Online Power Allocation for Sum Rate Maximization in TDD Massive MIMO Systems

Mingyu Li*, Tianyu Wang*†, and Shaowei Wang*
*School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China
†National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
Email: 151180066@smail.nju.edu.cn, {tianyu.alex.wang, wangsw}@nju.edu.cn

***Abstract*—In this paper, we investigate the power allocation problem with perfect channel state information (CSI) for downlink sum rate maximization in time duplex division massive MIMO systems. We note that the downlink sum rate is generally a non-convex function of the power allocation vector and the corresponding solution space increases exponentially with the number of simultaneous users, which make the optimal power allocation computationally intractable in general. Here, we introduce an online paradigm to achieve a tradeoff between computational complexity and sum rate performance, which utilizes the state-of-art online learning techniques to iteratively update the power allocation according to continuous CSIs. The computational complexity of the proposed algorithm is highly reduced by using the first order optimization technique and the system sum rate is improved by exploiting the time correlation of wireless channels. Simulation results show that the downlink sum rate can be increased by 15%-100% by using our proposed algorithm, compared with the conventional average and pathloss-based power allocation schemes.**

***Index Terms*—Massive MIMO, online learning, power allocation.**

I. Introduction

Massive MIMO has been proved to be an effective technology to increase the spectrum efficiency of cellular networks. There are two different duplex modes in massive MIMO, i.e., frequency division duplex (FDD) and time division duplex (TDD), respectively. In FDD, uplink and downlink transmissions are simultaneously performed in different frequency bands. In TDD, uplink and downlink transmissions are performed at different time slots in a common band. Compared with FDD systems, TDD systems utilize the reciprocity of wireless channels to obtain downlink channel state information (CSI) from the uplink pilots, which avoids the radio resource consumption of downlink CSI feedback [1]. Here, we focus on the power allocation problem in TDD massive MIMO systems.

Power allocation in TDD massive MIMO has been widely studied in the literature [2]–[5]. In [2], the authors investigate power allocation schemes for beam division multiple access transmission in multi-cell massive MIMO systems. They show that non-overlapping of different user beams is a necessary condition for optimal power allocation. In [3], the authors also consider a multi-cell power allocation problem. To avoid the high complexity of optimizing the entire network, the proposed algorithm only adjusts the transmission power of base stations that do not interfere with each other. In [4], the authors propose a max-min fairness power allocation method. The core idea is that users with bad channel conditions should have the same spectrum efficiency as other users. However, this fairness-based method highly reduces the downlink sum rate. In [5], the authors propose a pathloss-based power allocation method in which the power allocated to a mobile user is proportional to the received power of its uplink pilot.

In the existing literature, power allocation methods based on optimization techniques can highly improve the system performance [2], [3]. However, these methods always come up with high computational complexity. For low-complexity methods, such as average and pathloss-based power allocation schemes, the system performance is always sacrificed [4], [5]. In order to achieve a better tradeoff between computational complexity and system performance, we consider an online optimization scheme, in which the power allocation problem is solved online by learning the channel dynamics from continuous CSIs in each transmission time interval (TTI). The system performance can be improved by exploiting the time correlation of wireless channels, while at the same time, low computational complexity in each TTI can be maintained by using first order optimization techniques. We note that online learning techniques have been introduced to solve the optimal sub-carrier power allocation problem for energy efficiency maximization in MIMO systems [6]. However, our considered power allocation problem is for downlink sum rate maximization in massive MIMO systems, which makes it different from the existing work.

In this paper, we investigate the power allocation problem with perfect CSI for downlink sum rate maximization in TDD massive MIMO systems. We show that the downlink sum rate is a generally non-convex function of the power allocation vector. In order to achieve a better tradeoff between computational complexity and downlink sum rate, a power allocation algorithm based on online gradient descent is proposed. Simulation results show that the downlink sum rate can be increased by 15%-100% by using our proposed algorithm, compared with the conventional average and pathloss-based power allocation schemes.

