ROBUST SYMBOL-LEVEL PRECODING VIA AUTOENCODER-BASED DEEP LEARNING

Foad Sohrabi, Hei Victor Cheng, and Wei Yu

Department of Electrical and Computer Engineering
University of Toronto, Toronto, ON, Canada, M5S 3G4
Emails: fsohrabi@ece.utoronto.ca, hei.cheng@utoronto.ca, weiyu@ece.utoronto.ca

ABSTRACT

This paper proposes an autoencoder-based symbol-level precoding (SLP) scheme for a massive multiple-input multiple-output (MIMO) system operating in a limited-scattering environment. By recognizing that only imperfect channel state information (CSI) is available in practice, the goal of the proposed approach is to design the downlink SLP system robust to such imperfect CSI. Toward this goal, this paper leverages the concept of autoencoder wherein the end-to-end communications system is modeled by a deep neural network. By end-to-end training the proposed autoencoder, this paper shows that the downlink symbol-level precoder as well as the receivers' decision rule can be jointly designed in ways that are robust to channel uncertainty. Moreover, this paper introduces a novel two-step training procedure to design a robust precoding scheme for conventional modulations such as quadrature amplitude modulation (QAM) and phase shift keying (PSK). Numerical results indicate that the proposed autoencoder-based framework, either trained by the end-to-end approach in which the receive constellation is a design variable or by the proposed two-step training approach with QAM constellation, can efficiently design a SLP scheme for massive MIMO system which is robust to channel uncertainty.

1. INTRODUCTION

Massive multiple-input multiple-output (MIMO) [1, 2] is a promising technology for the next generations of cellular systems. The massive number of antennas at the base stations (BSs) creates a large degree of freedom for transmit precoding, which can be used to significantly enhance the system performance. While traditional multiuser precoding focuses on eliminating interference between different users, symbol-level precoding (SLP), proposed in [3, 4], seeks to exploit constructive interference for enhancing received signal power. In particular, unlike the traditional precoding methods, which only make use of the channel state information (CSI), SLP also exploits the knowledge of users' data symbols to improve the system performance by manipulating the interfering signals such that they add up constructively at the receivers [5]. To make sure that the received symbols for every user lie in the desired constructive region, the precoding vectors have to be carefully designed which involves formulating and solving non-trivial optimization problems. Most previous works on SLP focus on phase shift keying (PSK) modulation schemes as the decision boundaries are easier to characterize, e.g., [6, 7]. Some recent works consider SLP design for quadrature amplitude modulation (QAM) signaling, which are more widely used in modern communications systems, by formulating and solving the precoder design problem in a noise-less scenario [8, 9].

This work is supported by Huawei Technologies Canada.

Despite the gain achieved by adopting SLP, all these schemes depend on the perfect CSI assumption. However, CSI is never perfect in practice due to reasons such as imperfect channel estimation, limited/delayed feedback, and quantization errors. Therefore, designing symbol-level precoders that are robust to CSI errors is crucial. To this end, a robust precoding scheme has recently been proposed in [10, 11] by considering two types of CSI uncertainty model, namely, spherical bounded model and stochastic Gaussian model. Those robust SLP approaches in [10, 11] are based on the assumption that CSI uncertainty model is accurate. However, in practice, models for CSI errors are not accurate or not even available. Motivated by that, this paper proposes a data-driven framework to design a robust SLP scheme.

In particular, this paper considers a robust SLP design problem for a massive MIMO system operating in a limited scattering environment where only imperfect information about the sparse channel parameters are available. To address this problem, this paper models the end-to-end multi-user massive MIMO system as a deep neural network (DNN) autoencoder. Further, it is shown that a robust SLP design together with a robust receive constellation design can be obtained by end-to-end training of the proposed autoencoder. Such an end-to-end training approach exploits the full capacity of the neural networks and leads us to a robust receive constellation design. Further, to address the issue that the constellation designed by the autoencoder might not be easy to implement in practice, this paper also proposes a novel two-step training approach for conventional constellations. This paper numerically shows that the proposed autoencoder-based framework, either trained by the end-to-end training approach or by the proposed two-step training approach with QAM signaling, can design a SLP scheme for massive MIMO system that is robust to channel uncertainty.

