Performance Analysis of Near-Optimal Digital Precoding Algorithm for Massive MIMO Systems

Subitha D.

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
Email: subitha.d@vit.ac.in (S.D.)

***Abstract*—Precoder design for Massive Multiple-Input Multiple-Output (MIMO) systems is very important for improving efficiency and reliability. Computing the optimal precoder is a challenging task and many optimizations theory-assisted algorithms have been evolved to achieve the optimality. In general, achieving optimal performance is at the cost of complexity whereas heuristic beamforming is a technique that exploits the transmission scenario in its favor to simplify the problem. Here, we investigate the performance of linear precoders obtained by various heuristic techniques with the selection of parameters judicially. The existing heuristic algorithms such as Zero Forcing (ZF), Matched Filter (MF), and Minimum Mean Square Error (MMSE) are compared with the proposed heuristic algorithm. The Bit Error Rate (BER) performance in Urban Micro cell scenario for 128 × 4 and 64 × 4 massive MIMO System was simulated using WINNER II modelling with the above precoding algorithms. The experimental results on the above modelling has proven that the proposed heuristic approach performs better for larger antennas.**

***Keywords*—precoding, optimization, heuristic techniques**

I. Introduction

Precoding is a technique used in 4G and Beyond 5G wireless technologies to satisfy huge demands such as system capacity, spectral efficiency, and sum rate by allowing the transmission of multiple users simultaneously by suppressing interference among them [1, 2]. This is the so-called multiuser Multiple-Input Multiple-Output (MU-MIMO) system which is an integral part of all the upcoming wireless technologies. But the challenge here is the design of a precoding algorithm that satisfies the complexity performance trade-off which is the focus of this paper. To address this challenge, heuristic beamforming algorithms such as Zero Forcing (ZF), Regularized ZF (RZF), and Matched Filter (MF) [3, 4] have been the more commonly used precoding algorithms in recent wireless systems, but they have failed in terms of performance. Hence, hybrid precoding algorithms [5, 6] are preferred over the heuristic algorithms for their better performance. However, the optimality comes at the cost of increased complexity such as the higher dimensional matrix inversion and eigen value decomposition operations encountered in the computation of the precoding vector.

Various kinds of literature [7, 8] have obtained precoding matrix as the solution for different optimization strategies such as Weighted Sum Rate (WSR) maximization, transmit power minimization, and Signal Interference Noise Ratio (SINR) constraint. In [9], three different Beamforming Neural Network (BNN) architectures for each of the above problems have been proposed to solve the precoding weight vector. But the drawback here is the construction of the data set which uses complex iterative algorithms such as WMMSE Weighted Minimum Mean Square Error (MMSE) algorithm [10, 11].

Shi *et al*. [12] developed a Lagrangian Neural Network-based robust precoder that learns the Lagrangian multiplier (uplink power vector) from the channel vector. However, the drawback of this work is, only the Lagrangian multiplier is obtained from the neural network that computes only a part of the precoder matrix whereas the other parameters such as precoder directional vector and downlink power vector have to be solved using generalized eigenvalue problem, and closed-form expression [13] respectively. This involves again the high dimensional matrix inversion operations that are highly complex in real-time implementations. Koc *et al*. [14] proposed a deep neural network for the computation of a downlink power vector from the channel matrix. This work has trained the neural network using the Particle Swarm Optimization (PSO) based power allocation algorithm for downlink power vector calculation. Another deep neural network proposed in [15] trains the network to learn the power vector from the channel matrix. This paper also addresses the computation of power vector alone whereas the computation of other parameters is left uncovered. Hence, the above-discussed works have been focused only on the computation of either the downlink power vector or the virtual uplink vector (Lagrangian multiplier) but not on both except the work in [9]. But this work also suffers from certain computational drawbacks discussed above.

Manuscript received January 13, 2024; revised March 11, 2024; accepted April 1, 2024; published July 22, 2024.

doi:10.12720/jcm.19.7.325-330

Overall, rather than opting out of the neural network structures that contribute only a part of the complete precoder design problem, it would be better to perform the computations using the conventional approach.

