An improved dropping algorithm for line-of-sight massive MIMO with Tomlinson–Harashima precoding

DOI:
10.1109/LCOMM.2019.2934680

Document status and date:
Published: 01/11/2019

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2019.2934680

An Improved Dropping Algorithm for Line-of-Sight Massive MIMO with Tomlinson-Harashima Precoding

A. Farsaei, A. Alvarado, *Senior Member, IEEE*, F. M. J. Willems, *Fellow, IEEE*, and U. Gustavsson

***Abstract*—One of the problems in line-of-sight massive MIMO is that a few users can have correlated channel vectors. To alleviate this problem, a dropping algorithm has been proposed in the literature, which drops some of the correlated users to make the spatial correlation among the remaining users be less than a certain threshold. Thresholds were found by running a large number of simulations. In this paper, the same dropping algorithm is analyzed for a known *nonlinear* precoder: Tomlinson-Harashima precoder. Instead of simulation-based thresholds, closed-form analytical expressions are derived in this paper for two power allocation schemes: max-min and equal received power control schemes. It is shown that the derived thresholds are optimal in terms of achievable sum-rate when there is only one correlated pair of users. For channels with multiple pairs of correlated users, simulation results show that using the derived thresholds improves the 5th percentile sum-rate. Due to the fairness criterion of max-min, the improvement for max-min power control is much higher than equal received power control.**

***Index Terms*—Correlated users, line-of-sight, massive MIMO, max-min power control, Tomlinson-Harashima precoding.**

I. Introduction

One of the key properties of radio channels exploited in massive MIMO systems is favorable propagation (FP) [1], [2]. FP is defined as the mutual orthogonality among the channel vectors from the base station (BS) to the users [2]. FP can be observed both in line-of-sight (LOS) and independent and identically distributed Rayleigh fading environments [3]. In LOS environments, there is a nonnegligible probability that the channel vectors of a few users become highly correlated [2], which leads to a reduction in the downlink capacity and the achievable sum-rates of linear precoders [4], [5].

To improve the achievable sum-rates of linear precoders such as conjugate beamforming (CB) and zero-forcing (ZF) with max-min power control, a correlation-based dropping (CD) algorithm was proposed in [3] for LOS environments. In the CD algorithm, the BS drops some users and reschedules them in another coherence interval to make the spatial correlation between the channel vectors of the remaining users be less than a certain threshold. The threshold is found by repeating extensive numerical simulations for the same scenario. A similar criterion was used in [6], [7]. Instead of simulation-based thresholds as in [3], thresholds are derived in [8] for CB and ZF with max-min power control.

To improve the achievable sum-rate of a BS, which uses the CD algorithm, nonlinear precoders can be used instead of linear precoders. Tomlinson-Harashima precoding (THP) is a known nonlinear precoder [9] for which an efficient implementation with a computational complexity order equal to that of ZF exists [10, Sec. 4.3.1]. In this paper, we analyze the combination of CD algorithm and THP from a theoretical viewpoint for two different power allocation schemes: max-min and equal received power control schemes. To the best of our knowledge, this is the first time the CD algorithm is analytically characterized in combination with THP.

The contribution of this paper is twofold. First, we derive closed-form analytical thresholds for the CD algorithm with THP and the two aforementioned power allocation schemes. Furthermore, we prove that using the derived thresholds when there is only one correlated pair of users, results in the optimal dropping strategy. The second contribution of this paper is to use the derived thresholds for channels with multiple pairs of correlated users. Numerical simulations show that near-optimum performance is achieved if a 100-antenna BS serving 10 single-antenna users employs the derived thresholds. By employing the CD algorithm for a BS with THP, the achievable sum-rates of both power control schemes are improved.[1]

II. System Model and Preliminaries

In this section, THP with two power allocation schemes is reviewed. The THP sum-rates of the two power allocation schemes are derived for a "correlated channel of size K" defined in [8, Definition 1]. In this type of channel, user $K-1$ and user K are the only correlated pair of users. Perfect channel state information (CSI) is assumed at the BS. We then compare the aforementioned sum-rates with the sum-rate achieved when the BS simply drops one of the correlated users.

