A Data Augmentation based DNN Approach for Outage-Constrained Robust Beamforming

Minglei You
Department of Engineering
Durham University
Durham, UK, DH1 3LE
minglei.you@durham.ac.uk

Gan Zheng
Wolfson School of Mechanical, Electrical and Manufacturing Engineering
Loughborough University
Loughborough, UK, LE11 3TU
g.zheng@lboro.ac.uk

Hongjian Sun
Department of Engineering
Durham University
Durham, UK, DH1 3LE
hongjian.sun@durham.ac.uk

***Abstract*—This paper studies the long-standing problem of outage-constrained robust downlink beamforming in the multi-user multi-antenna wireless communications systems. State of the art solutions have very high computational complexity which poses a major challenge to meet the latency requirement in the future communications systems, e.g., the targeted 1 ms end-to-end latency in 5G. By transforming the robust beamforming problem into a deep learning problem, we propose a new unsupervised data augmentation based deep neural network (DNN) method to address the outage-constrained robust beamforming problem with uncertain channel state information at the transmitter. Simulation results demonstrate that our proposed data augmentation based DNN method for the robust beamforming problem is capable to satisfy the required outage probability, and most importantly, compared to the benchmark Bernstein-Type Inequality (BTI) method, it is less conservative, more power efficient and several orders of magnitude faster.**

I. Introduction

In the multi-user and multi-antenna systems, the beamforming method has shown its excellent performance in enhancing the quality of service (QoS) and improving the system throughput [1]. By exploiting the channel state information (CSI) at the transmitter side (CSIT), the beamforming algorithms are capable to optimize the resource allocations between multiple users, while keeping the interference low [2]. However, perfect CSIT is not always available in practical systems. Even under the assumption of perfect CSI at the receiver side, the feedback is prone to quantization errors. Moreover, the time-varying nature of the wireless channels makes the estimated CSI vulnerable to channel error and uncertainty.

With imperfect CSIT, the resource allocation algorithms including the beamforming algorithms are subject to performance deterioration. This makes the optimal beamforming under perfect CSIT suboptimal with regard to the designed targets, such as throughput and signal-to-interference plus noise ratio (SINR). It may even fail the communications due to the uncertainty in the estimated CSIT, which makes the whole communications system unreliable. Traditionally, this challenge is addressed via robust beamforming, which has been well studied to provide the worst-case or probabilistic performance guarantee using semidefinite relaxation [2], Bernstein-Type Inequality (BTI) method and Large Deviation Inequality (LDI) method [3]. These methods have been then extensively applied in the study of systems with CSI uncertainties, including simultaneous wireless information and power transfer (SWIPT) [4] and multicell interference networks [5].

However, existing numerical methods are subject to high computational complexity and cannot meet the latency requirement in the fifth-generation (5G) and beyond networks. The targeted latency in 5G networks is below 1 millisecond (ms), which is in contrast to the 10 ms performance in 4G networks [6]. This is essential in supporting various time-critical services, such as self-driving cars and industrial automation applications, but also poses a more stringent constraint to the advanced signal processing techniques to be applicable in these time-critical systems, whose advantages such as throughput improvement might be invalid due to the high complexity. Together with the potential performance degradation due to uncertainty in CSI, the robust beamforming problem is challenging from both latency and reliability aspects.

To address the aforementioned challenges, we propose an unsupervised data augmentation based Deep Neural Network (DNN) approach for the outage constrained robust beamforming in multi-user wireless networks. The proposed training method transforms the robust beamforming problem under uncertainties to a deep learning problem, where the beamforming output from the DNN is trained with the data augmentation method to be robust against potential CSI errors, and the estimated CSIT is the DNN's input. Comparing to the benchmarking BTI algorithm, the simulations show that the proposed data augmentation based DNN approach is capable to guarantee the required outage probability, and achieve better power efficiency and 800 times faster time performance.

In the rest of the paper, we will first formulate the robust beamforming problem in multi-user wireless communication networks in Section II. The robust beamforming problem is then transformed to a deep learning problem, and a novel data augmentation based DNN method is proposed in Section III. The simulation results are presented in Section IV. Finally, the conclusions are drawn in Section V.