In this paper we have analyzed the problem of precoder design with two different objective functions; maximization of sum rate and minimization of power. The problem of power minimization is chosen as it is convex and the solution of this objective function is the optimal precoder vector. The optimal precoder vector computation is further simplified by applying the orthogonal property of the MIMO channel under symmetric channel conditions. The proposed low-complexity heuristic algorithm overcomes the drawback of complex large-scale matrix inverse operations required for the computation of precoder matrix by the method explained below. The low complexity algorithm is validated under urban microcell scenario using WINNER II channel modelling. This paper is structured as follows: Section II presents the system design and problem formulation. The Proposed low complexity algorithm for near-optimal precoder is presented in section III. The MATLAB simulation results and performance analysis is discussed in Section IV. Finally, the concluding remarks and future work is included in Section V.

II. System Design

Let us consider a model of a downlink system with multiple transmitting antennas at the Base Station (BS) equipped with N_t antennas serving K single-antenna users. The channel vector between the k^{th} user and the multiple BS antennas is denoted as $\boldsymbol{h}_k \in C^{N_t \times 1}$. The transmitted symbol after modulation is denoted as satisfying the constraint $E\{\| x_k |^2\} = 1$.

The signal received at the k^{th} user is given by

$$\boldsymbol{y}_k = \boldsymbol{h}_k^H \boldsymbol{p}_k \boldsymbol{x}_k + \underbrace{\sum_{j \neq k}^{K} \boldsymbol{h}_k^H \boldsymbol{p}_i \boldsymbol{x}_j + n_k}_{Interference + noise} \tag{1}$$

where, $\boldsymbol{p}_k \in C^{N_t \times 1}$ is the precoder (beamformer) vector that maps the input transmitted symbol of each user with the output transmitted symbol vector of the antenna array. In (1) the first term refers to the intended signal of user k and the other two terms refer to the interference due to the adjacent users and channel noise of Gaussian distribution with variance σ_n^2 respectively. Hence, the interference has a variance equal to $r_k = \sigma_n^2 + \sum_{i \neq k}^{K} E\{\boldsymbol{h}_k^H \boldsymbol{p}_i \boldsymbol{p}_i^H \boldsymbol{h}_k\}$.

A. Problem Formulation- Problem 1

The objective of the system is to maximize the WSR r_k which is the function of SINR of all the k users. This maximization is achieved by the selection of an optimal precoding weight vector which is subject to the power constraint.

$$\max_{\boldsymbol{p}_k} \sum_{k=1}^{K} r_k = \max_{\boldsymbol{p}_k} \sum_{k=1}^{K} \log(1 + SINR_k) \tag{2}$$

s.t. to $\sum_{k=1}^{K} \| \boldsymbol{p}_k \|^2 \leq P$ *for* k=1, 2,..., K where $SINR$ of the k^{th} user is given by

$$SINR_k = \frac{\boldsymbol{p}_k^H \boldsymbol{R}_k \boldsymbol{p}_k}{\sigma^2 + \sum_{i \neq k}^{K} \boldsymbol{p}_i^H R_k \boldsymbol{p}_i} \tag{3}$$

where $\boldsymbol{R}_k = E\left(\boldsymbol{h}_k \boldsymbol{h}_k^H\right) \in C^{N_t \times N_t}$

The objective function defined in Problem 1 is non-convex and hence it is difficult to obtain an exact solution for this problem. Only iterative algorithms can be used to solve this non-convex optimization function [16]. To reduce the higher computational complexity incurred by the iterative algorithms, this high dimensional problem is transformed into a lower dimensional space as given below.

B. Problem Formulation- Problem 2

The non-convex problem 1 is reformulated into a convex Problem 2 with the objective function of power minimization and constraint on SINR as given in Eq. (4) [12]. As we have kept reasonably large SINR values for all users such that the value exceeds the threshold value, it intuitively maximizes the WSR stated in problem 1. The reformulated power minimization problem is given as

$$\min \sum_{k=1}^{K} \boldsymbol{p}_k^H \boldsymbol{p}_k \tag{4}$$

s.t. to for k=*1,2..., K.*

The above constraint is denoted as $C_k \leqslant 0$, where C_k is written as

$$C_k = \gamma_k - SINR_k \leq 0 \tag{5}$$

The above problem is transformed into a Lagrangian problem using a Lagrangian multiplier μ_k

$$L_R = \sum_{k=1}^{K} \boldsymbol{p}_k^H \boldsymbol{p}_k + \sum_{k=1}^{K} \boldsymbol{\mu}_k C_k \tag{6}$$