A. Tomlinson-Harashima Precoding

The model for the downlink channel from an M-antenna BS, which uses THP to communicate with K single-antenna

A. Farsaei, A. Alvarado and F. M. J. Willems are with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (email: {a.farsaee,a.alvarado,f.m.j.willems}@tue.nl).

U. Gustavsson is with Ericsson Research, Ericsson AB, 417 56 Gothenburg, Sweden (email: ulf.gustavsson@ericsson.com).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 721732.

[1] Lowercase, bold lowercase, and bold uppercase letters denote scalars, column vectors, and matrices, respectively. $|\cdot|$ and $\|\cdot\|$ denote the absolute value and l^2-norm operators. The superscripts *, ^{T}, and ^{H} denote complex conjugate, un-conjugated transpose, and conjugated transpose, respectively. $\text{diag}(\boldsymbol{p})$ denotes a diagonal matrix with diagonal entries taken from $\boldsymbol{p}$.

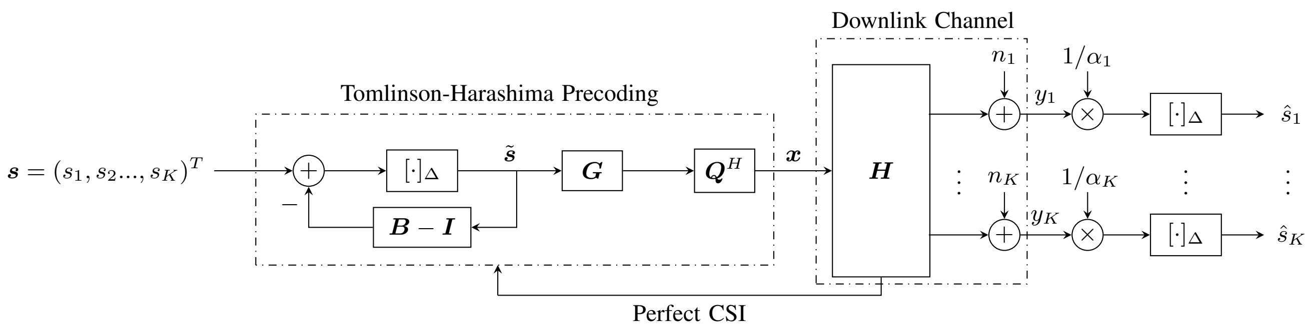

Fig. 1. The model of the downlink channel with Tomlinson-Harashima precoding.

users, is shown in Fig. 1. The intended zero-mean, uncorrelated and unit variance symbols $\boldsymbol{s} = (s_1, ..., s_K)^T \in \mathbb{C}^{K\times 1}$ are encoded to $\tilde{\boldsymbol{s}} = (\tilde{s}_1, ..., \tilde{s}_K)^T \in \mathbb{C}^{K\times 1}$ by using the feedback filter $\boldsymbol{B} - \boldsymbol{I} \in \mathbb{C}^{K\times K}$ and the modulo operator $[\cdot]_\Delta$ with divisor Δ. Then, by using the power control matrix $\boldsymbol{G} = \text{diag}(\sqrt{d_1}, ..., \sqrt{d_K})$ and feedforward filter $\boldsymbol{Q}^H \in \mathbb{C}^{M\times K}$, the precoded vector $\tilde{\boldsymbol{x}} \in \mathbb{C}^{M\times 1}$ is generated (for more details on THP see [10, Sec. 5.4.4]). The average power constraint at the BS is $\mathbb{E}[\|\boldsymbol{x}\|^2] = P_{\text{tot}}$. The transmitted vector $\boldsymbol{x}$ at the BS goes through the channel $\boldsymbol{H} = (\boldsymbol{h}_1, ..., \boldsymbol{h}_K)^T \in \mathbb{C}^{K\times M}$, where $\boldsymbol{h}_i$ is the channel vector from the BS to user i.

