Towards a Learning-Only Approach for Non-Convex Sum-Rate Maximization

QINGYU SONG, The Chinese University of Hong Kong, China
GUOCHEN LIU, Huawei Technologies Co., Ltd., China
HONG XU, The Chinese University of Hong Kong, China

In this study, we try to further improve optimality of non-convex optimization problems by exploring machine learning-only approaches. We propose a general framework We introduce several white-box models within black-box neural networks to enhance learning efficiency and achieve iteration-free implementation. We evaluate the accuracy and efficiency of the proposed model using simulated data of varying magnitudes. Our model outperforms all other learning-only state-of-the-art (SOTA) approaches and achieves faster inference speeds than other learning-augmented SOTAs.

1 INTRODUCTION

Optimization Problem. This study investigates a classical non-convex optimization problem in wireless communication, namely the sum-rate maximization problem [5] (WMMSE paper). Our goal is to solve the following maximization problem to reduce interference between users in shared frequency subcarriers. Specifically, we consider a cellular network comprising of B base stations (BS) each with N_t transmitting antennas. Each base station serves U users (UE), each equipped with N_r receiving antennas. The channel (coefficient) between UE u and BS b is denoted by matrix $\mathbf{H}_{b,u}$, while the precoding (objective variable) is denoted by matrix $\mathbf{V}_{b,u}$. We optimize each precoding $\mathbf{V}$ with given channels $\mathbf{H}$ and total power P to maximize the sum rate of overall users. Notably, the non-convexity in the objective function is caused by the inversity of the outer product, which reflects the interference among all users, including intra-BS and inter-BS interference. As established in [5], this problem is known to be NP-hard.

$$\max_{\mathbf{V}} \sum_{b=1}^{B}\sum_{u=1}^{U} \log_2 \det \left[\mathbf{I}_{N_r} + \left(\mathbf{H}_{b,u}\mathbf{V}_{b,u}\right)\left(\mathbf{H}_{b,u}\mathbf{V}_{b,u}\right)^H \cdot \left[\sum_{\tilde{b},\tilde{u};\tilde{b},\tilde{u}\neq b,u}^{B,U} \left(\mathbf{H}_{\tilde{b},u}\mathbf{V}_{\tilde{b},\tilde{u}}\right)\left(\mathbf{H}_{\tilde{b},u}\mathbf{V}_{\tilde{b},\tilde{u}}\right)^H + \sigma^2 \mathbf{I}_{N_r} \right]^{-1} \right], \quad (1)$$
$$\text{s.t.} \sum_{u=1}^{U} \mathrm{Tr}\left(\mathbf{V}_{b,u}\mathbf{V}_{b,u}^H\right) \le P, b = 1, 2, \ldots, B.$$

Numerical Optimization Algorithm. The Weighted Mean-Square Error (WMMSE) algorithm [5] is a widely employed iterative numerical algorithm for problem (1). Specifically, WMMSE formulates a dual problem with two auxiliary variables, making it numerically solvable by applying a blocked coordinate descent algorithm. We benchmark our method's performance against WMMSE as the state-of-the-art.

Learning-assisted Algorithms. Many *learning-augmented* approaches have been developed based on WMMSE [1, 2], achieving optimality levels around 90% with a fixed number of iterations. They are, however, limited by WMMSE's constraints. First, there is no robustness guarantee as WMMSE requires random initialization. Although the latest version of WMMSE [6] suggests using Zero-Forcing with SVD decomposition, it remains non-deterministic. Second, the optimality of these approaches is upper-bounded by WMMSE by nature.

In response to the limitations above, *learning-only* methods have also been studied recently [3, 4]. Nonetheless, most of them simplify the problem in (1) to a rudimentary power allocation problem on scalar variables. Their performance is inadequate for the original optimization scenario. We aim to advance this approach by exploring its potential beyond

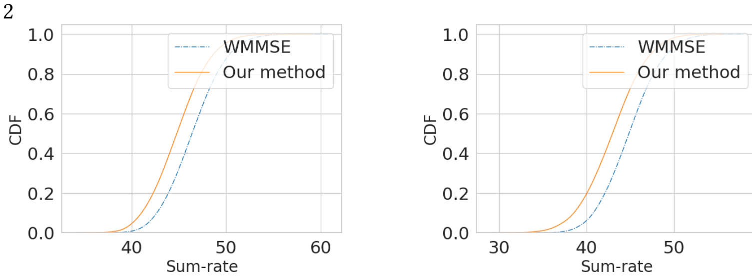

Fig. 1. Cumulative Distribution Function of Optimal Objective Values Across Varied Inter-Cell Interference Scenarios. Left: Small Interference. Right: Large Interference.