Substituting Eq. (5) in Eq. (6), we get

$$L_R = \sum_{k=1}^{K} \| \boldsymbol{p}_k \|^2 + \sum_{k=1}^{K} \boldsymbol{\mu}_k \left(\sum_{i \neq k} \frac{1}{\gamma_i \sigma^2} \boldsymbol{p}_i^H \boldsymbol{R}_i \boldsymbol{p}_i + 1 - \frac{1}{\gamma_k \sigma^2} \boldsymbol{p}_k^H \boldsymbol{R}_k \boldsymbol{p}_k \right) \tag{7}$$

The above convex optimal condition is solved using the three Karush-KuhnTucker (KKT) conditions given below

$$\frac{\delta L_R}{\delta \boldsymbol{p}_k} = 0, \quad \text{for } k = 1,\dots,K \qquad (8)$$
$$\mu_k C_k = 0, \mu_k \geq 0$$

The optimal precoding vector is obtained by solving the above conditions. The precoding vector which $\tilde{\boldsymbol{p}}_k$ is the normalized beamforming directional vector satisfying the condition $\boldsymbol{\rho}_k$ is the beamforming power vector. The above two components of the precoder are calculated separately by solving the KKT conditions in Eq. (8). Differentiating LR and substituting the first KKT condition in Eq. (7) we obtain

$$\boldsymbol{p}_k + \sum_{i \neq k} \frac{\boldsymbol{\mu}_i}{\sigma^2} \boldsymbol{p}_k \boldsymbol{R}_k \boldsymbol{p}_i - \frac{\boldsymbol{\mu}_k}{\gamma_k \sigma^2} \boldsymbol{R}_k \boldsymbol{p}_k = 0 \qquad (9)$$

$$\Leftrightarrow \left(I_N + \sum_{i=1}^{K} \frac{\boldsymbol{\mu}_i}{\sigma^2} \boldsymbol{R}_k \boldsymbol{p}_i \right) \boldsymbol{p}_k = \frac{\boldsymbol{\mu}_k}{\sigma^2} \left(1 + \frac{1}{\gamma_k} \right) \boldsymbol{R}_k \boldsymbol{p}_k$$

From Eq. (9), we get

$$\boldsymbol{p}_k = \underbrace{\left(I_N + \sum_{i=1}^{K} \frac{\boldsymbol{\mu}_i}{\sigma^2} \boldsymbol{h}_i \boldsymbol{h}_i^H \right)^{-1} \boldsymbol{h}_k}_{precodingd\,irection} \underbrace{\frac{\boldsymbol{\mu}_i}{\sigma^2} \left(1 + \frac{1}{\gamma_k} \right) \boldsymbol{h}_k^H \boldsymbol{p}_k}_{scalar} \qquad (10)$$

From Eq. (10), the optimal beamforming direction is obtained. Hence, the beamforming vector is represented as given below:

$$\boldsymbol{p}_k = \underbrace{\sqrt{\boldsymbol{\rho}_k}}_{precoding\ power} \underbrace{\frac{\left(I_N + \sum_{i=1}^{K} \frac{\boldsymbol{\mu}_i}{\sigma^2} \boldsymbol{h}_i \boldsymbol{h}_i^H \right)^{-1} \boldsymbol{h}_k}{\left\| \left(I_N + \sum_{i=1}^{K} \frac{\boldsymbol{\mu}_i}{\sigma^2} \boldsymbol{h}_i \boldsymbol{h}_i^H \right)^{-1} \boldsymbol{h}_k \right\|}}_{=normalized\,precoding\ direction} \qquad (11)$$

The next step is the calculation of the beamforming power vector $\boldsymbol{\rho}_k$. This is obtained by taking the second KKT condition given in Eq. (8) → $C_k = 0$, and substituting Eq. (3) and Eq. (5) in this condition, we get

$$\frac{1}{\gamma_k} \tilde{\boldsymbol{p}}_k^H \boldsymbol{R}_k \tilde{\boldsymbol{p}}_k \rho_k - \sum_{i \neq k} \tilde{\boldsymbol{p}}_i^H \boldsymbol{R}_k \tilde{\boldsymbol{p}}_i \rho_i = \sigma^2 \text{ for } k = 1,\dots,K \qquad (12)$$

Representing Eq. (12) in matrix form to get an extended covariance matrix as below

$$[\boldsymbol{T}]_{ki} = t_{ki} = \begin{cases} \frac{1}{\gamma_k} \tilde{\boldsymbol{p}}_i^H \boldsymbol{R}_k \tilde{\boldsymbol{p}}_i & k = i, \\ -\tilde{\boldsymbol{p}}_i^H \boldsymbol{R}_k \tilde{\boldsymbol{p}}_i & k \neq i. \end{cases} \qquad (13)$$