The rest of this paper is organized as follows: In Section II,

This work was partially supported by the National Natural Science Foundation of China (61671233, 61801208), the Jiangsu Science Foundation (BK20170650), the Postdoctoral Science Foundation of China (BX201700118, 2017M621712), the Jiangsu Postdoctoral Science Foundation(1701118B), and the open research fund of National Mobile Communications Research Laboratory (2019D02).

978-1-7281-0962-6/19/$31.00 ©2019 IEEE

we formulate the considered power allocation problem as an optimization problem and show its non-convexity. In Section III, we propose the online power allocation algorithm. In Section IV, we present the simulation results of the proposed algorithm and compare it with the conventional average and pathloss-based power allocation schemes. Finally, we conclude our paper in Section V.

II. System Model

In this section, we formulate the power allocation problem of a downlink TDD massive MIMO system as an optimization problem, and we show that it is generally a non-convex optimization problem. We use $(.)^T$ and $(.)^H$ to indicate the transpose and the conjugate transpose of a matrix, respectively.

A. Power Allocation

We consider a single-cell downlink TDD massive MIMO system with one base station equipped with M antennas serving K single-antenna mobile users. We denote $\mathbf{H} \in \mathbb{C}^{M\times K}$ as the channel matrix between the base station and K users, and $\mathbf{W} \in \mathbb{C}^{M\times K}$ as the linear precoding matrix of the base station. The power allocation coefficient for user k is denoted by p_k, and $\mathbf{P} \triangleq diag\,(p_1, p_2, ..., p_k)$ is the power allocation coefficient matrix. Given $\mathbf{q} \triangleq (q_1, q_2, ..., q_K)^T$ as the original symbol to be transmitted for K users, which satisfies $\mathbf{E}\left\{\mathbf{q}\mathbf{q}^H\right\} = \mathbf{I}_K$, and P_0 as the total downlink transmission power. The linearly precoded signal vector transmitted by the base station is given by

$$\mathbf{x} = \mathbf{W}\sqrt{P_0\mathbf{P}}\mathbf{q}. \tag{1}$$

The power allocation coefficient matrix $\mathbf{P}$ must satisfy the power constraint $\mathbf{E}\left\{||\mathbf{x}||^2\right\} \le P_0$. Thus, by substituting (1) into the power constraint function, we have

$$\sum_{i=1}^{K} p_i \sum_{j=1}^{K}(W_{ij}W_{ij}^*) \le 1. \tag{2}$$

We denote $\mathbf{h}_i$ and $\mathbf{w}_i$ as the i^{th} columns of $\mathbf{H}$ and $\mathbf{W}$, respectively. The received signal at user k is then given by

$$y_k = \mathbf{h}_k^T\mathbf{w}_k\sqrt{P_0 p_k}q_k + \sum_{i=1,i\neq k}^{K}\mathbf{h}_i^T\mathbf{w}_i\sqrt{P_0 p_i}q_i + z_k, \tag{3}$$

where z_k represents the additive white Gaussian noise with zero mean and variance σ^2 at user k. Note that $\mathbf{h}_k^T\mathbf{w}_k\sqrt{P_0 p_k}q_k$, $\sum_{i=1,i\neq k}^{K}\mathbf{h}_i^T\mathbf{w}_i\sqrt{P_0 p_i}q_i$ and z_k represent the signal part, the interference part and the noise part, respectively. We have the downlink sum rate is given by

$$R = \sum_{k=1}^{K}\log_2(1 + \frac{\frac{P_0}{\sigma^2}p_k|\mathbf{h}_k^T\mathbf{w}_k|^2}{\frac{P_0}{\sigma^2}\sum_{i=1,i\neq k}^{K}p_i|\mathbf{h}_k^T\mathbf{w}_i|^2 + 1}). \tag{4}$$