II. Problem Formulation

In this work, we consider a multi-user wireless communications network, where the transmitter is equipped with n_T antennas and each receiver (user) has a single antenna. Specifically, an M-user multiple-input and single-output (MISO) system is considered as follows [2]:

$$y_m = \mathbf{h}_m^H \left(\sum_{n=1}^{M} \mathbf{t}_n x_n \right) + N_0, m = 1, \ldots, M, \tag{1}$$

where $(\cdot)^H$ is the conjugate transpose operation and x_n is the complex digital symbol sent to user n with $\mathbb{E}_{x_n}\{|x_n|^2\} = 1$ and $\mathbb{E}_x\{\cdot\}$ denotes the expectation operation with respect to x, y_m (complex scalar) is the received symbol at the user m, $\mathbf{h}_m$ ($n_T \times 1$ complex vector) is the MISO channel coefficients from the transmitter to user m , $\mathbf{t}_m$ ($n_T \times 1$ complex vector) is the transmit beamforming vector for user m and N_0 is the power of the noise at user m that follows a circularly symmetric complex Gaussian distribution.

For each user m, it decodes its own information signal and treats others' information signals as interference. Hence the SINR γ_m can be given as follows:

$$\begin{aligned} \gamma_m &= \frac{|\mathbf{h}_m^H \mathbf{t}_m|^2}{\sum_{n=1, n\neq m}^{M} |\mathbf{h}_m^H \mathbf{t}_n|^2 + N_0} \\ &\triangleq \psi(\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}, \{\mathbf{h}_1, \ldots, \mathbf{h}_M\}). \end{aligned} \tag{2}$$

The transmit power to user m is given by $||\mathbf{t}_m||_2^2$, therefore the total transmit power P for all M users is given as follows:

$$P(\mathbf{t}_1, \ldots, \mathbf{t}_M) = \sum_{m=1}^{M} ||\mathbf{t}_m||_2^2, \tag{3}$$

where $|| \cdot ||_2$ denotes the Euclidean norm. The problem of finding the optimal configurations of $\mathbf{t}_1, \ldots, \mathbf{t}_M$ in (2) with targeted performances such as throughput or SINR are called beamforming optimization, whose solutions under perfect CSIT have been well studied in the existing literature [7].

In practical systems, the transmitter cannot obtain the perfect CSIT, because of the estimation errors at the transmitter/receiver side, or quantization errors in the feedback from the receivers. In this paper, it is assumed that CSIT is not perfect, which is modeled as follows:

$$\mathbf{h}_m = \tilde{\mathbf{h}}_m + \Delta \mathbf{h}_m, \tag{4}$$

where the CSIT estimation $\tilde{\mathbf{h}}_m$ is known at the transmitter side, and $\Delta \mathbf{h}_m$ is the channel error.

The channel error $\Delta \mathbf{h}_m$ is a random variable, we cannot know the actual CSI $\mathbf{h}_m$ and need to make beamforming decisions based on the estimated CSI $\tilde{\mathbf{h}}_m$ and the statistical information about $\Delta \mathbf{h}_m$. In this paper, the robust beamforming problem is formulated as the minimization of the expectation of transmit power, while satisfying the outage constraints for each user due to different channel error $\{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M\}$ under a given set of estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$. Mathematically, this problem is formulated as

$$\begin{aligned} \min_{\mathbf{t}_1, \ldots, \mathbf{t}_M} \quad & \mathbb{E}_{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M} \{P(\mathbf{t}_1, \ldots, \mathbf{t}_M)\} \\ \text{s.t.} \quad & \mathbb{E}_{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M} \{\Pr\{\gamma_m \leq \Gamma_m\}\} \leq \rho_{\max}, \forall m, \end{aligned} \tag{5}$$

where $\Pr\{x\}$ denotes the probability of the event x, $\rho_{\max}$ denotes the maximum outage probability and Γ_m denotes the targeted SINR threshold. The outage is defined as the case that the user's SINR γ_m is below the targeted SINR threshold Γ_m, where the constraints in (5) is to stress that for each user this outage probability should be statistically bounded by $\rho_{\max}$ with regard to all potential channel error $\{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M\}$.