Using the above matrix form, Eq. (12) is rewritten as:

$$\sum_{k=1}^{K} t_{ki} \rho_k = \sigma^2$$
$$\Rightarrow \boldsymbol{T\rho} = \sigma^2 [1]_{k \times 1} \qquad (14)$$

where the downlink power vector is $\boldsymbol{\rho} = [\rho_1\ \rho_2 \dots \rho_K]^T$ Rewriting Eq. (13) using Eq. (14) as:

$$\sum_{k=1}^{K} t_{ik} \mu_k = \sigma^2$$
$$\Rightarrow \boldsymbol{\mu} = \sigma^2 (\boldsymbol{T}^{-1})^H [1]_{k \times 1} \qquad (15)$$

where virtual uplink power parameter $\boldsymbol{\mu} = [\mu_1\ \mu_2 \dots \mu_K]^T$.

Using Eq. (14) and Eq. (15), we can find the near-optimal downlink and uplink power vectors respectively. Subsequently using Eq. (12) the precoder weight vector can also be obtained.

III. Proposed Low Complexity Algorithm for Near-Optimal Precoder Design

As discussed above, to compute the precoder vector it is essential to have both downlink and virtual uplink power vectors. However, solving them from extended covariance matrix $\boldsymbol{T}$ using Eq. (14) and Eq. (15) is highly complex as it requires the complex matrix inversion operation. In this section, we are going to develop a low-complexity algorithm that achieves near-optimal performance by the assumption of these two parameters $\boldsymbol{\mu}$ (also referred to as a Lagrangian parameter in some literature) and judiciously rather than computing them in a complex manner.

A. Symmetric Channel Conditions

Consider a symmetric channel where all the virtual channels between the BS antennas and user terminals are equally strong and well-separated beam directivity. In such scenarios where the channels are assumed to be equal and independent we can have a single value for the uplink power parameter ($\boldsymbol{\mu}$) for all the K users rather than having K independent parameters, i.e., $\mu_1 = \mu_2 = \dots = \mu_K$. These channels exhibit symmetric properties only when the number of transmitting antennas is greater which in turn makes the channels highly orthogonal and independent of each other [17]. Hence, for the massive MIMO downlink scenario with a large number of transmitting antennas in the order of Hundreds at the BS we have applied a heuristic beamforming approach using which we let the Lagrangian multiplier μ_k for *k=1,2…K* equal to the average transmit power per user, i.e., $\mu_k = \frac{P}{K}$ since $\sum_{i=1}^{K} \boldsymbol{\mu}_k = P$, where P is the total power constraint for K number of users. Hence, this approach of Heuristic beamforming reduces the complexity of computation by taking a single value of the Lagrangian multiplier rather the having all K degrees of values to find the optimal solution.

Next is the computation of the downlink power vector Rho which can be found in Eq. (14). But this requires an inverse of high dimensional complex matrix $\boldsymbol{T}$, a challenging task as far as the hardware implementation is concerned. To overcome these drawbacks, there is a direct method of finding $\boldsymbol{\mu}$ as given below. The value $\mu_k = \frac{P}{K}$ obtained in the previous step is substituted in (15) and so we obtain

$$(\boldsymbol{T}^{-1})^H [1.......1]^T = \frac{1}{\sigma^2}\left[\frac{P}{K}.......\frac{P}{K}\right]^T \tag{16}$$

Let $(\boldsymbol{T}^{-1})^H = \boldsymbol{M}$

Upon solving the above equation, we get

$$\boldsymbol{M} = \begin{bmatrix} \frac{P}{\sigma^2 K} & 0 & \cdots\cdots & 0 \\ 0 & \frac{P}{\sigma^2 K} & \cdots\cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots\cdots & \frac{P}{\sigma^2 K} \end{bmatrix} = \boldsymbol{T}^{-1} \tag{17}$$

We get it since P, and K, σ^2 are all real. From Eq. (14) we can now directly compute the downlink power vector as follows

$$\boldsymbol{\rho} = \boldsymbol{T}^{-1}\sigma^2[1]_{K\times 1} = \left[\frac{P}{K}.......\frac{P}{K}\right]^T, \Rightarrow \quad \rho_k = \frac{P}{K} \tag{18}$$

Finally, we have for k=1, 2, ... K.