For THP, the LQ decomposition of the channel $\boldsymbol{H} = \boldsymbol{LQ}$ is used, where $\boldsymbol{Q}^H$ is used for the feedforward filter, and $\boldsymbol{L}$ is a lower triangular matrix with positive diagonal elements l_{ii}. The matrix $\boldsymbol{L}$ is used to find $\boldsymbol{B} = \boldsymbol{LG}$ for the feedback filter. The received signal for user i can be shown to be:

$$y_i = l_{ii}\sqrt{d_i}\tilde{s}_i + \sum_{j=1}^{i-1} l_{ij}\sqrt{d_j}\tilde{s}_j + n_i, \tag{1}$$

where n_i is complex AWGN noise with variance N_0. By using the scalar $\alpha_i = l_{ii}\sqrt{d_i}$, and using the modulo operator at the receiver, the estimated symbol for user i is:

$$\hat{s}_i = \left[\frac{y_i}{\alpha_i}\right]_\Delta = \left[\tilde{s}_i + \sum_{j=1}^{i-1} \frac{l_{ij}\sqrt{d_j}}{\alpha_i}\tilde{s}_j + \frac{n_i}{\alpha_i}\right]_\Delta. \tag{2}$$

By encoding $\tilde{s}_i = [s_i - \sum_{j=1}^{i-1} \frac{l_{ij}\sqrt{d_j}}{\alpha_i}\tilde{s}_j]_\Delta$ (see the feedback loop in Fig. 1), the estimated symbol becomes:

$$\hat{s}_i = \left[s_i + \frac{n_i}{\alpha_i}\right]_\Delta. \tag{3}$$

In high signal to noise ratios (SNRs), the modulo loss [11] can be ignored. Therefore, the following sum-rate (in bits/channel use) is achieved using inflated lattice strategies [12], where the shaping loss is neglected:

$$R_{\text{THP}} = \sum_{i=1}^{K} \log_2\left(1 + \frac{d_i}{N_0} l_{ii}^2\right). \tag{4}$$

Max-min power control coefficients are found by maximizing the minimum SNR among the users as:

$$d_{i,\text{max-min}} = \frac{P_{\text{tot}}}{l_{ii}^2 \sum_{j=1}^{K} \frac{1}{l_{jj}^2}}, \quad i = 1, ..., K. \tag{5}$$

The equal received power control coefficients considering the difference in the channel norms of the users are found by:

$$d_{i,\text{equ.}} = \frac{P_{\text{tot}}}{\|\boldsymbol{h}_i\|^2 \sum_{j=1}^{K} \frac{1}{\|\boldsymbol{h}_j\|^2}}. \tag{6}$$

In the case of mutual orthogonality among the K users, $d_{i,\text{max-min.}}$ and $d_{i,\text{equ.}}$ are equivalent since $l_{ii}^2 = \|\boldsymbol{h}_i\|^2$ (see Appendix A for more details). In this case, each user achieves the same SNR denoted by γ:

$$\gamma = \frac{P_{\text{tot}}}{N_0 \sum_{i=1}^{K} \frac{1}{\|\boldsymbol{h}_i\|^2}}. \tag{7}$$

For THP, the order of users for encoding $\boldsymbol{s}$ to $\tilde{\boldsymbol{s}}$ has to be optimized to maximize the sum-rate given the power control schemes. In this paper, the algorithm described in [10, Sec. 5.4.5] is used for THP with max-min power control to find an appropriate order of users. For equal received power control, we use the algorithm in [10, Sec. 5.4.8].