It is noteworthy that the use of DNNs in designing precoding schemes and/or designing the receive constellations has been adopted in some recent works, e.g. [12–14]. However, to the best of our knowledge, this paper is the first to focus on robustness in an end-to-end autoencoder-based SLP systems.

2. SYSTEM MODEL

Consider the downlink of a unicast massive MIMO system in which a BS with M transmit antennas serves K single-antenna users by communicating a message m_k of B bits to user k in each channel use, i.e., $m_k \in \{1, \ldots, 2^B\}$. For such a system, if the perfect instantaneous CSI is available at the BS, the transmitted signal $\mathbf{x}$ can be written as a function of instantaneous CSI and the intended messages $\mathbf{m} = [m_1, \ldots, m_K]$:

$$\mathbf{x} = \tilde{\mathcal{P}}\left(\mathbf{H}, \mathbf{m}\right), \tag{1}$$

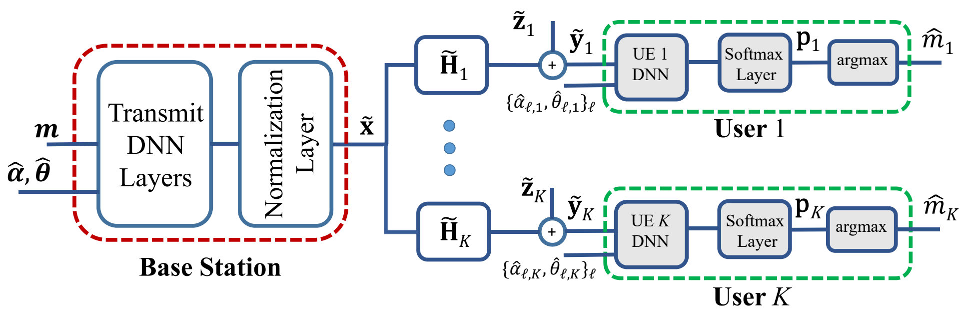

Fig. 1. A DNN autoencoder representing a K-user symbol-level precoding massive MIMO system.

where $\tilde{\mathcal{P}}$ is the precoding function, $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_K]^T$, and $\mathbf{h}_k \in \mathbb{C}^M$ is the vector of channel gains between the BS and user k. However, in practice, the BS has only access to imperfect/partial CSI. In this paper, we aim to design a SLP scheme which is robust to channel uncertainty in a propagating environment with sparse channels between user terminals and the BS. In such an environment, the channel of the k^{th} user can be modeled with L propagation paths [15]:

$$\mathbf{h}_k = \tfrac{1}{\sqrt{L}} \textstyle\sum_{\ell=1}^{L} \alpha_{\ell,k} \mathbf{a}_t\left(\theta_{\ell,k}\right), \tag{2}$$

where $\alpha_{\ell,k}$ is the complex gain of the ℓ^{th} path between the BS and user k, $\theta_{\ell,k}$ is the corresponding AoD, and $\mathbf{a}_t\left(\cdot\right)$ is the transmit array response vector. For a uniform linear array with half-wavelength antenna spacing, we have $\mathbf{a}_t\left(\theta\right) = \left[1, e^{j\pi \sin(\theta)}, \ldots, e^{j\pi(M-1)\sin(\theta)}\right]$.

In this paper, we assume that the available CSI at the BS is in the form of imperfect estimation of the sparse channel parameters as:

$$\hat{\alpha}_{\ell,k} = \alpha_{\ell,k} + \Delta\alpha_{\ell,k}, \tag{3a}$$

$$\hat{\theta}_{\ell,k} = \theta_{\ell,k} + \Delta\theta_{\ell,k}, \tag{3b}$$

where $\Delta\alpha_{\ell,k}$ and $\Delta\theta_{\ell,k}$ are the estimation errors in channel gains and AoDs, respectively. As an example, such a CSI model is valid for frequency-division duplex systems in which the parameters of the sparse channels are first estimated by the users through downlink training phase and then fed back to the BS. In such a scenario, the estimation error in channel gains is typically modeled by a Gaussian distributed random variable, i.e., $\Delta\alpha_{\ell,k} \sim \mathcal{CN}\left(0, \sigma_{\Delta\alpha}^2\right)$, while the estimation error in AoDs is characterized by a uniform distributed random variable, i.e., $\Delta\theta_{\ell,k} \sim \mathcal{U}\left(-\Delta\theta_{\max}, \Delta\theta_{\max}\right)$ [16, 17].