Therefore, the optimal power allocation problem for downlink sum rate maximization is formulated as follows:

$$\begin{aligned}\max_{p_1,p_2,...,p_K} \quad & \sum_{k=1}^{K}\log_2(1 + \frac{\frac{P_0}{\sigma^2}p_k|\mathbf{h}_k^T\mathbf{w}_k|^2}{\frac{P_0}{\sigma^2}\sum_{i=1,i\neq k}^{K}p_i|\mathbf{h}_k^T\mathbf{w}_i|^2 + 1}),\\ s.t. \quad & \sum_{i=1}^{K}p_i\sum_{j=1}^{K}(W_{ij}W_{ij}^*) \le 1, \quad p_i \ge 0, \quad i = 1, 2, ..., K.\end{aligned} \tag{5}$$

B. Non-convexity

To simplify our notions, we define

$$\widetilde{p}_i \triangleq p_i\sum_{j=1}^{K}(W_{ij}W_{ij}^*), \tag{6}$$

and

$$a_{ki} \triangleq \frac{P_0}{\sigma^2}|\mathbf{h}_k^T\mathbf{w}_i|^2 / \sum_{j=1}^{K}(W_{ij}W_{ij}^*). \tag{7}$$

Thus, by substituting equations (6) and (7) into problem (5), we can rewrite our considered power allocation problem as follows:

$$\begin{aligned}\max_{\widetilde{p}_1,\widetilde{p}_2,...,\widetilde{p}_K} \quad & \sum_{k=1}^{K}\log_2(1 + \frac{\widetilde{p}_k a_{kk}}{\sum_{i=1,i\neq k}^{K}\widetilde{p}_i a_{ki} + 1}),\\ s.t. \quad & \sum_{i=1}^{K}\widetilde{p}_i \le 1, \quad \widetilde{p}_i \ge 0, \quad i = 1, 2, ..., K.\end{aligned} \tag{8}$$

In order to determine the convexity of problem (8), we need to know whether the objective function is convex to the power allocation vector. We denote $\mathbf{G}$ as the Hessian matrix of R, defined as

$$\begin{bmatrix} \frac{\partial^2 R}{\partial \widetilde{p}_1^2} & ... & \frac{\partial^2 R}{\partial \widetilde{p}_1 \partial \widetilde{p}_K} \\ ... & ... & ... \\ \frac{\partial^2 R}{\partial \widetilde{p}_K \partial \widetilde{p}_1} & ... & \frac{\partial^2 R}{\partial \widetilde{p}_K^2} \end{bmatrix},$$

where the k-th element on the principal diagonal is given by

$$\begin{aligned}\frac{\partial^2 R}{\partial \widetilde{p}_k^2} =& \frac{-a_{kk}^2}{(\sum_{i=1}^{K}\widetilde{p}_i a_{ki} + 1)^2} + \\ & \sum_{i=1,i\neq k}^{K}\frac{\widetilde{p}_i a_{ii} a_{ik}^2(\sum_{j=1}^{K}\widetilde{p}_j a_{ij} + \sum_{j=1,j\neq i}^{K}\widetilde{p}_j a_{ij} + 2)}{(\sum_{j=1}^{K}\widetilde{p}_j a_{ij} + 1)^2(\sum_{j=1,j\neq i}^{K}\widetilde{p}_j a_{ij} + 1)^2}.\end{aligned} \tag{9}$$

For any semi-definite matrix, all elements on its principal diagonal should be non-negative. Therefore, the objective function of problem (8) is convex only if all elements on the principal diagonal of $\mathbf{G}$ is non-negative, i.e., $\partial^2 R/\partial \widetilde{p}_k^2 >= 0, k = 1, 2, ...K$. We set $K = 2$, $\widetilde{p}_1 = 1$, $\widetilde{p}_2 = 0$ and $\mathbf{m} \triangleq (1, 0)$, and we have

$$\begin{aligned}\mathbf{mGm^T} =& (1,0)\begin{bmatrix} \frac{\partial^2 R}{\partial \widetilde{p}_1^2} & \frac{\partial^2 R}{\partial \widetilde{p}_1 \partial \widetilde{p}_2} \\ \frac{\partial^2 R}{\partial \widetilde{p}_2 \partial \widetilde{p}_1} & \frac{\partial^2 R}{\partial \widetilde{p}_2^2} \end{bmatrix}(1,0)^T \\ =& -\frac{a_{11}^2}{(a_{11}+1)^2} \\ <& 0.\end{aligned} \tag{10}$$

Therefore, the Hessian matrix $\mathbf{G}$ is not semi-definite in general, and the objective function R is a generally non-convex function of the power allocation vector. Since the objective function is non-convex, it is computationally intractable to get a global optimal solution that maximizes the sum rate.