B. Asymmetric Channel Conditions

Under asymmetric channel conditions that happen when the number of transmitting antennas is not very high, the above-explained heuristic approach would fail. Hence, the most general way to start computation is from the downlink power vector based on per-antenna power constraint and total power constraint [18]. We, now recall power constraint P given in Eq. (2), $\sum_{k=1}^{K} \| \boldsymbol{p}_k \|^2 \leq P$ and we also know that, and hence we can write $\sum_{k=1}^{K} \| \sqrt{\boldsymbol{\rho}_k}\, \tilde{\boldsymbol{p}}_k \|^2 \leq P$.

Under a symmetric channel scenario, all are equal as the gains of all the independent channels are symmetric $(\rho_k = \frac{P}{K})$. However, for the asymmetric channel condition, the downlink power parameters are unequal concerning the L1 norm of the corresponding precoder vector for k=1, 2, ... K as given below

$$\boldsymbol{\rho} = [\rho_1 \ \rho_2 \ \ldots \ \rho_K]^T = \left[\frac{\sqrt{P/K}}{\| p_1 \|} \cdots \frac{\sqrt{P/K}}{\| p_K \|}\right] \tag{19}$$

Next is the computation of the $\boldsymbol{T}$ matrix from which the finding using Eq. (15). requires complex matrix inversion operation and so we conclude that the former method of parameter computation is less complex under the symmetric channel scenario of massive MIMO system.

The above two sections explained the ways of finding downlink and virtual uplink power parameters with low computational complexity. Next is the evaluation of the normalized precoder vector from Eq. (11).

$$\tilde{\boldsymbol{p}}_k = \left(I_N + \sum_{i=1}^{K} \frac{\mu_i}{\sigma^2} \boldsymbol{h}_i \boldsymbol{h}_i^H\right)^{-1} \boldsymbol{h}_k = \left(I_N + \sum_{i=1}^{K} \frac{P}{\sigma^2 K} \boldsymbol{h}_i \boldsymbol{h}_i^H\right)^{-1} \boldsymbol{h}_k \tag{20}$$

This requires the inversion of the matrix in the order of $\boldsymbol{N}_t \times \boldsymbol{N}_t$ as $\boldsymbol{h}_k \in C^{N_t \times 1}$ which is computationally intensive as the number of transmitting antennas in a massive MIMO system is in the order of Hundreds. To simplify this task, the following matrix identity is applied in Eq. (20).

$$(\boldsymbol{I} + \boldsymbol{AB})^{-1}\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{BA})^{-1} \tag{21}$$

Rewriting Eq. (20), we get the expression with a reduced matrix size of the order 1×1 to be inverted as given in Eq. (22)

$$\left(I_{\boldsymbol{N}t\times\boldsymbol{N}t} + \sum_{i=1}^{K} \frac{P}{\sigma^2 K} \boldsymbol{h}_i \boldsymbol{h}_i^H\right)^{-1} \boldsymbol{h}_k = \boldsymbol{h}_k \left(I_1 + \sum_{i=1}^{K} \frac{P}{\sigma^2 K} \boldsymbol{h}_i^H \boldsymbol{h}_i\right)^{-1} \in C_{1\times 1} \tag{22}$$

Thus $\boldsymbol{h}_i \boldsymbol{h}_i^H$ a complex ($\boldsymbol{N}_t \times \boldsymbol{N}_t$) matrix is changed by 1×1 using the above manipulation. This eventually reduces the computations involved in the matrix inversion from $2\boldsymbol{N}_t^2$ into single scalar operation. Thus, the final precoder vector is computed by combining the results of Eq. (18) and Eq. (22) for the power and directional precoder vector components respectively with reduced complexity.

IV. Results and Discussion

The simulation is carried out for a downlink massive MIMO urban micro cell scenario having BS with 64 and 128 uniform circular array serving K=4 users. This scenario is simulated using the WINNER phase-II model [19] in MATLAB software. The coverage area of BS is 500 meters and the height of the BS antennas antenna heights is well above the surrounding objects as the scenario is an urban environment. Random data sequences of 4000 bits are generated and modulated using 16-QAM. Bit Error Rate (BER) performance of heuristic ZF, MF, and proposed heuristic MMSE precoding algorithms are compared against the optimal MMSE precoding [20] for *K*=4 users with different transmit antenna configurations of 64 and 128 given in Fig. 1 and Fig. 2 respectively.