Under such a system model, the transmitted signal should now be designed as a function of the estimated channel parameters and the intended messages as:

$$\mathbf{x} = \mathcal{P}\left(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\theta}}, \mathbf{m}\right), \tag{4}$$

where $\mathcal{P}$ is the SLP function, $\hat{\boldsymbol{\alpha}} \triangleq [\hat{\alpha}_{1,1}, \ldots, \hat{\alpha}_{L,K}]$, and $\hat{\boldsymbol{\theta}} \triangleq [\hat{\theta}_{1,1}, \ldots, \hat{\theta}_{L,K}]$. The received signal at user k can then be modeled as:

$$y_k = \mathbf{h}_k^T \mathbf{x} + z_k, \tag{5}$$

where $z_k \sim \mathcal{CN}(0, 2\sigma^2)$ is the additive white Gaussian noise. Finally, user k seeks to recover the intended message based on a set of predefined decision rules which are functions of the received signal, y_k, and the available CSI at user k, $\{\hat{\alpha}_{\ell,k}, \hat{\theta}_{\ell,k}\}_{\ell=1}^{L}$.

The ultimate goal of this paper is to design the transmitter's SLP function satisfying a per-transmission power constraint, i.e., $\|\mathbf{x}\|^2 \leq P$, as well as the receivers' decision rules such that they are robust to the channel estimation error in (3). In the next section, we propose a DNN autoencoder framework to tackle this problem.

3. END-TO-END SYMBOL-LEVEL PRECODING SYSTEM DESIGN USING AUTOENCODER

In this section, we show how to represent an end-to-end unicast massive MIMO system explained in Section 2 as an autoencoder and explain how to train this autoencoder to find the SLP mapping as well as the receivers' decision rules.

3.1. Autoencoder Representation

In order to use existing deep learning libraries which only support real-value operations, we first transform the complex signal model in (5) into its equivalent real representation as:

$$\underbrace{\begin{bmatrix} \Re\{y_k\} \\ \Im\{y_k\} \end{bmatrix}}_{\widetilde{\mathbf{y}}_k} = \underbrace{\begin{bmatrix} \Re\{\mathbf{h}_k^T\} & -\Im\{\mathbf{h}_k^T\} \\ \Im\{\mathbf{h}_k^T\} & \Re\{\mathbf{h}_k^T\} \end{bmatrix}}_{\widetilde{\mathbf{H}}_k} \underbrace{\begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix}}_{\widetilde{\mathbf{x}}} + \underbrace{\begin{bmatrix} \Re\{z_k\} \\ \Im\{z_k\} \end{bmatrix}}_{\widetilde{\mathbf{z}}_k}.$$

According to (4), the transmitted signal needs to be designed as a function of the estimated channel parameters, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\theta}}$, and the intended messages $\mathbf{m}$. In this work, we develop such a function by employing a T-layer neural network in which the real-valued transmitted signal can be written as:

$$\widetilde{\mathbf{x}} = \sigma_T\left(\mathbf{W}_T \sigma_{T-1}\left(\cdots \mathbf{W}_2 \sigma_1\left(\mathbf{W}_1 \mathbf{v} + \mathbf{b}_1\right) + \cdots\right) + \mathbf{b}_T\right),$$

where σ_t, $\mathbf{W}_t$, and $\mathbf{b}_t$ are the activation function, the weights, and the biases in the t^{th} layer, respectively, and $\mathbf{v} \triangleq [\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\theta}}, \mathbf{m}]$ is the input vector to the neural network which contains the information about the available CSI and the intended messages. In order to ensure that the per-transmission total power constraint, i.e., $\|\mathbf{x}\|^2 \leq P$, is satisfied, a normalization layer with activation function $\sigma_T(x) = \min(\sqrt{P}, \|x\|)\frac{x}{\|x\|}$ is used at the last layer of the transmit neural network.