Note that the statistical formulation in (5) is different from beamforming problem with perfect CSI. The exact channel error $\{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M\}$ cannot be known in advance, where usually only historical or statistical distributions can be known. Therefore with a given estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$, solutions $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$ via the robust beamforming (5) are functions of both channel error $\{\Delta \mathbf{h}_1, \ldots, \Delta \mathbf{h}_M\}$ and estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$, while the traditional beamforming solutions with perfect CSI only rely on the estimated CSI.

The problem (5) is highly involved to solve because there is no closed-form expression for the outage constraint. In the existing literature, this problem is partially solved by robust beamforming methods such as BTI method and LDI method [1]. However due to the complexity, these existing numerical algorithms are subject to the high computational latency and the solutions are conservative which requires much more power than necessary to guarantee the robustness. Due to the complexity of the problem, the required time for traditional methods are large (over hundreds of milliseconds), which cannot meet the 5G's latency requirement of sub-millisecond time performance. In this paper, we address the time and reliability challenges under imperfect CSIT using a data augmentation based DNN method as detailed in the next section.

III. Data Augmentation based DNN Method

Compared to the traditional numerical methods, the deep learning based solutions are attractive since the inference of the trained models is competitive in its time performance, which is especially desirable in the scenarios with stringent latency requirements. In this section, a data augmentation based DNN method is proposed to solve the robust beamforming problem under imperfect CSIT.

Assume there is a perfect robust beamforming function $f(\cdot)$, where for each estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$, its outputs are the robust beamforming solutions $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$ that satisfy the constraints defined in the optimization problem (5), which can be written as follows:

$$\{\mathbf{t}_1, \ldots, \mathbf{t}_M\} = f(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M). \tag{6}$$

Next assume there is a DNN network $f_{\boldsymbol{\theta}}(\cdot)$ with parameter set $\boldsymbol{\theta}$ that can represent (6) as follows:

$$\{\mathbf{t}_1, \ldots, \mathbf{t}_M\} = f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta}). \tag{7}$$

In this way, the general objective of the DNN based method is to use deep learning method to find the parameter set $\boldsymbol{\theta}$, such that (7) produces solutions for (5). The deep learning method cannot be directly applied to solve the optimization problem (5) because of the constraints and the channel errors, where several transformations are performed as follows.

Firstly, it is noticed that the outage performance needs to be derived in the constraints in (5). To facilitate the calculation of the outage, the quantile function $g_m(\gamma_m, \rho_{\max})$ is introduced as follows [8]:

$$g_m(\gamma_m, \rho_{\max}) = \inf\{\gamma | \Pr\{\gamma_m \leq \gamma\} \leq \rho_{\max}\}, \tag{8}$$

where the quantile function $g_m(\gamma_m)$ calculates the minimum SINR value γ that makes the probability of $\gamma_m \leq \gamma$ no more than the maximum outage probability $\rho_{\max}$. This transforms the original constraint in (5) to its equivalent form as follows:

$$\begin{aligned} \min_{\{\mathbf{t}_1,\ldots,\mathbf{t}_M\}} \quad & \mathbb{E}_{\{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}\{P(\mathbf{t}_1,\ldots,\mathbf{t}_M)\} \\ \text{s.t.} \quad & \mathbb{E}_{\{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}\{g_m(\gamma_m, \rho_{\max}) - \Gamma_m\} \leq 0, \forall m. \end{aligned} \tag{9}$$

Next, in (5), for every given estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$, the traditional optimization method requires a new numerical search. Different from that, this paper is to find the a universal parameter set $\boldsymbol{\theta}$ such that the DNN network $f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta})$ will provide solutions to all estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$ as inputs, instead of training different sets $\boldsymbol{\theta}$ for every combination of $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$. By using the relation between γ_m and the beamforming solutions $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$ in (2) as well as the definition in (7), the problem in (9) is further transformed to find the optimal parameter set $\boldsymbol{\theta}$ as follows:

$$\begin{aligned} \min_{\boldsymbol{\theta}} \mathbb{E}_{\substack{\{\tilde{\mathbf{h}}_1,\ldots,\tilde{\mathbf{h}}_M\} \\ \{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}} & \{P(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta}))\} \\ \text{s.t. } \mathbb{E}_{\substack{\{\tilde{\mathbf{h}}_1,\ldots,\tilde{\mathbf{h}}_M\} \\ \{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}} & \{g_m(\psi(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta}), \mathbf{h}_1, \ldots, \mathbf{h}_M), \\ & \rho_{\max}) - \Gamma_m\} \leq 0, \forall m. \end{aligned} \tag{10}$$