III. Online Power Allocation

In this section, we first provide some preliminaries on online learning, and then propose an online power allocation method based on the online gradient descent algorithm.

A. Preliminaries on Online Learning

Since the actual environment can be extremely complicated, it is sometimes infeasible to solve practical problems by establishing a comprehensive theoretical model in which classical mathematical optimization techniques can be utilized. Online learning is a process of continuous learning and decision. It can quickly adjust the model according to online feedback data, so as to reflect real-time changes that help to make accurate online decisions. Due to its real-time characteristic and low computational complexity, online learning has achieved stupendous success in various fields [7].

Specifically, in each time instance t, the learner takes an action $x_t \in X$, and receives a feedback from the environment $f_t(x_t)$, in which $f_t : X \to \mathbb{R}$ represents the loss of action x_t at time t. Then, the learner updates its strategy x_{t+1} according to the feedback. The performance of online learning algorithms is evaluated by the regret, which is defined as

$$\text{Regret} = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} f_t(x), \tag{11}$$

which represents the extra loss of the online strategy compared to the optimal static strategy in hindsight.

B. Power Allocation Based on Online Gradient Descent

For the considered power allocation problem, the power allocation $\widetilde{\mathbf{P}}(n)$ at TTI n can be viewed as the action, and the instantaneous sum rate R can be viewed as the benefit received from action $\widetilde{\mathbf{P}}(n)$, which corresponds to the loss function in online learning. Therefore, we can utilize the online framework to update the power allocation $\widetilde{\mathbf{P}}(n+1)$ in the next TTI.

We assume that the base station employs the maximum ratio transmission (MRT) precoding method. Thus, the precoding matrix is given by

$$\mathbf{W} = \mathbf{H}^*, \tag{12}$$

where $(.)^*$ indicates the conjugate of the matrix. By employing the MRT precoding, $\widetilde{p}_i$ in (6) is given by

$$\widetilde{p}_i \triangleq p_i \sum_{j=1}^{K} (H_{ij}^* H_{ij}), \tag{13}$$

and a_{ki} in (7) is given by

$$a_{ki} \triangleq \frac{P_0}{\sigma^2} |\mathbf{h}_k^T \mathbf{w}_i|^2 / \sum_{j=1}^{K} (H_{ij}^* H_{ij}). \tag{14}$$

Algorithm 1 Online Power Allocation

Parameter: variable step-size sequence $\gamma_n > 0$.
Initialize: $n \leftarrow 0$; Choose $\widetilde{\mathbf{P}}$ so that the constraint $\mathbf{E}\left\{||\mathbf{x}||^2\right\} \leq P_0$ is satisfied.

```
repeat
    n ← n + 1;
    get H;
    SINR_k ← p̃_k a_kk / Σ_{i=1,i≠k}^K p̃_i a_ki + 1;
    R ← Σ_{k=1}^K log_2(1 + SINR_k);
    Q ← diag(∂R/∂p̃_1, ∂R/∂p̃_2, ..., ∂R/∂p̃_K);
    Update and project: P̃ ← Π(P̃ + γ_n Q);
until transmission ends.
```

- **repeat**
  - $n \leftarrow n+1$;
  - get $\mathbf{H}$;
  - $SINR_k \leftarrow \widetilde{p}_k a_{kk} / \sum_{i=1, i\neq k}^{K} \widetilde{p}_i a_{ki} + 1$;
  - $R \leftarrow \sum_{k=1}^{K} \log_2(1 + SINR_k)$;
  - $\mathbf{Q} \leftarrow diag(\partial R/\partial \widetilde{p}_1, \partial R/\partial \widetilde{p}_2, ..., \partial R/\partial \widetilde{p}_K)$;
  - Update and project: $\widetilde{\mathbf{P}} \leftarrow \mathbf{\Pi}(\widetilde{\mathbf{P}} + \gamma_n \mathbf{Q})$;
- **until** transmission ends.