Each receiver is also implemented by a feedforward neural network with R dense layers, where the input to the k^{th} user's DNN is the received signal, y_k, together with the available CSI at user k, $\{\hat{\alpha}_{\ell,k}, \hat{\theta}_{\ell,k}\}_{\ell=1}^{L}$. The last layer of the receive neural networks employs softmax activation function in order to output the probability vector, $\mathbf{p}_k \in (0,1)^{|2^B|}$, where its i^{th} element indicates the probability of the intended message for user k being i. The decoded message at user k, denoted by $\hat{m}_k$, corresponds then to the index of the element of $\mathbf{p}_k$ with the highest probability.

The block diagram of the proposed autoencoder that models a unicast massive MIMO system is shown in Fig. 1. The procedure of learning the symbol-level precoding function and the receivers' decision rules using the concept of autoencoder is discussed next.

3.2. Autoencoder Training

We propose two different training approaches to designing the transmit and receive operations in the DNN autoencoder in Fig. 1. In the first approach presented in Section 3.2.1, we seek to jointly design the transmit precoding scheme and the receive constellations by end-to-end training the autoencoder in Fig.1. In order to enable the proposed autoencoder to work with the conventional constellations such as QAM and PSK constellations, we propose a second approach based on a two-step training process in Section 3.2.2.

3.2.1. End-to-End Training

The ultimate goal of the considered communications network, which is modeled as an autoencoder in Fig. 1, is to successfully recover the intended messages at the users. This communication task can be treated as a classification problem for which the categorical cross-entropy is a common choice of loss function [18]. Following this observation, the proposed autoencoder can be trained end-to-end by employing stochastic gradient descent (SGD) algorithms in order to minimize the average cross-entropy between the one-hot representation of the intended messages and the probability vectors generated by the users, which can be written as:

$$\mathcal{L}_{\text{CE}} = -\mathbb{E}_{\mathbf{u}} \left[\frac{1}{K2^B} \sum_{k=1}^{K} \sum_{m=1}^{2^B} \log p_{k,m} \right], \tag{6}$$

where $p_{k,m}$ is the m^{th} element of the probability vector $\mathbf{p}_k$, and the expectation is over all the stochastic parameters in the system, i.e., $\mathbf{u} \triangleq [\widetilde{\mathbf{z}}, \boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\theta}, \hat{\boldsymbol{\theta}}, \mathbf{m}]$.

Such an end-to-end training enables us to jointly design the transmit precoding scheme as well as the decision rules at each user. However, the final designed decision rules in this approach (as we will see in Section 4) can be quite different from the conventional decision rules. As such, they might not be easy to implement in practice. To address this issue, we next consider training the proposed autoencoder for conventional constellations.

3.2.2. Two-Step Training for Conventional Constellations

In this section, we propose a two-step training approach to learn the parameters of the autoencoder in Fig.1 for a scenario in which the general shape of the receive constellation is given. In this case, the receive neural networks are first trained to characterize the decision boundaries of the constellation of interest, then the network in Fig. 1 is trained to learn the parameters of the transmit DNN while the receivers' parameters are fixed. Note that for the fixed conventional constellation, we could have used a conventional receiver with the posterior probabilities computed based on the traditional decision boundaries and the noise statistics. However, it is not easy to incorporate such a receiver in the Tensorflow environment for the training of the robust precoder. For this reason, we prefer to replace the conventional receiver by a neural network trained in the first step, followed by the training of the precoder in the second step.

To illustrate the first step, let us assume that the intended message m for a user[1] is modulated using one of the conventional modulations, e.g., QAM or PSK, and the real-value representation of the corresponding modulation symbol is denoted by $\mathbf{s}_m \in \{\mathbf{s}_1, \ldots, \mathbf{s}_{2^B}\}$. Further, for correct demodulation, we consider that the received signal $\widetilde{\mathbf{y}}$ is scaled as:

$$\bar{\mathbf{y}} \triangleq \beta \widetilde{\mathbf{y}} = \beta \widetilde{\mathbf{H}} \widetilde{\mathbf{x}} + \beta \widetilde{\mathbf{z}} = \mathbf{s}_m + \mathbf{n}, \tag{7}$$