B. Sum-Rate Comparison

Consider a correlated channel of size K with the spatial correlation given by:

$$\rho = \frac{\boldsymbol{h}_K^H \boldsymbol{h}_{K-1}}{\|\boldsymbol{h}_K\|\|\boldsymbol{h}_{K-1}\|}, \tag{8}$$

where without loss of generality we assume that $\|h_{K-1}\| \geq \|h_K\|$. If the BS drops user (DU) K,[2] user 1 to user $K-1$ will experience FP. Thus, due to mutual orthogonality of the users, the following sum-rate R_{DU} is achieved (with either max-min or equal received power control) using (7) for $K-1$ users:

$$R_{\text{DU}} = (K-1)\log_2\left(1 + \frac{P_{\text{tot}}}{N_0 \sum_{i=1}^{K-1} \frac{1}{\|\boldsymbol{h}_i\|^2}}\right). \tag{9}$$

If the BS does not drop user K, the following sum-rates are achieved replacing (5) and (6) in (4):

$$R_{\text{max-min}} = K\log_2\left(1 + \frac{P_{\text{tot}}}{N_0 \sum_{i=1}^{K} \frac{1}{l_{ii}^2}}\right), \tag{10}$$

$$R_{\text{equ.}} = \sum_{i=1}^{K} \log_2\left(1 + \frac{P_{\text{tot}} l_{ii}^2}{N_0 \sum_{j=1}^{K} \frac{\|\boldsymbol{h}_i\|^2}{\|\boldsymbol{h}_j\|^2}}\right). \tag{11}$$

[2]Dropping user K, results in a higher R_{DU} when $\|\boldsymbol{h}_{K-1}\| > \|\boldsymbol{h}_K\|$. In case of $\|\boldsymbol{h}_{K-1}\| = \|\boldsymbol{h}_K\|$, user $K-1$ or user K can be dropped.

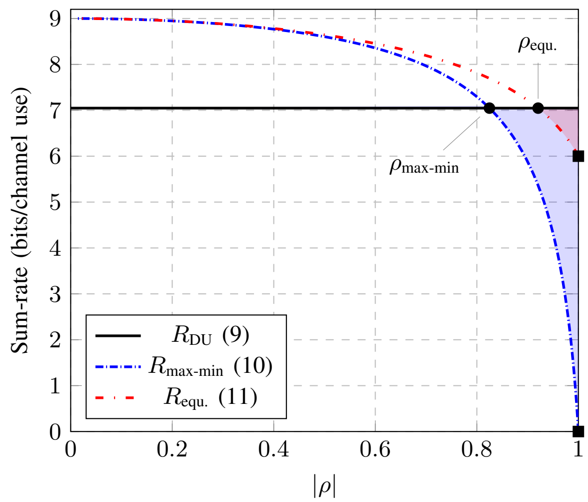

Fig. 2. R_{DU}, $R_{\text{max-min}}$, and R_{equ} as a function of $|\rho|$ for a correlated channel of size 3 when $M = 100$ and $\gamma = 7$. The blue (red) shaded area shows the extra sum-rate that the BS can achieve by dropping the user with max-min power control (equal received power control).