The core idea of our method is to track the gradient matrix of R to maximize the real-time downlink sum rate and then project back to the feasible area when the power limits are violated, as shown in Fig 1. To that end, we denote $\mathbf{Q}$ as the

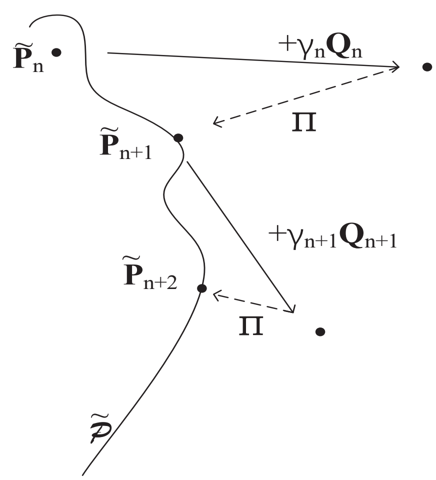


Fig. 1. Schematic diagram of recursive process in online gradient descent.

gradient matrix, given by

$$\mathbf{Q} = diag(\frac{\partial R}{\partial \widetilde{p}_1}, \frac{\partial R}{\partial \widetilde{p}_2}, ..., \frac{\partial R}{\partial \widetilde{p}_K}), \tag{15}$$

where,

$$\begin{aligned} \frac{\partial R}{\partial \widetilde{p}_k} = & \frac{a_{kk}}{\sum_{i=1}^{K} \widetilde{p}_i a_{ki} + 1} - \\ & \sum_{i=1, i\neq k}^{K} \frac{\widetilde{p}_i a_{ii} a_{ik}}{(\sum_{j=1}^{K} \widetilde{p}_j a_{ij} + 1)(\sum_{j=1, j\neq i}^{K} \widetilde{p}_j a_{ij} + 1)}. \end{aligned} \tag{16}$$

Here, we utilize the online gradient descent method [7], in which the updated strategy is given by

$$\widetilde{\mathbf{P}}(n+1) = \mathbf{\Pi}(\widetilde{\mathbf{P}}(n) + \gamma_n \mathbf{Q}(n)), \tag{17}$$

where $\gamma_n > 0$ is a variable step-size sequence, and Π represents the projection of the matrix:

$$\mathbf{\Pi}(\mathbf{Y}) = \arg\min_{\widetilde{\mathbf{P}}\in\widetilde{\mathcal{P}}} ||\widetilde{\mathbf{P}} - \mathbf{Y}||^2. \tag{18}$$

Note that equation (18) is equivalent to the following optimization problem

$$\begin{aligned} &\max_{\widetilde{p}_1,\widetilde{p}_2,...,\widetilde{p}_K} \quad \mathrm{tr}(\mathbf{Y}\widetilde{\mathbf{P}}) - \frac{1}{2}||\widetilde{\mathbf{P}}||^2, \\ &s.t. \quad \sum_{i=1}^{K} \widetilde{p}_i \le 1, \quad \widetilde{p}_i \ge 0, \quad i = 1, 2, ..., K. \end{aligned} \tag{19}$$

Since $\mathbf{Y}$ and $\widetilde{\mathbf{P}}$ are diagonal matrices, we have

$$\mathrm{tr}(\mathbf{Y}\widetilde{\mathbf{P}}) - \frac{1}{2}||\widetilde{\mathbf{P}}||^2 = \sum_j y_j\widetilde{p}_j - \frac{1}{2}\sum_j \widetilde{p}_j^2. \tag{20}$$