Since the deep learning method cannot directly include the constraint in the training procedure, the Lagrange dual method [9] is applied and the Lagrangian of (10) is derived as follows:

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\lambda}) \triangleq & \mathbb{E}_{\substack{\{\tilde{\mathbf{h}}_1,\ldots,\tilde{\mathbf{h}}_M\} \\ \{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}} \{P(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta}))\} + \\ & \sum_{m=1}^{M} \lambda_m \mathbb{E}_{\substack{\{\tilde{\mathbf{h}}_1,\ldots,\tilde{\mathbf{h}}_M\} \\ \{\Delta\mathbf{h}_1,\ldots,\Delta\mathbf{h}_M\}}} \{g_m(\psi(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M; \boldsymbol{\theta})), \rho_{\max}) - \Gamma_m\} \end{aligned} \tag{11}$$

where $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_M\}$ are the Lagrange dual variables associated with each constraint in (10). Then the dual problem of (10) can be given as follows:

$$\begin{aligned} \max_{\boldsymbol{\lambda}} \quad & \min_{\boldsymbol{\theta}} L(\boldsymbol{\theta}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \lambda_m \geq 0, \forall m. \end{aligned} \tag{12}$$

By using $L(\boldsymbol{\theta}, \boldsymbol{\lambda})$ as the loss function, it is seen that the transformed problem in (12) provides an unsupervised deep learning approach to solve the original outage constrained robust beamforming under imperfect CSIT problem in (9).

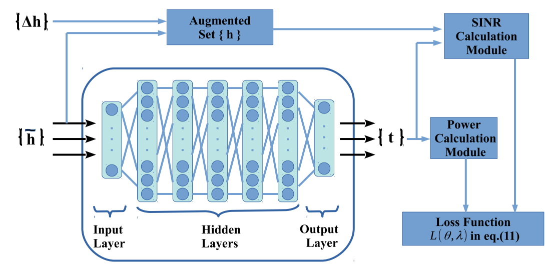

Fig. 1. An illustration of the proposed DNN training method for the robust beamforming problem in the URLLC networks.

Recall that during the operation, only the estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$ are available to calculate the beamforming $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$, while the actual CSI $\{\mathbf{h}_1, \ldots, \mathbf{h}_M\}$ are not known. To make the beamforming outputs $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$ robust to all possible $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$ and $\{\Delta\mathbf{h}_1, \ldots, \Delta\mathbf{h}_M\}$, we propose a data augmentation based DNN method, which is illustrated in Fig. 1. During the training, the inputs are the estimated CSI $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$, while the outputs of the DNN are the beamforming $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$. To evaluate the performance of the DNN outputs and calculate the loss function, we need to calculate the outage performance. Since there is no closed-form expression for the outage constraint in (5), we propose that each $\{\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_M\}$ is combined with a set of potential errors to facilitate the evaluation of the outage probability, which results in an augmented CSI data set containing a group of the possible CSI realizations $\{\mathbf{h}_1, \ldots, \mathbf{h}_M\}$. Then this augmented CSI data set is used to calculate their corresponding SINR with regard to the same beamforming output $\{\mathbf{t}_1, \ldots, \mathbf{t}_M\}$ via (2), which is referred to as the SINR Calculation Module in Fig. 1. Meantime, the transmit power $P(\{\mathbf{t}_1, \ldots, \mathbf{t}_M\})$ is calculated via (3), which is referred to as the Power Calculation Module in Fig. 1. Finally the SINR and transmit powers are calculated via (11) to form the loss function.