Max-min power control equalizes the SNR among the users by sacrificing the good users for the worst users. However, equal received power control only takes the channel norms into account, which is not as strict as max-min power control. Thus, it is expected that max-min power control achieves a lower sum-rate compared to equal received power control, i.e., $R_{\text{max-min}} \leq R_{\text{equ.}}$. In Fig. 2, R_{DU}, $R_{\text{max-min}}$, and $R_{\text{equ.}}$ are shown as a function of $|\rho|$ for a correlated channel of size $K = 3$, where the users have the same channel norms (i.e., $\|\boldsymbol{h}_1\| = \|\boldsymbol{h}_2\| = \|\boldsymbol{h}_3\|$). The horizontal black line shows R_{DU}, where the BS drops one of the correlated users regardless of $|\rho|$. For low $|\rho|$, $R_{\text{max-min}}$ (blue curve) and $R_{\text{equ.}}$ (red curve) are above R_{DU} since the users are almost orthogonal, and there is no need for dropping. By increasing $|\rho|$, both $R_{\text{max-min}}$ and $R_{\text{equ.}}$ reduce. When $|\rho| > \rho_{\text{max-min}}$ (shown by a black circle), $R_{\text{max-min}}$ falls below R_{DU}. Thus, the BS with max-min power control should drop the user only when $|\rho| > \rho_{\text{max-min}}$ to avoid the loss in the sum-rate (blue shaded area). Similarly, when $|\rho| > \rho_{\text{equ.}}$ (shown by a black circle), $R_{\text{equ.}}$ falls below R_{DU}, and the BS with equal received power control should drop the user. When $|\rho| \to 1$, $R_{\text{max-min}}$ converges to 0 bits/channel use (shown by a black square), which shows it is essential for the BS with the max-min power control to drop the correlated users. However, when $|\rho| \to 1$, $R_{\text{equ.}}$ becomes close to 6 bits/channel use (shown by a black square). By dropping the user when $|\rho| > \rho_{\text{max-min}}$ ($|\rho| > \rho_{\text{equ.}}$), the BS can achieve R_{DU}, which shows that the sum-rate with max-min power control (or equal received power) is maximized. Comparing the sum-rates reveals the goal of the CD algorithm, which is to drop the correlated user only when it leads to a higher sum-rate for the remaining users with a given power allocation scheme.

III. Analytical Thresholds for the CD Algorithm

In this section, analytical expressions for the thresholds $\rho_{\text{max-min}}$ and $\rho_{\text{equ.}}$ are derived for a correlated channel of size K, which has only one pair of correlated users. Then, simulations are presented to study the derived thresholds and how the CD algorithm performs with the derived thresholds for channels with multiple pairs of correlated users.

A. Analytical Expressions for the Thresholds

The following Theorem gives the optimal dropping strategy in terms of the THP sum-rate with max-min and equal received power control.

Theorem 1. *Consider a correlated channel of size K, with spatial correlation ρ between user $K-1$ and user K ($\|\boldsymbol{h}_{K-1}\| \geq \|\boldsymbol{h}_K\|$). The sum-rate with max-min and equal received power control is maximized if the BS drops user K, when $|\rho| > \rho_{\textit{max-min}}$ and $|\rho| > \rho_{\textit{equ.}}$, respectively, where*

$$\rho^2_{\textit{max-min}} = 1 - \frac{1}{\|\boldsymbol{h}_K\|^2 \left(\frac{P_{tot}}{N_0(2^a - 1)} - \sum_{j=1}^{K-1} \frac{1}{\|\boldsymbol{h}_j\|^2} \right)}, \quad (12)$$

with $a = R_{DU}/K$ (R_{DU} is given by (9)), and

$$\rho^2_{\textit{equ.}} = 1 - \frac{N_0 \eta (2^b - 1)}{P_{tot}}, \quad (13)$$

with $\eta = \sum_{i=1}^{K} \frac{1}{\|\boldsymbol{h}_i\|^2}$, and b:

$$b = R_{DU} - (K-1) \log_2 \left(1 + \frac{P_{tot}}{N_0 \eta} \right). \quad (14)$$

Proof. See Appendix A. □

The derivation of thresholds is true for any channel matrix with only one pair of correlated users. Our emphasis is on LOS environments in which there is a nonnegligible probability that a few users become correlated [2, Sec. 4.3].

B. Simulation Results

In this section, the derived thresholds in Theorem 1 are studied as a function of γ. Then, the use of the derived thresholds in the CD algorithm for channels with multiple pairs of correlated users is studied. The details of the simulation scenario are as follows. A BS with a uniform linear array of $M = 100$ antennas with a half-wavelength spacing is assumed, which serves $K = 10$ single-antenna users. A single-cell is assumed with a field-of-view of $[30°, 150°]$. The carrier frequency is set to 30 GHz. The users are uniformly distributed at the cell-edge (200 m) to study the worst-case performance of the system (with no shadowing). The minimum distance between two users is assumed to be a wavelength. For LOS environments, we use [13, eq. (7.26)] for the channel matrix.