Therefore, problem (19) can be rewritten as follows:

$$\begin{aligned} &\max_{\widetilde{p}_1,\widetilde{p}_2,...,\widetilde{p}_K} \quad \sum_j y_j\widetilde{p}_j - \frac{1}{2}\sum_j \widetilde{p}_j^2, \\ &s.t. \quad \sum_j \widetilde{p}_j \le 1, \quad \widetilde{p}_j \ge 0, \quad j = 1, 2, ..., K. \end{aligned} \tag{21}$$

Writing the Lagrange multiplier $\lambda_j \ge 0$ of the constraint $\widetilde{p}_j \ge 0$ and the Lagrange multiplier $\lambda \ge 0$ of the constraint $\sum_j \widetilde{p}_j \le 1$, the first-order Karush-Kuhn-Tucker (KKT) conditions for (21) are given by

$$y_j = \widetilde{p}_j + \lambda - \lambda_j, \tag{22a}$$

$$\lambda_j\widetilde{p}_j = 0, \quad \lambda(1 - \sum_j \widetilde{p}_j) = 0. \tag{22b}$$

By solving (22a) and (22b), we get

$$\widetilde{p}_i = \begin{cases} 0 & if \quad y_i < 0 \\ y_i & if \quad y_i \ge 0 \quad and \quad \sum_j [y_j]_+ < 1 \\ [y_i - \lambda]_+ & if \quad y_i \ge 0 \quad and \quad \sum_j [y_j]_+ \ge 1. \end{cases} \tag{23}$$

In the above formula, $\lambda > 0$ is obtained by

$$\sum_{i:y_i\ge 0} [y_i - \lambda]_+ = 1. \tag{24}$$

We note that $\mathbf{\Pi}(\mathbf{Y})$ can be expressed as

$$\mathbf{\Pi}(\mathbf{Y}) = \widetilde{\mathbf{P}}^*, \tag{25}$$

where $\widetilde{\mathbf{P}}^*$ denotes the optimal value of $\widetilde{\mathbf{P}}$. Note that the projection matrix can be written as

$$\mathbf{\Pi}(\mathbf{Y}) = \mathbf{\Pi}(diag(\mathbf{y})). \tag{26}$$

We define $\mathbf{\Pi}(diag(\mathbf{y}))$ as follows:

$$\mathbf{\Pi}(diag(\mathbf{y})) = diag(\pi(\mathbf{y})), \tag{27}$$

then we have $\pi_i(\mathbf{y}) = \widetilde{p}_i$. The algorithm is summarized in Alg. 1. Compared to the optimization algorithm that achieves convergence through multiple iterations, our proposed algorithm only requires several iterations, which greatly reduces the computational complexity.

IV. Simulation Results

In this section, we evaluate the performance of the proposed online power allocation algorithm and compare it with the two low-complex power allocation methods. The first is the pathloss-based algorithm, in which the downlink power allocation is proportional to the inverse of the pathloss of uplink channels [5]. The second is the average power allocation algorithm, in which the total power is equally distributed to each user [8].

We utilize the COST 2100 channel model to simulate the propagation parameters with time correlation. Specifically, we consider a single cell with 10 MHz at central frequency 1 GHz. The number of base station antennas is $M = 64$ and the number of users is $K = 4$. The speed of mobile users is set at 0 m/s, 1 m/s and 10 m/s, respectively, to simulate different levels of time correlation. The TTI is set at 0.5 ms [9]. The parameters are summarized in Table I.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
TTI	0.5ms
Central frequency	1GHz
Total bandwidth	10MHz
User speed	$(0, 1, 10)$m/s
Number of transmit antennas	M=64
Number of users	K=4

