Deep Learning Based Beamforming for MISO Systems with Dirty-Paper Coding

Xingliang Lou¹, Wenchao Xia¹, Zhao Haitao¹, Wanli Wen², Xiaohui Li³, and Bin Wang⁴

¹Nanjing University of Posts and Telecommunications
²Chongqing University
³Taiyuan University of Technology
⁴Hangzhou Hikvision Digital Technology Co Ltd

October 20, 2022

Abstract

Beamforming technique can effectively improve the spectrum utilization of multi-antenna systems, while the dirty-paper coding (DPC) technique can reduce inter-user interference. In this letter, we aim to maximize the weighted sum-rate under power constraint in a multiple-input-single-output (MISO) system with the DPC. However, the existing methods of beamforming optimization mainly rely on customized iterative algorithms, which have high computational complexity. To address this issue, by utilizing the deep learning technique and the uplink-downlink duality, and carefully exploring the optimal solution structure, we devise a beamforming neural network (BFNNet), which includes a deep neural network module and a signal processing module. Besides, we use the modulus of the channel coefficients as the input of deep neural network, which reduces the input size. Simulation results show that a well-trained BFNNet can achieve near-optimal solutions, while significantly reducing computational complexity.
Deep Learning Based Beamforming for MISO Systems with Dirty-Paper Coding

Xingliang Lou, Wenchao Xia, Wanli Wen, Haitao Zhao, Xiaohui Li and Bin Wang

Beamforming technique can effectively improve the spectrum utilization of multi-antenna systems, while the dirty-paper coding (DPC) technique can reduce inter-user interference. In this letter, we aim to maximize the weighted sum-rate under power constraint in a multiple-input-single-output (MISO) system with the DPC. However, the existing methods of beamforming optimization mainly rely on customized iterative algorithms, which have high computational complexity. To address this issue, by utilizing the deep learning technique and the uplink-downlink duality, and carefully exploring the optimal solution structure, we devise a beamforming neural network (BFNNet), which includes a deep neural network module and a signal processing module. Besides, we use the modulus of the channel coefficients as the input of deep neural network, which reduces the input size. Simulation results show that a well-trained BFNNet can achieve near-optimal solutions, while significantly reducing computational complexity.

Introduction: Beamforming technique can improve the spectrum efficiency of multi-antenna systems while the dirty-paper coding (DPC) technique [1] can reduce inter-user interference. Thus, beamforming strategies using the DPC technique are a potential way to maximize the weighted sum-rate under power constraint in a multi-antenna system. However, finding the optimal beamforming to maximize the weighted sum-rate is a non-convex problem. There have been some methods of beamforming design studied in existing literature. For example, the weighted minimum mean square error (WMMSE) algorithm was proposed in [2, 3]. Since the uplink-downlink duality was proved in [4], the downlink sum-rate maximization problem can be solved by considering the dual uplink problem. [5] has found the achievable rate of multi-antenna downlink, and [6, 7] have established the conversion structure to reduce the prediction complexity. Using the power budget as side information, [13] investigated the influence of power constraint on the received SINR at user i, when using the DPC. Because decoding/encoding is performed in sequence, the interference of user k (k > i) has no effect on the demodulated received SINR of user i. Thus, the received SINR at user i can be expressed as

$$\text{SINR}_{i}^{\text{DL}} = \frac{|h_i^T u_i|^2}{\sum_{k=1}^{K} |h_k^T u_i|^2 + \sigma^2}$$  (2)

Define $U = [u_1, u_2, \ldots, u_K]$, then the downlink sum-rate maximization problem under power constraint is formulated as

$$\textbf{P1:} \max_{U} \sum_{i=1}^{K} \log_2(1 + \text{SINR}_{i}^{\text{DL}})$$

s.t. $|u_i|^2 \leq P_m$,  (3)

where $P_m$ is the power budget. Note that, $\textbf{P1}$ is a challenging non-convex problem, which can be solved using the WMMSE algorithm or the uplink-downlink duality based algorithms [6, 7]. But these algorithms relying on iterative processes are difficult to meet the implementation requirements. Thus, we propose to solve it using a DL-based beamforming framework, which will be described in next Sections.

Expert Knowledge: Before giving the DL based beamforming framework, we first establish a concept of expert knowledge [4] for the purpose of reducing prediction complexity.