where β is the scaling factor and $\mathbf{n}$ is the effective noise. The objective of the receive DNN is to generate a probability vector $\mathbf{p}$ containing the probability for each of the possible modulation symbols given the observation $\bar{\mathbf{y}}$. To develop such a DNN, we first collect a large set of labeled data consisting of $\bar{\mathbf{y}} \in \mathbb{R}^2$ as the input and $\bar{\mathbf{p}} \triangleq [\bar{p}_1, \ldots, \bar{p}_{2^B}]$ as the output, where $\bar{p}_i = \Pr(\mathbf{s}_m = \mathbf{s}_i | \bar{\mathbf{y}})$ is computed for $\mathbf{n} \in \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$. This labeled data set is then used to train the receive DNN by applying SGD methods to minimize the average MSE, i.e., $\mathcal{L}_{\text{MSE}} = \mathbb{E}\left[\|\bar{\mathbf{p}} - \mathbf{p}\|^2\right]$.

After the receive DNNs are designed, we can apply the training approach in Section 3.2.1 to further obtain the transmit precoding scheme $\tilde{\mathcal{P}}$ as well as the scaling factor β.

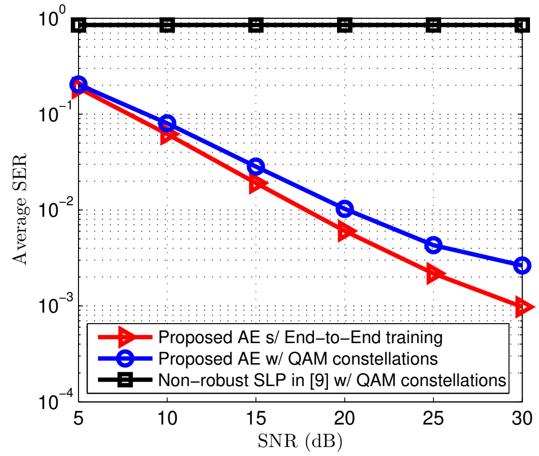

Fig. 2. Average SER versus SNR.

4. NUMERICAL RESULTS

In this section, we illustrate the performance of the proposed robust autoencoder-based SLP and compare that with the performance of the non-robust SLP scheme for QAM signaling in [9]. The reason that we only compare our proposed method with the non-robust SLP scheme is that, to the best of our knowledge, there is no work on robust SLP for limited-scattering environments. In our numerical experiments, we consider a unicast massive MIMO system in which a BS with $M = 128$ antennas transmits independent 4-bit messages to $K = 3$ users in a single path environment, i.e., $L_k = 1, \forall k$. Further, we assume that the channel parameters are modeled as $\alpha_k \sim \mathcal{CN}(0.5 + 0.5\imath, 1)$ and $\theta_k \sim \mathcal{U}(\phi_k - 5^\circ, \phi_k + 5^\circ), \forall k$, with $\{\phi_1, \phi_2, \phi_3\} = \{-30^\circ, 0^\circ, +30^\circ\}$.[2]

We implement the proposed network in Fig. 1 on TensorFlow [19] by employing Adam optimizer [20] with a learning rate progressively decreasing from 10^{-3} to 10^{-5}. We consider 4-layer neural networks at the transmitter and receivers, while the number of transmit and receive hidden neurons at different layers are $[1024, 512, 512, 256]$ and $[256, 128, 64, 16]$, respectively. Further, we adopt the rectified linear unit (ReLU) activation function

[1] In this part, we drop the user index k for the ease of notation.

[2] In this section, we drop the index ℓ since single-path channel model is considered.