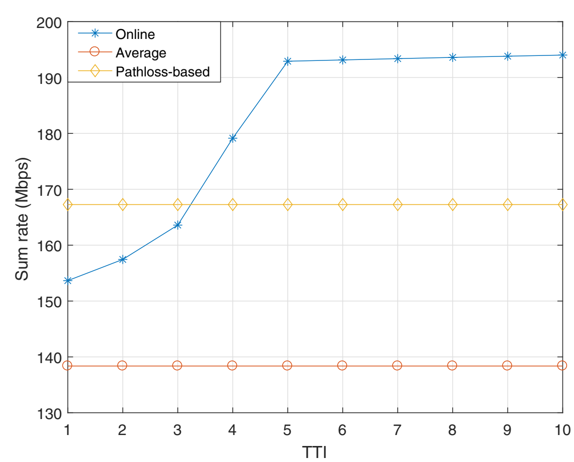
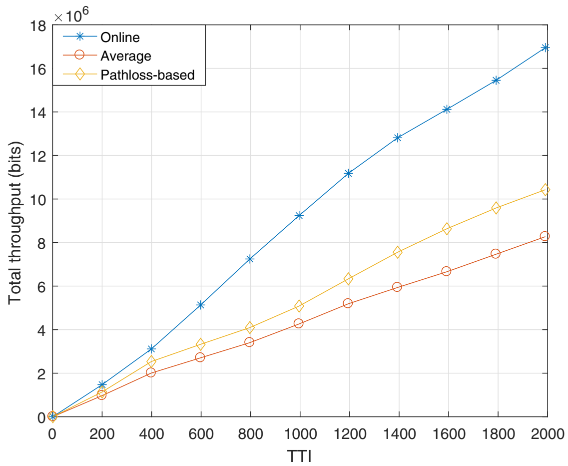
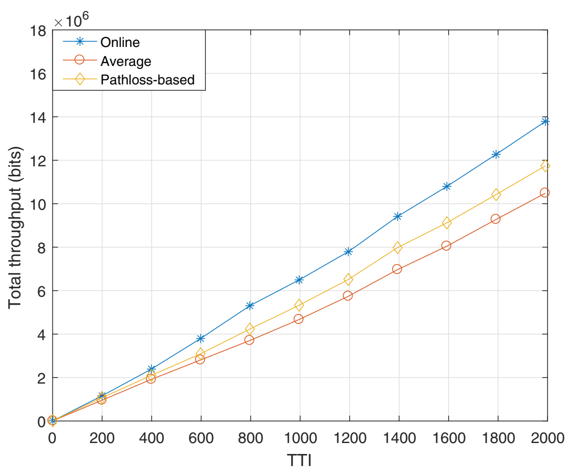


Fig. 2. Sum rate as a function of time with static users.

In Fig. 2, we show the sum rate as a function of time with static users. As we can see, the sum rates of the pathloss-based and average power allocation algorithms remain 168 Mbps and 138 Mbps, since the channel is static over time. For the proposed online power allocation algorithm, the sum rate increases from 154 Mbps to 195 Mbps after 5 TTIs of learning. The sum rate of the proposed online power allocation algorithm is 40% higher than the average power allocation scheme and is 17% higher than the pathloss-based power allocation scheme.

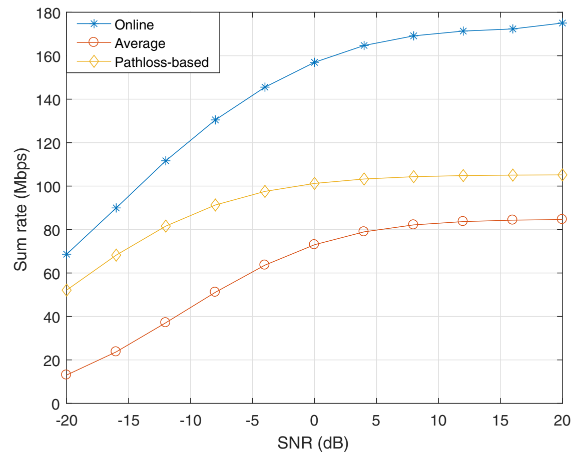


(a) Pedestrian (V = 1m/s)