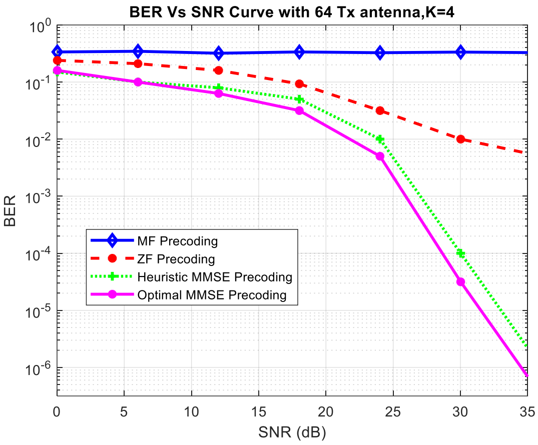

Fig. 1. BER Performance in Urban Microcell scenario for 64 × 4 massive MIMO system.

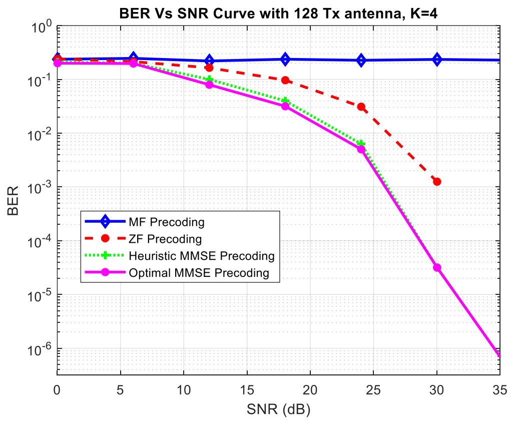

Fig. 2. BER Performance in Urban Microcell scenario for 128 × 4 massive MIMO System.

From the results, it is clear that MF is comparable with ZF and Heuristic MMSE only for low Signal to Noise Ratio (SNR). As SNR improves the performance of ZF improves whereas the proposed heuristic MMSE outperforms the other two overall ranges of SNR. Moreover, the heuristic algorithm performance is improved as the number of antennas increases. This is because as the number of transmitting antennas increases the orthogonality of the channel improves which creates independent virtual channels between the transmitting and receiving antennas. Considering this symmetric channel condition where all have equal gain, the virtual uplink parameter is also chosen to be equal to *P/K* which is in the case of heuristic beamforming. This makes the sense that transmits heuristic MMSE precoding reaches optimal beamforming with asymmetric as the number of antennas is more. Hence it is concluded that the low complexity near-optimal heuristic beamforming algorithms produce optimal results in massive MIMO systems. From the above results, it is evident that the heuristic algorithm tries to attain the optimal beamforming algorithm in our massive MIMO scenario and in [20], but there exists a small gap between the two because the optimal MMSE precoding algorithm is an iterative algorithm that computes the precoder vector without making any assumptions. This optimal algorithm is suitable for all types of channels, like symmetric or asymmetric, and works well for various ranges of SNR. Moreover, the results in [20] were simulated for the number of transmitting antennas 4 and 8, whereas our simulation environment was created using the Winner II model with 64 and 128 transmitting antennas.

Fig 3. shows the BER performance of 4 users and 8 users. This illustrates that as the number of users increases there is a mild shift in characteristics evenly over the entire range of SNR.

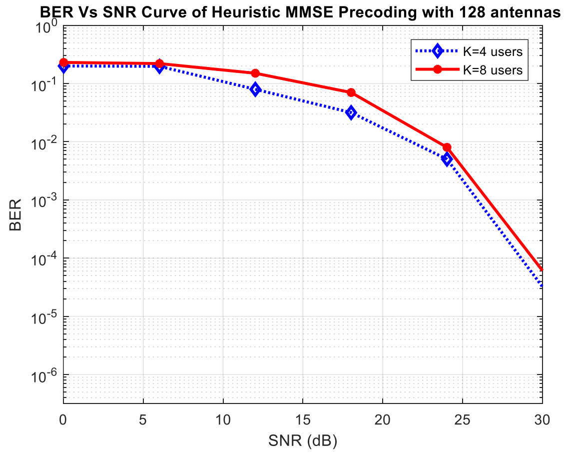

Fig. 3. BER Performance proposed Heuristic precoding with 128 antennas for *K*=4 and 8 users.