Lemma 1: The achievable uplink sum-rate is equal to the achievable downlink sum-rate, i.e.,

$$C_{\text{DL sum-r}}^{\text{DL}} = C_{\text{UL sum-r}}^{\text{DL}}$$

where

$$C_{\text{DL sum-r}}^{\text{DL}} = \max_{U_{\text{DL}}} \sum_{i=1}^{K} \log_2(1 + \text{SINR}_{i}^{\text{DL}})$$

s.t. $|p_i| \leq P_m$,

$$|u_i|^2 = 1, \forall i,$$  (5)

and

$$C_{\text{UL sum-r}}^{\text{DL}} = \max_{U_{\text{UL}}} \sum_{i=1}^{K} \log_2(1 + \text{SINR}_{i}^{\text{UL}})$$

s.t. $|q_i| \leq P_m$,

$$|u_i|^2 = 1, \forall i,$$  (6)

with

$$\text{SINR}_{i}^{\text{UL}} = \frac{|p_i h_i^T u_i|^2}{\sum_{k=1}^{K} |p_k h_k^T u_i|^2 + \sigma^2}$$

and

$$\text{SINR}_{i}^{\text{DL}} = \frac{|q_i h_i^T u_i|^2}{\sum_{k=1}^{K} |q_k h_k^T u_i|^2 + \sigma^2}$$  (8)

in which $U = [u_1, u_2, \ldots, u_K]$ is the normalized beamforming, $p = [p_1, \ldots, p_K]^T$ and $q = [q_1, \ldots, q_K]^T$ are downlink and uplink power allocation vectors, respectively.

Proof: The proof is similar to the proof of Theorem 2 in [4] and thus is omitted here. In addition, [4] also proves that the optimal normalized beamforming of the uplink is also optimal for the downlink.
Signal Processing

I

I

R

I

(h

(h

(h

(h

(h

(h

(h

(h

(h

1,1

1,2

K,M

1

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

\(K,M\)

\(1,1\)

\(1,2\)

\(K,M\)

\(B\)

\(\text{Layer} \)

\(\text{Input Layer} \)

\(\text{Fully Connected Layer} \)

\(\text{Batch Normalization Layer} \)

\(\text{Activation Layer} \)

\(\text{Queue} \)

\(\text{ConvNet Module} \)

\(\text{Signal Processing Module} \)

\(\text{Deep Neural Network Module} \)

\(\text{Chib} \)

\(U^*\)

\(R_{\text{sum}}(h) = \max_{\|q\|_2 \leq \rho_m} \log_2(1 + \sin(\text{NIR}_{p1}))\)

\(= \max_{\|q\|_2 \leq \rho_m} \log_2(1 + \frac{1}{\sigma} \sum_{k=1}^{K} q_k h_k h_k^H)\)  

(9)

where the second equation is obtained due to \(|\hat{u}_i^*|_2 = 1, \forall i\). The problem in (9) can be solved using the IWF algorithm [8] until convergence. Knowing the optimal \(q_i^*\), the optimal beamforming vectors are given as the MMSE solutions [3], i.e.,

\(\hat{u}_i^* = \frac{\sigma^2 1 + \sum_{k=1}^{K} q_k^* h_k h_k^H - 1 h_i^*}{\|\sigma^2 1 + \sum_{k=1}^{K} q_k^* h_k h_k^H - 1 h_i^*\|_2}\)  

(10)

Then, we can find the optimal power allocation vector \(p^*\) of the downlink problem in (5) according to the following lemma.

**Lemma 2:** Given the optimal transmit power vector \(q_i^*\) and beamforming matrix \(U^*\) of the uplink problem in (6), then we can obtain the optimal transmit power vector \(p^*\) of the downlink problem in (5) as

\[ p_i^* = \begin{cases} B_i^{-1/2} q_i^* B_i^{-1/2}, & i = 1,2, \ldots, K, \sum_{i=1}^{K} p_i^* = 1 \end{cases} \]

(11)

where \(B_i = \sigma^2 + \sum_{k=1}^{K} q_k^* h_k^H h_k h_k^H - 1 \) represent the interference experienced by user \(i\) in the uplink and the interference experienced by user \(i\) in the downlink, respectively.