Table 1. Comparing Percentage of Optimal Objective Value to WMMSE across Different Scales of Matrix Variables (Columns, (Small (8 8-by-2), Moderate (16 16-by-2), and Large (32 32-by-2)) on Varied Inter-Cell Interference Scenarios (Rows).

	Small	Moderate	Large
Small	95.6%	96.5%	93%
Large	95.3%	95.5%	91.5%

the current scope. As will be shown soon, our learning-only method outperforms all existing work in this category, and offers comparable optimality and efficiency to learning-augmented algorithms.

2 PRELIMINARY DESIGN AND EVALUATION

Structural Learning Framework. We aim to improve the sum-rate performance by emulating the first-order derivative condition in WMMSE through a neural network model. WMMSE derives a local optimum as an inverse matrix dot-product with a coefficient column vector. Our approach interprets the inverse matrix as a non-unitary 'basis' and the coefficient vector as a corresponding coefficient. We propose a learning framework to construct these bases and coefficients. Similar to other SOTAs [2, 3], we utilize a fixed number of iterative procedures within neural networks, which significantly reduces model size. Further, our model incorporates pooling techniques and message-passing networks, making it scalable to different scenario scales.

Virtual Variables. It is widely recognized that when the number of rows of the matrix $\mathbf{V}$ is greater than or equal to $N_r \times U_b \times B$, the problem (1) can be solved using singular value decomposition [5]. This implies that when the number of transmitting antennas is sufficiently large, all interference can be eliminated. For learning efficiency, we employ a learnable parameter matrix to linearly construct virtual high-dimensional variables that map to solvable dimensions. The linearity of this approach preserves all inner product directions in (1) and also enables the retrieval of the original variable from the inner product with the Hermitian of the parameter matrix.

We comprehensively evaluate our proposed framework in simulated scenarios of varying scales, i.e., small, moderate, and large scenarios, with eight, sixteen, and thirty-two variables arranged on 8-by-2, 16-by-2, and 32-by-2 matrix spaces. The number of parameters was scaled accordingly for each scenario. The number of parameters was scaled accordingly for each scenario. We followed the method outlined in WMMSE [5] for all other configurations. Additionally, we introduce a scenario with significant inter-cell interference. Currently, our results demonstrate over 90% optimality compared to WMMSE. The outcomes of our evaluation are assessed using randomly generated testing data, and the results are presented in Figure 1 and Table 1.

REFERENCES

[1] Q. Hu, Y. Cai, Q. Shi, K. Xu, G. Yu, and Z. Ding. Iterative Algorithm Induced Deep-Unfolding Neural Networks: Precoding Design for Multiuser MIMO Systems. *IEEE Trans. Wirel. Commun.*, 20(2):1394–1410, 2021.

[2] L. Schynol and M. Pesavento. Deep Unfolding in Multicell MU-MIMO. In *2022 30th European Signal Processing Conference (EUSIPCO)*, 2022.

[3] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief. Graph Neural Networks for Scalable Radio Resource Management: Architecture Design and Theoretical Analysis. *IEEE J. Sel. Areas Commun.*, 39(1):101–115, 2021.

[4] Y. Shen, J. Zhang, S. Song, and K. B. Letaief. Graph Neural Networks for Wireless Communications: From Theory to Practice. *IEEE Trans. Wirel. Commun.*, pages 1–1, 2022.

[5] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. *IEEE Trans. Signal Process.*, 59(9):4331–4340, 2011.

[6] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo. Rethinking WMMSE: Can Its Complexity Scale Linearly With the Number of BS Antennas? *IEEE Trans. Signal Process.*, 71:433–446, 2023.