V. Conclusion

The computation of optimal multiuser beamforming vector is very difficult as the problems are generally non-convex, and hence, to simplify the approach, heuristic beamforming schemes have been proposed with reduced complexity. In this work, we have shown the BER performance of the proposed heuristic MMSE precoding for a massive MIMO base station having 64 and 128 transmit antennas with 4 mobile users under an urban microcell scenario of WINNER II model. The proposed heuristic beamforming approach sets a single constant value of *P*/*K* for both virtual uplink and downlink power parameters rather than computing *K* different values. This approach performs well in a massive MIMO scenario where the number of antennas is higher. The simulation results prove that the above approach nears optimal MMSE precoding under symmetric channel conditions where the number of transmitting antennas is large in comparison with the number of receiving antennas, which makes the channels completely orthogonal and independent of each other. The simulation results performed under urban microcell scenario prove that the above discussed method of low complexity precoding algorithm works well under symmetric channel conditions.

The above results provide a heuristic approach for the computation of the precoder vector, which seems to be a very complex task under asymmetric channel conditions. In general, as the number of transmitting antennas increases in massive MIMO systems, the orthogonality between the antennas increases, which in turn makes the virtual channels between the transmitting and receiving antennas completely independent. In the future, considering the heuristic approach for massive MIMO systems, we can design precoders using machine learning algorithms by predicting precoder weights using the model trained using the optimal MMSE algorithms. This makes the system more robust to variations in channel conditions.

Conflict of Interest

The author declares no conflict of interest.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun*., vol. 9, no. 11, pp. 3590–3600, 2010.

[2] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[3] C. S. Park, Y. S. Byun, A. M. Bokiye, and Y. H. Lee, "Complexity reduced zero-forcing beamforming in massive MIMO systems," *IEEE Information Theory and Applications Workshop*, pp. 1–5, 2014.

[4] X. Zhu *et al*., "Performance analysis of linear precoding in massive MIMO systems with finite-alphabet inputs," *IEEE Access*, vol. 7, 2019.

[5] X. Mo, L. Gui, K. Ying, X. Sang, and X. Diao, "Low complexity joint hybrid precoding for RIS-assisted wideband wireless systems," *Digital Signal Processing*, vol. 140, 2023, 104138.

[6] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," in *Proc. 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 800–805.

[7] M. Zhang, J. Gao, and C. Zhong, "A deep learning-based framework for low complexity multiuser MIMO precoding design," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, 2022, pp. 11193–11206.

[8] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 142–148, 2014.

[9] W. Xia *et al*., "A deep learning framework for optimization of MISO downlink beamforming," *IEEE Transactions on Communications*, vol. 68, no. 3, 2020.

[10] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process*, vol. 59, no. 9, pp. 4331–4340, 2011.

[11] S. S. Christensen, R. Agarwal, E. D. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMOBC beamforming design," *IEEE Trans. Wireless Commun*., vol. 7, no. 12, pp. 4792–4799, 2008.

[12] J. Shi, W. Wang, X. Yi, X. Gao, and G. Y. Li, "Deep learning based robust precoding for massive MIMO," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7429–7443, 2021.

[13] W. Yu, "Sum-capacity computation for the Gaussian vector broadcast channel via dual decomposition," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 754–759, 2006.

[14] A. Koc, M. Wang, and T. L. Ngoc, "Deep learning based multi-user power allocation and hybrid precoding in massive MIMO systems," in *Proc. ICC 2022 IEEE International Conference on Communications*, 2022, pp. 5487–5492.

[15] H. Sun *et al*., "Learning to optimize: Training deep neural networks for wireless resource management," in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun*., 2017, pp. 1–6.

[16] Q. Hu *et al*., "Iterative algorithm induced deep-unfolding neural networks: Precoding design for multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 20, 2020, pp. 1394–1410.

[17] E. Bjornson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Processing Magazine*, vol. 31, no. 4, 2014.

[18] E. Bobrov, B. Chinyaev, V. Kuznetsov, *et al*., "Power allocation algorithms for massive MIMO systems with multi-antenna users," *Wireless Networks*, vol. 29, 2022.

[19] L. Hentila, P. Kyosti, M. Kaske, M. Narandzic, and M. Alatossava. (2007). Matlab implementation of the WINNER phase II channel model ver 1.1. [Online]. Available: https://www.winner-project.eu/.

[20] E. Bjornson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2, pp. 113–381, 2013.