \boldsymbol{\pi}_R, \boldsymbol{\pi}_D\} \Rightarrow \{\boldsymbol{\pi}_S, \boldsymbol{\pi}_{R,DE}, \boldsymbol{\pi}_{R,EN}, \boldsymbol{\pi}_D\} \tag{5}$$

(P2) First stage: $\min_{\boldsymbol{\pi}_S, \boldsymbol{\pi}_{R,DE}} \quad L_{SR}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_{R,DE})$

Second stage: $\min_{\boldsymbol{\pi}_{R,EN}, \boldsymbol{\pi}_D} \quad L_{SD}(\boldsymbol{\pi}_{R,EN}, \boldsymbol{\pi}_D)$

Training SNR

- Fixed SNR: γ
- Mixed SNR: $\gamma \in \{\gamma_l, \gamma_l + \Delta, \cdots, \gamma_u - \Delta, \gamma_u\}$

Training Algorithm

Algorithm 1 Two-stage training of the proposed AE model

Input Number of channel uses n, number of information bits (per message) k; SNR parameters Δ, γ_l and γ_u

FIRST STAGE: TRAINING OF THE SOURCE-RELAY LINK

Construct a partial model for the source-relay link;
Randomly generate $\gamma_{SR} \in \{\gamma_l, \gamma_l + \Delta, \cdots, \gamma_u - \Delta, \gamma_u\}$;
Train this partial model to minimize $L_{SR}(\boldsymbol{\pi}_S, \boldsymbol{\pi}_{R,DE})$;
Save EncoderS and DecoderR;

SECOND STAGE: TRAINING OF THE ENTIRE NETWORK

Load EncoderS and DecoderR;
Incorporate the loaded components to construct the complete AE model;
Randomly generate $\gamma_{IJ} \in \{\gamma_l, \gamma_l + \Delta, \cdots, \gamma_u - \Delta, \gamma_u\}$ for $(I, J) \in \{(S, R), (R, D), (S, D)\}$;
Train the proposed AE model to minimize $L_{SD}(\boldsymbol{\pi}_{R,EN}, \boldsymbol{\pi}_D)$;
Obtain EncoderR and DecoderD.

Outline

1 Background and Motivation

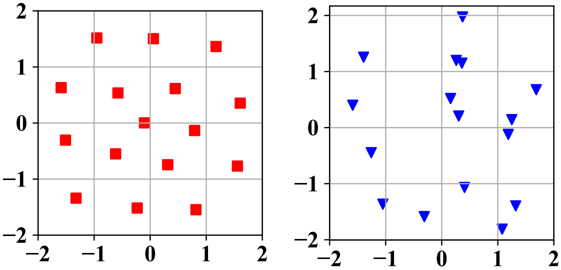
2 Relay-Assisted Cooperative Communication System

3 Learning the Cooperative System

4 Simulation Results

5 Conclusion

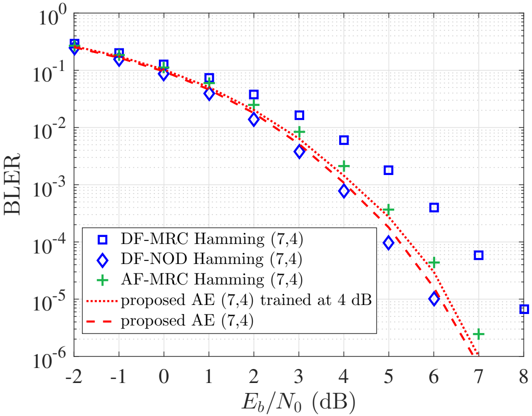
Learned Constellations



Constellations of $\mathbf{x}_S$ (red squares) and $\mathbf{x}_R$ (blue triangles) with an average power constraint for $(n, k) = (2, 4)$

- $\mathbf{x}_S$: APSK-like $\Rightarrow$ Shaping gain
- $\mathbf{x}_R$: Irregular Overlapping Non-conventional

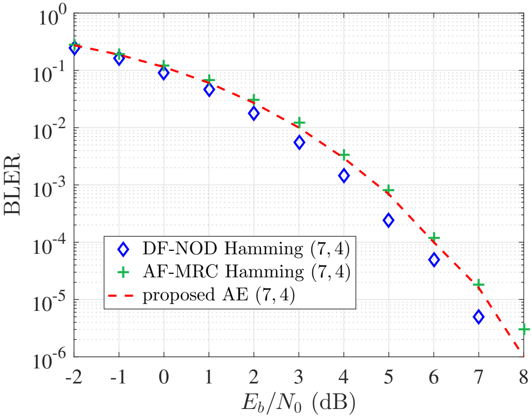
BLER Performance



BLER performance comparison of the proposed AE and the baseline schemes for $(n, k) = (7, 4)$

$\Rightarrow$ Competitive BLER performance

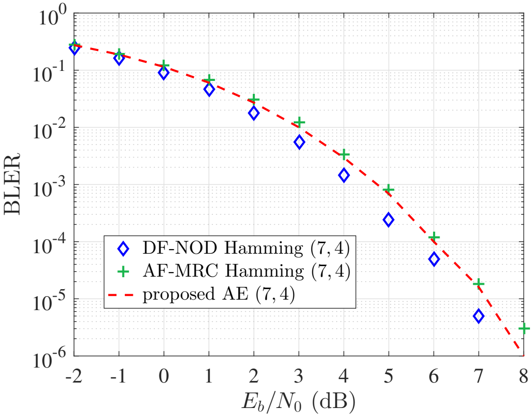
Robustness under Non-Gaussian Channels



BLER performance comparison of the proposed AE and the baseline schemes under the impulse noises

e.g., interference produced by radar signals

Robustness under Non-Gaussian Channels



BLER performance comparison of the proposed AE and the baseline schemes under the impulse noises

e.g., interference produced by radar signals
$\Rightarrow$ Competitive BLER performance

Outline

1. Background and Motivation
2. Relay-Assisted Cooperative Communication System
3. Learning the Cooperative System
4. Simulation Results
5. Conclusion

Conclusion

Scheme	**Noise amplification**	**Hard decision**	**Channel estimation**	**Decoding complexity per block**
DF	No	Yes	Yes	$\mathcal{O}(n \cdot 2^k \cdot 2^k)$
AF	Yes	No	Yes	$\mathcal{O}(n \cdot 2^k)$
AE	No	No	No	$\mathcal{O}(n \cdot 2^k)$

$\Rightarrow$ The proposed AE is a competitive alternative for the conventional relaying techniques DF and AF

Conclusion

Scheme	**Noise amplification**	**Hard decision**	**Channel estimation**	**Decoding complexity per block**
DF	No	Yes	Yes	$\mathcal{O}(n \cdot 2^k \cdot 2^k)$
AF	Yes	No	Yes	$\mathcal{O}(n \cdot 2^k)$
AE	No	No	No	$\mathcal{O}(n \cdot 2^k)$

$\Rightarrow$ The proposed AE is a competitive alternative for the conventional relaying techniques DF and AF

Future works:

A theoretical perspective and performance guarantee need to be provided!

Consider other relay networks, e.g., two-way, full-duplex, ...

Take Away

Carefully designed training algorithm, loss functions, and structure
$\Rightarrow$ AE works

Take Away

Carefully designed training algorithm, loss functions, and structure
$\Rightarrow$ AE works

More general scenarios, Theoretical perspective
$\Rightarrow$ a longer journal version of this work :)

Thanks!

Email: {ylubg}@ust.hk