Fig. 3 shows $\rho_{\text{max-min}}$ (see (12)) and $\rho_{\text{equ.}}$ (see (13)) as a function of γ for $K = 10$. As can be seen in Fig. 3, by increasing γ, both $\rho_{\text{max-min}}$ and $\rho_{\text{equ.}}$ increase. This is due to using a higher transmit power at the BS, which can overcome a higher correlation. Furthermore, $\rho_{\text{equ.}}$ is higher than $\rho_{\text{max-min}}$ for all SNRs. More importantly, by increasing γ, $\rho_{\text{equ.}}$ converges to 1, which means that dropping the correlated user may not be required for equal received power control at high SNRs.

The CD algorithm is applied on 100000 realizations of the channel to find the cumulative distribution function (CDF) of THP sum-rate. In Fig. 4, the CDF of the THP sum-rate

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2019.2934680

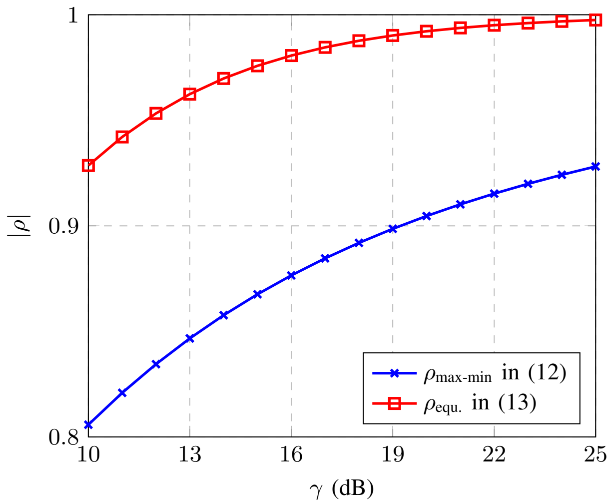

Fig. 3. The thresholds for the max-min $\rho_{\text{max-min}}$ and equal received power control $\rho_{\text{equ.}}$ as a function of γ for $K = 10$.

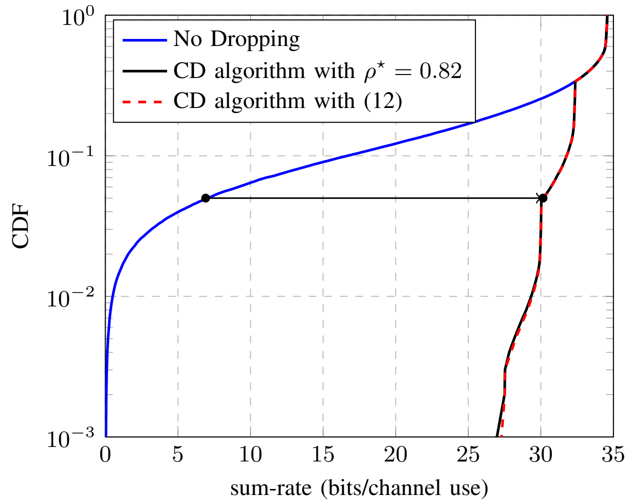

Fig. 4. CDF of THP sum-rate with max-min power control with no dropping, compared to CD algorithm with two different thresholds for $K = 10$, $M = 100$, and $\gamma = 10$. The first threshold is found by (12), whereas the second threshold $\rho^\star = 0.82$ is found by repeating the simulations. The arrow shows the improvement of the 5th percentile sum-rate by using CD algorithm.