In (11), the expectation calculation can be fulfilled by the arithmetic mean over a subset of all potential combinations of estimated CSIs and channel errors. With the assumption of independence between estimated CSIs and the channel errors, the training samples for the estimated CSIs and channel errors can be created independently and much reduced in size. In practice, we can sample the estimated CSIs into a set $\mathbb{S}_{\tilde{\mathbf{h}}}$ and the potential channel errors into a set $\mathbb{S}_{\Delta\mathbf{h}}$, either via the historical measurements from the real-world system, or the random samples from the known distributions.

During the training procedure, both the $\mathbb{S}_{\tilde{\mathbf{h}}}$ and $\mathbb{S}_{\Delta\mathbf{h}}$ can be divided into batch-wise subsets $\mathcal{S}_{\tilde{\mathbf{h}}}$ and $\mathcal{S}_{\Delta\mathbf{h}}$, respectively. Then the augmented data set $\mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}}$ can be created via $\mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}} = \{\mathbf{h}_i | \mathbf{h}_i = \tilde{\mathbf{h}}_i + \Delta\mathbf{h}_j, \forall\tilde{\mathbf{h}}_i \in \mathcal{S}_{\tilde{\mathbf{h}}}, \forall\Delta\mathbf{h}_j \in \mathcal{S}_{\Delta\mathbf{h}}\}$.

Then at the training step t, the DNN parameter $\boldsymbol{\theta}$ and the dual parameters $\boldsymbol{\lambda}$ are updated using (13) and (14) at the top of this page based on the Gradient Decent (GD) method [10]. where the clamp operation is defined as $(x)_+ \triangleq \max\{x, 0\}$,

$$\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)} - \frac{\alpha_{\mathrm{DNN}}}{|\mathcal{S}_{\tilde{\mathbf{h}}}|} \sum_{\tilde{\mathbf{h}} \in \mathcal{S}_{\tilde{\mathbf{h}}}} \nabla_{\boldsymbol{\theta}} P(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}; \boldsymbol{\theta}^{(t-1)})) + \sum_{m=1}^{M} \frac{\alpha_{\mathrm{DNN}} \lambda_m^{(t-1)}}{|\mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}}|} \sum_{\mathbf{h} \in \mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}}} \nabla_{\boldsymbol{\theta}} \left(g_m(\psi(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}; \boldsymbol{\theta}^{(t-1)}), \mathbf{h}), \rho_{\max}) - \Gamma_m \right) \quad (13)$$

$$\lambda_m^{(t)} = \left(\lambda_m^{(t-1)} + \frac{\alpha_{\mathrm{L}} \lambda_m^{(t-1)}}{|\mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}}|} \sum_{\mathbf{h} \in \mathcal{S}_{\tilde{\mathbf{h}}+\Delta\mathbf{h}}} \left(g_m(\psi(f_{\boldsymbol{\theta}}(\tilde{\mathbf{h}}; \boldsymbol{\theta}^{(t-1)}), \mathbf{h}), \rho_{\max}) - \Gamma_m \right) \right)_+ , \quad (14)$$

α_{DNN} and α_{L} are the learning rates for DNN network and Lagrange dual parameters, respectively.

IV. Simulation

In order to validate the performance of the proposed data augmentation based DNN method for the robust beamforming problem, a scenario with a 2-antenna transmitter and two single-antenna users is considered, which is referred to as 2×2 scenario in the rest of the paper. In the simulation, the estimated CSIs and the channel errors are assumed to be mutually independent and follow the independent and identical distribution of (i.i.d.) standard circularly symmetric complex Gaussian distribution with the variance of 1 and 0.002, respectively [1]. The systems' statistical performance requirement is set as a maximum outage probability $\rho_{\max} = 0.05$ with an SINR threshold of 2 dB. In order to compare the performance of the proposed DNN based method with the state-of-the-art methods, the BTI method in [1] is used as the benchmark algorithm, which is solved with *SeDuMi* via *CVX* [11].

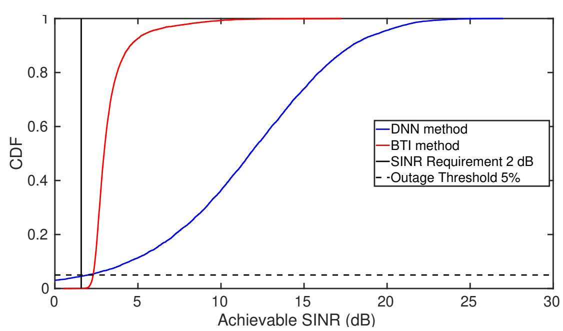

Fig. 2. The CDF performance comparison between the BTI method and the proposed DNN method for user 1 in a 2x2 scenario.