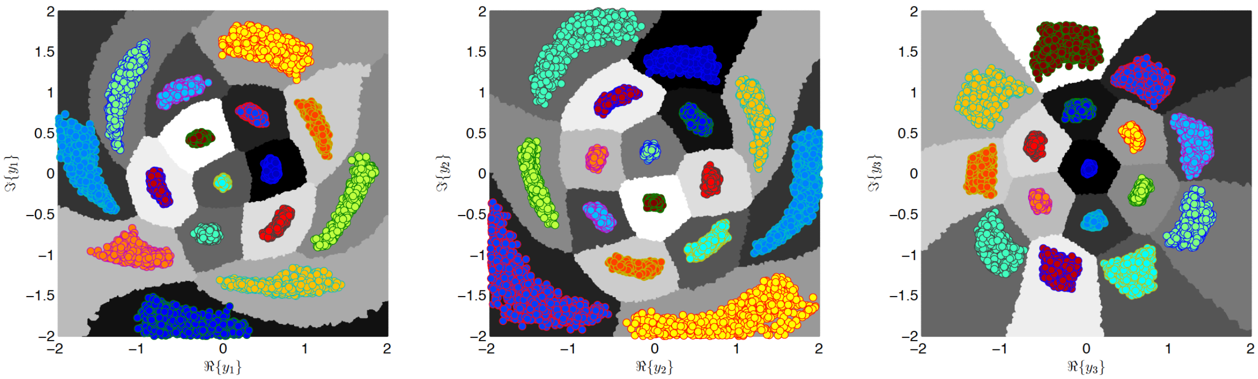

Fig. 3. The decision boundaries (in grey scale) designed by the autoencoder together with the noiseless received signal (as circles) for a robust SLP with $K = 3$ users.

at the hidden layers. In the training stage, it is assumed that the noise variance, σ^2, is randomly generated so that the signal-to-noise ratio, $\text{SNR} \triangleq 10\log_{10}(\frac{P}{\sigma^2})$, is uniformly distributed in a reasonable range, i.e., $\text{SNR} \sim \mathcal{U}(5, 30)$dB. This strategy enables the trained network to operate on a wide range of SNRs. Finally, in order to produce a robust precoder and/or robust constellation, 10^5 channel realizations are used in the training phase; the parameters of CSI uncertainty are $\sigma_{\Delta\alpha} = 0.001$ and $\Delta\theta_{\text{max}} = 1^\circ$.

As the first experiment, we plot the average symbol error rate (SER) against SNR in Fig. 2 for a scenario with channel uncertainty parameters $\sigma_{\Delta\alpha} = 0.001$ and $\Delta\theta_{\text{max}} = 1^\circ$. It can be seen from Fig. 2 that the proposed autoencoder-based framework, either with trainable receive constellations or with fixed 16-QAM constellations, achieves a reasonable SER by designing a robust symbol-level precoder while the non-robust SLP approach in [9] fails to do so. Moreover, Fig. 2 indicates that the autoencoder-based SLP, in which the receive constellation is designed via end-to-end training, achieves a better performance as compared to the autoencoder which is trained for QAM signaling. This suggests that the end-to-end training of the proposed autoencoder can help us design a more robust receive constellation.

To show that the final design of the receive constellation is indeed robust to the CSI error, Fig. 3 plots the final decision boundaries designed by the autoencoder together with the noiseless received signal in this case for all $K = 3$ users. Note that the noiseless points are not fixed due to the channel uncertainty and due to the fact that they depend on the data symbols of the other users. From Fig. 3, it can be seen that the receive constellation boundaries are non-uniform where the decision boundary of inner constellations are smaller than that of outer constellations. This can be justified by considering the fact that the outer constellation regions are more sensitive to the phase error in AoDs which is the main source of CSI error in the considered setup.

Next, in Fig. 4, we plot the average SER against the maximum range of estimation error in AoDs, $\Delta\theta_{\text{max}}$, in a high SNR regime, i.e., $\text{SNR} = 30$dB, while we set $\sigma_{\Delta\alpha} = 0.001$. It can be seen from Fig. 4 that there is a sharp phase transition around $\Delta\theta_{\text{max}} = 0.008$ in the performance of the non-robust SLP algorithm in [9], indicating that although such a SLP approach can achieve a very good performance when perfect CSI is available, it completely fails to work when there is a small error in AoD estimations. This means that non-robust SLP schemes such as the one in [9] can be extremely sensitive to the CSI uncertainty. Fig. 4 also illustrates that the proposed autoencoder-based framework achieves a decent constant SER of 10^{-3} over the range of $\Delta\theta_{\text{max}} \in [0, 1]$, which is exactly the range considered in the training stage. This indicates that the proposed SLP scheme can indeed provide a robust design to the channel uncertainty. Finally, Fig. 4 shows that a relatively good performance can still be obtained by the proposed autoencoder framework even if the actual estimation error in AoDs is larger than that in the training phase.