**Proof:** The achievable rate of user \(i\) in the uplink is given by

\[ R_{UL}^i = \log_2(1 + B_i^{-1/2} q_i^* h_i^H h_i h_i^H q_i^*) \]

(12)

Using matrix knowledge, we have the simplified formula as

\[ R_{UL}^i = \log_2(1 + B_i^{-1/2} q_i^* h_i^H h_i A_i^{-1/2} A_i^{1/2} q_i^* A_i^{1/2} B_i^{-1/2} q_i^*) \]

Treating \(B_i^{-1/2} q_i^* h_i^H h_i A_i^{-1/2}\) as the effective channel of the system, we flip the channel and get

\[ R_{UL}^i = \log_2(1 + A_i^{-1/2} h_i^H q_i^* B_i^{-1/2} A_i^{1/2} q_i^* A_i^{1/2} B_i^{-1/2} h_i^H h_i A_i^{-1/2}) \]

(13)

Now, consider the achievable rate of user \(i\) in the downlink and we have

\[ R_{DL}^i = \log_2(1 + A_i^{-1/2} h_i^H q_i^* B_i^{-1/2} A_i^{1/2} q_i^* A_i^{1/2} B_i^{-1/2} h_i^H h_i A_i^{-1/2}) \]

(14)

By setting the downlink transmit power as in (11), then we have \(R_{DL}^i = R_{UL}^i\). [6] has proved that \(\sum_{i=1}^{K} q_i^* = \sum_{i=1}^{K} p_i^*\).

Note that \(p_i^*\) only depends on \(p_1^*, \ldots, p_{i-1}^*\), thus the transmit power can be calculated sequentially in ascending order.

**BFNNet Structure:** The proposed BFNNet for the sum-rate maximization problem is shown in Fig. 1, which includes a deep neural network module and a signal processing module.

The deep neural network module includes an input layer, multiple hidden layers, and an output layer. The channel coefficients

\[ \text{Fig. 1. BFNNet for the sum-rate maximization problem.} \]

**Simulation Results:** In this section, we use the scene in [12] to conduct some numerical simulations to evaluate the performance of the proposed BFNNet. In order to train the deep neural network module, we use the IWF algorithm [8] to generate 20000 training samples and 5000 testing samples, respectively. In our simulations, We use a network with three hidden layers, one input layer and one output layer for the deep neural network module. The first hidden layer contains 256 neurons and the second hidden layer contains 128 neurons and the third hidden layer contains 64 neurons. For comparison, several baseline solutions are introduced, including zero-forcing (ZF) beamforming, regularized ZF (RZF) beamforming [14], and the WMMSE algorithm with the RZF initialization [2]. Moreover, the DPC used for all the baseline solutions.

**Fig. 2** Sum-rate performance averaged over 5000 samples under

\[ \{ K = 4, M = 4 \}. \]

Fig. 2 shows that with the increase of normalized transmission power, the sum-rate performance of all solutions are improved. We observe that the performance of the BFNNet solution is very close to the IWF algorithm, and better than the WMMSE algorithm which can find the locally optimal solution to the sum-rate maximization problem. This is
because that the BFNet is trained using the samples generated by the IWF algorithm which can achieve the optimal solution to problem $\textbf{P1}$.

Fig. 3 shows the sum-rate performance of the five beamforming solutions, where $P_m = 30$ dBm and $M = K$. In Fig. 3(a), as the number of transmitting antennas increases, the sum-rate performance of the five schemes increases at the same time and the BFNet solution outperforms the other solutions except the IWF algorithm. In addition, as the number of transmit antennas increases, the performance gap becomes greater.

In Fig. 3(b), the computational complexity, in terms of the execution time, of the BFNet solution is higher than that of the ZF beamforming solution as well as the RZF beamforming solution. The reason is that ZF beamforming and RZF beamforming solutions do not require any iterative process, the BFNet solution needs to perform neural network iterations. The above observations validate that the BFNet solution provides a good balance between system performance and computational complexity.

**Conclusion:** In this letter, we considered the MISO system with the DPC and formulated the problem of sum-rate maximization under a total power constraint.

**Fig. 3** Comparison of five different beamforming solution: (a) sum-rate performance and (b) execution time of each sample averaged over 5000 samples under $[K = M, P_m = 30$ dBm].

**Acknowledgments:** This work was supported by the National Key Research and Development Program (No. 2020YFB1806608); the Natural Science Foundation on Frontier Technology Basic Research Project of Jiangsu under Grant BK20212001; the National Natural Science Foundation of China (No. 92067201, 61871446); the Natural Science Research Project of Jiangsu Higher Education Institutions (No. 21KJB510034, 21KJB510027); Future Network Scientific Research Fund Project (FNSRF-2021-ZD-8, FNSRF-2021-YB-31).

X. Lou, W. Xia, H. Zhan/Jiangsu Key Laboratory of Wireless Communications, Nanjing University of Posts and Telecommunications, Nanjing, China

W. Wen (The School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China)

X. Li (Taiyuan University of Technology, Taiyuan, China)

B. Wang (College of Electrical Engineering, Zhejiang University, Hangzhou, China)

E-mail: zhaoht@njupt.edu.cn, wangbin2@hikvision.com

**References**