For the training purpose, a set of CSI estimation is randomly generated with a total size of 10^4, while the size of the channel error set is 4×10^5. The batch size for the estimated CSIs is set as 10^3, while the batch size for the channel errors is set as 10^5, which is equivalent to an augmentation ratio of 100 in Fig. 1. The complex channels are represented in real forms, where each estimated CSI and channel errors are reshaped to a 1×8 row vector with the imagine parts concatenated after the real parts.

A DNN network with 5 hidden layer is constructed, where all hidden layers are fully connected layers with width 200, and the PReLU activation function [12] is applied as the activation layer, which is defined as $\mathrm{PReLU}(x) = \max\{x, 0\}$, if $x \geq 0$, and $0.25x$ otherwise. The batch normalization [13] for each hidden layer and Adam algorithm [14] are used as the optimizer with a learning rate of 10^{-5} and an L2 penalty of 10^{-5}. The Lagrangian in (11) is used as the loss function for the unsupervised learning procedure as detailed in Fig. 1, with a learning rate of 10^{-2} for the Lagrange dual parameters. The models were trained with the open-source machine learning framework PyTorch on GPU Nvidia Titan RTX.

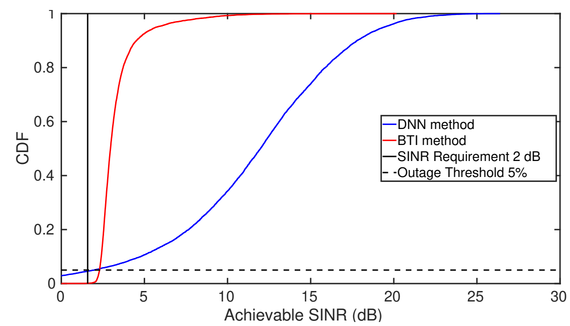

Fig. 3. The CDF performance comparison between the BTI method and the proposed DNN method for user 2 in a 2x2 scenario.

The trained models are evaluated against 10^4 randomly generated estimated CSIs and 10^4 randomly generated channel errors. The cumulative distribution function (CDF) of the achievable SINR under the considered 2×2 scenario are considered, where the results for User 1 and User 2 are shown in Fig. 2 and 3, respectively. In this considered scenario, the equivalent SINR threshold Γ_m for both users are the same as 2 dB, which is illustrated as the vertical solid line in the figures. It can be seen that both the proposed data augmentation based DNN method and the BTI method have achieved the robust beamforming objective, where the CDF at 2 dB are both below the designed goal of 5% as indicated by the horizontal dashed line. Compared to the proposed DNN based method as indicated in Fig. 2 and 3, the BTI method has a much lower probability than the design threshold of 5% outage below 2dB. As detailed in Table I, the proposed DNN method also shows a 1.63 dB average transmit power performance saving against the BTI method, which indicates that BTI provides much conservative performance at the cost of more average consumed power as compared to the proposed DNN method. This demonstrates that the proposed DNN method achieves a better performance with regard to the outage-constrained robust beamforming targets, where the objective is to minimize the average transmit power. It is also noticed that the CDF performances for User 1 and User 2 are sharing similar trends, which demonstrates that the proposed DNN method is capable

TABLE I
TIME AND AVERAGE TRANSMIT POWER PERFORMANCE FOR THE COMPARED METHODS

	Average Time (s)	Min. Time (s)	Max. Time (s)	Average Transmit Power
DNN method	3.23e-4	3.17e-4	4.00e-4	12.47 dB
BTI method [1]	2.65e-1	2.46e-1	3.23e-1	14.10 dB

to provide a fair performance trade-off between the two users.