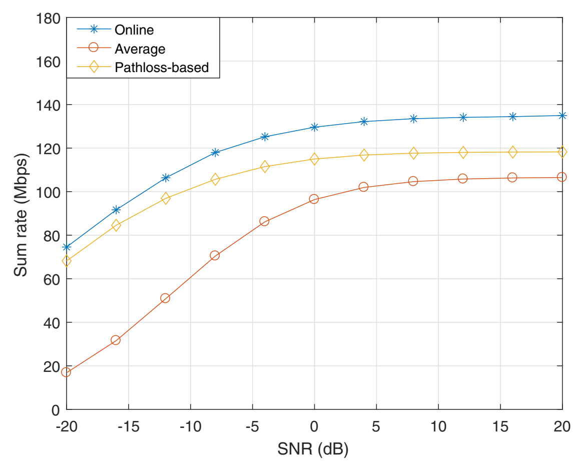


(b) Vehicular (V = 10m/s)

Fig. 3. Total throughput as a function of time with mobile users.



(a) Pedestrian (V = 1m/s)



(b) Vehicular (V = 10m/s)

Fig. 4. Sum rate as a function of SNR with mobile users.

In Fig. 3, we show the total throughput as a function of time with mobile users. Fig. 3(a) and Fig. 3(b) represent pedestrians (1 m/s) and vehicles (10 m/s), respectively. As we can see, the total throughput of the three methods increases linearly with increasing TTI. Our online power allocation algorithm can track the change of radio channels accurately when the users move slowly, therefore, the total throughput is much higher than the average and pathloss-based power allocation schemes. Specifically, after 1s (2000TTIs), the total throughput of online power allocation is 100% higher than the average power allocation and 60% higher than the pathloss-based power allocation scheme. In the case of vehicular users, the total throughput of our algorithm is only about 35% higher than the average power allocation and 15% higher than the pathloss-based power allocation scheme. This is because the channels may change before the online learning converges.

In Fig. 4, we show the sum rate as a function of SNR with mobile users. Fig. 4(a) and Fig. 4(b) represent pedestrians (1 m/s) and vehicles (10 m/s), respectively. As we can see, the sum rate increases with SNR and the proposed method outperforms the other two schemes. Specifically, when the SNR is 20dB, the sum rate of online power allocation is 100% higher than the average power allocation and 60% higher than the pathloss-based power allocation scheme when the users move slowly. In the case of vehicular users, because the channels may change before the online learning converges, the sum rate of our algorithm is only about 25% higher than the average power allocation and 13% higher than the pathloss-based power allocation scheme.

V. Conclusion

In this paper, we investigated the power allocation problem with perfect CSI for downlink sum rate maximization in TDD massive MIMO systems, which is shown to be a non-convex optimization problem in general. Given that, we propose an online power allocation algorithm based on online learning techniques. The proposed algorithm can achieve an efficient tradeoff between computational complexity and downlink sum rate. By using COST 2100 outdoor channel model, we show that the proposed online power allocation algorithm can increase the downlink sum rate by 15%-100%, compared with the conventional average and pathloss-based power allocation schemes.

REFERENCES

[1] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Tech. Journal*, vol. 20, pp. 11–22, 2015.

[2] C. Sun, X. Gao, and Z. Ding, "BDMA in multicell massive MIMO communications: Power allocation algorithms," *IEEE Trans. Signal Process.*, vol. 65, pp. 2962–2974, Jun. 2017.

[3] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez, and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 5941–5955, Nov. 2015.

[4] E. Björnson, J. Hoydis, L. Sanguinetti, *et al.*, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends ® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[5] S. Zarei, W. Gerstacker, R. Muller, and R. Schober, "Low-complexity linear precoding for downlink large-scale MIMO systems," in *Proc. IEEE PIMRC'13*, pp. 1119–1124, Sept. 2013.

[6] P. Mertikopoulos and E. V. Belmega, "Learning to be green: Robust energy efficiency maximization in dynamic mimoofdm systems," *IEEE J. Sel. Areas Commun.*, vol. 34, pp. 743–757, Apr. 2016.

[7] E. Hazan *et al.*, "Introduction to online convex optimization," *Foundations and Trends ® in Optimization*, vol. 2, no. 3-4, pp. 157–325, 2016.

[8] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. IEEE Allerton'13*, pp. 293–298, Oct. 2013.

[9] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, pp. 92–99, Dec. 2012.