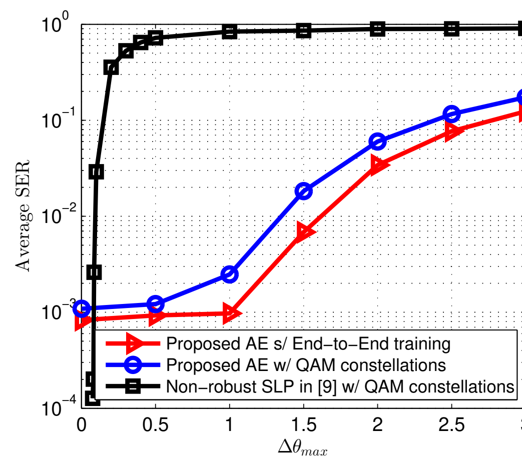

Fig. 4. Average SER versus $\Delta\theta_{\text{max}}$.

5. CONCLUSION

In this work, we propose a DNN autoencoder framework for symbol-level precoding which is robust to channel uncertainty. Two different approaches for training the proposed autoencoder are presented. In the first approach, we train the proposed autoencoder end-to-end to jointly design the transmitter's precoding scheme and the receivers' decision rules. The second training approach which involves two-steps, namely, receiver training followed by transmitter training, is also presented to enable the proposed autoencoder to operate with conventional modulation techniques. By injecting random samples of channel variations into the training process, the proposed DNN architecture is able to design an autoencoder-based SLP scheme for massive MIMO system which is robust to channel uncertainty.

6. REFERENCES

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[4] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, July 2015.

[5] A. Li, D. Spano, J. Krivochiza, S. Domouchtsidis, C. G. Tsinos, C. Masouros, S. Chatzinotas, Y. Li, B. Vucetic, and B. Ottersten, "Interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions," July 2019. [Online]. Available: https://arxiv.org/pdf/1907.05530.pdf

[6] A. Li and C. Masouros, "Interference exploitation precoding made practical: Optimal closed-form solutions for PSK modulations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7661–7676, Sept. 2018.

[7] K. L. Law and C. Masouros, "Symbol error rate minimization precoding for interference exploitation," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5718–5731, Nov. 2018.

[8] A. Kalantari, C. G. Tsinos, M. Soltanalian, S. Chatzinotas, W. Ma, E. G. Larsson, and B. Ottersten, "M-QAM precoder design for MIMO directional modulation transceivers," Nov. 2018. [Online]. Available: https://arxiv.org/abs/1702.06878

[9] A. Li, C. Masouros, Y. Li, B. Vucetic, and A. L. Swindlehurst, "Interference exploitation precoding for multi-level modulations: Closed-form solutions," Nov. 2018. [Online]. Available: https://arxiv.org/abs/1811.03289

[10] A. Haqiqatnejad, F. Kayhan, and B. Ottersten, "Robust design of power minimizing symbol-level precoder under channel uncertainty," in *IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[11] A. Haqiqatnejad, F. Kayhan, and B. Ottersten, "Robust SINR-constrained symbol-level multiuser precoding with imperfect channel knowledges," Mar. 2019. [Online]. Available: https://arxiv.org/abs/1903.03371

[12] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[13] F. Sohrabi and W. Yu, "One-bit precoding constellation design via autoencoder-based deep learning," in *IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2019.

[14] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," May 2019. [Online]. Available: https://arxiv.org/abs/1905.13212

[15] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[16] L. Sun, Y. Qin, Z. Zhuang, R. Chen, Y. Zhang, J. Lu, F. Shu, and J. Wang, "A robust secure hybrid analog and digital receive beamforming scheme for efficient interference reduction," *IEEE Access*, vol. 7, pp. 22 227–22 234, Feb. 2019.

[17] C. Pradhan, A. Li, L. Zhuo, Y. Li, and B. Vucetic, "Beam misalignment aware hybrid transceiver design in mmWave MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 306–10 310, Oct. 2019.

[18] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[19] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," Mar. 2016. [Online]. Available: https://arxiv.org/abs/1603.04467

[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014. [Online]. Available: https://arxiv.org/abs/1412.6980