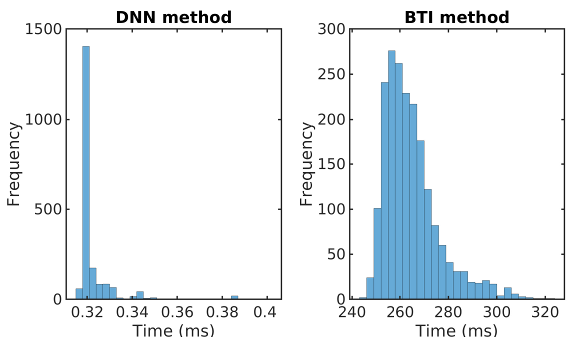

Fig. 4. The frequency of the 2000 execution time performance, compared between the proposed DNN method and the BTI method.

The executing time for the proposed DNN method and the benchmarking BTI method are also investigated, which is of great importance to the time-critical applications in the 5G and beyond networks. For a fair comparison, both methods exploit the same set of randomly generated 2000 estimated CSIs as inputs, and executed on an Intel Core i7 CPU instead of GPU. The statistical results are presented in Fig. 4, while the details of the minimum, average and maximum time to complete the algorithms, as well as the average transmit power, are reported in Table I. It can be seen that the minimum time for the BTI method to complete the beamforming calculation is 246 ms, while in contrast, the maximum execution time for the proposed DNN method is 0.4 ms. On average, our proposed DNN methods can reduce the executing time of the BTI method by 800 times.

From Fig. 4, it can be seen that the execution time for the proposed DNN based method has less variation compared to that for the BTI method. This can be indicated by two aspects, a) the variation range of the proposed DNN based method is within 0.1 ms, while that of the BTI method is more than 70 ms, and b) the execution time for the proposed data augmentation based DNN method is more concentrated to its mean value (0.32 ms), while the time performance for BTI method is more diverse around its mean value (265 ms) as indicated in Fig. 4. This can be explained by the formulation difference between the BTI method via (6) and the proposed DNN based method via (7), because the numerical searching procedure in BTI method varies with different estimated CSIs, but the proposed data augmentation based DNN method applies the same calculation procedure to all possible estimated CSIs. This indicates a more reliable executing time performance of the proposed data augmentation based DNN method compared to the BTI method.

V. CONCLUSION

In this paper, the outage-constrained robust beamforming problem was studied in multi-user multi-antenna wireless communications systems, where the time complexity and statistical outage performance were addressed simultaneously. The robust beamforming problem was transformed to a deep learning problem by applying Lagrange dual method, and a novel augmentation-based DNN training approach was proposed. The simulation results verified that the proposed data augmentation based DNN method satisfied the outage-constrained robust beamforming targets, while it outperformed the benchmarking robust beamforming algorithms in terms of executing time and average transmit power performance.

REFERENCES

[1] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Transactions on Signal Processing*, vol. 62, pp. 5690–5705, Nov. 2014.
[2] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 609028, Jul. 2008.
[3] Y. Yuan, P. Xu, Z. Yang, Z. Ding, and Q. Chen, "Joint robust beamforming and power-splitting ratio design in SWIPT-based cooperative NOMA systems with CSI uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 68, pp. 2386–2400, Mar. 2019.
[4] M. R. Khandaker, K.-K. Wong, Y. Zhang, and Z. Zheng, "Probabilistically robust SWIPT for secrecy MISOME systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 211–226, Sept. 2016.
[5] X. Zhang and M. R. Nakhai, "Robust chance-constrained distributed beamforming for multicell interference networks," in *IEEE International Conference on Communications, Kuala Lumpur, Malaysia*, pp. 1–6, 2016.
[6] 3GPP TR 38.913, "Study on scenarios and requirements for next generation access technologies," June 2016.
[7] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1165–1175, Jul. 2004.
[8] L. Adam and M. Branda, "Machine learning approach to chance-constrained problems: An algorithm based on the stochastic gradient descent," *arXiv preprint arXiv:1905.10986*, 2019.
[9] H. Lee, S. H. Lee, and T. Q. S. Quek, "Constrained deep learning for wireless resource management," in *IEEE ICC*, pp. 1–6, May 2019.
[10] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in *International conference on machine learning*, pp. 1139–1147, 2013.
[11] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, pp. 95–110, Springer-Verlag Limited, 2008.
[12] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.
[13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv:1502.03167*, 2015.
[14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, USA*, pp. 1–15, May 2015.