with max-min power control with no dropping is compared to CD algorithm with two different thresholds. The first threshold denoted by $\rho^\star$ is found by repeating the simulations to maximize the 5th percentile sum-rate, whereas the second threshold is found by (12). The same curves are presented for equal received power control in Fig. 5. In Fig. 4 and Fig. 5, the 5th percentile sum-rates are shown by black circles. As can be seen in Fig. 4, dropping the correlated users is essential for the max-min power control since the 5th percentile sum-rate is improved from 6.91 to 30.16 bits/channel use. This very high improvement is due to the max-min criterion in which the users are sacrificed for the highly correlated users (see the blue shaded area in Fig. 2). On the other hand, the 5th percentile sum-rate for equal received power control is improved from 30.49 to only 31.15 bits/channel use (see Fig. 5). This improvement can be explained by the red shaded area in Fig. 2. Furthermore, as can be seen in Fig. 4 and 5, the obtained CDF using the derived threshold for the CD algorithm, almost overlaps with the CDF obtained by $\rho^\star$ for both power allocation schemes.

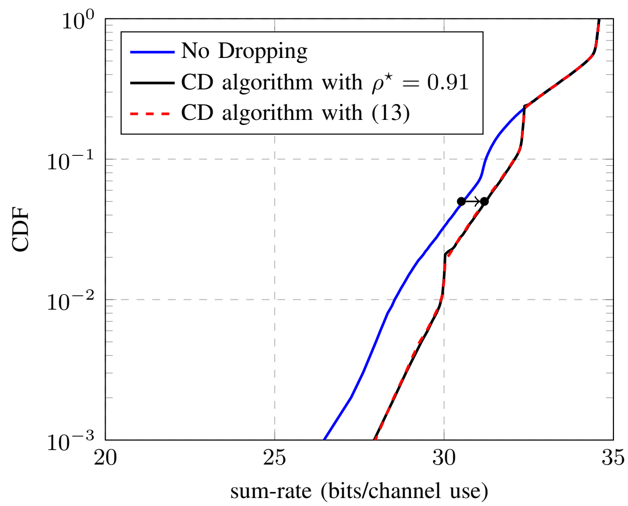

Fig. 5. Same as Fig. 4 for equal received power control with $\rho^\star = 0.91$.

IV. Conclusions

In this paper, a previously proposed dropping algorithm for LOS massive MIMO is combined with THP with max-min and equal received power control. The main contribution of our paper is to derive analytical thresholds for the CD algorithm with THP, which avoids extensive numerical simulations. The derived thresholds are optimal when there is only one correlated pair of users. For scenarios with multiple pairs of correlated users, simulation results showed that using the derived thresholds for the dropping algorithm has near-optimum performance. Furthermore, simulation results showed the CD algorithm with the derived threshold improved the 5th percentile sum-rate for both max-min and equal received power control, while the improvement for the max-min power control was much higher due to the max-min criterion. Thus, using the CD algorithm with the derived threshold is essential for max-min power control.

Appendix A
Proof of Theorem 1

To derive $l_{ii}^2, i = 1, ..., K$ for a correlated channel of size K, recall the LQ decomposition of the channel $\boldsymbol{H} = \boldsymbol{LQ}$. Then, $\boldsymbol{HH}^H = \boldsymbol{LQQ}^H\boldsymbol{L}^H = \boldsymbol{LL}^H$. For a correlated channel of size K, the diagonal elements of $\boldsymbol{HH}^H$ are $\|\boldsymbol{h}_i\|^2$ and the off-diagonal elements are 0 except:

$$\left[\boldsymbol{HH}^H\right]_{K-1,K} = \left[\boldsymbol{HH}^H\right]^H_{K,K-1} = \|\boldsymbol{h}_{K-1}\|\|\boldsymbol{h}_K\|\rho. \quad (15)$$

By comparing the first $K-2$ rows of $\boldsymbol{HH}^H$ and $\boldsymbol{LL}^H$, the corresponding elements of $\boldsymbol{L}$ are found. The diagonal elements of $\boldsymbol{L}\boldsymbol{L}^H$ are found as $l_{ii}^2 = \|\boldsymbol{h}_i\|^2, i = 1, ..., K-2$, and all the off-diagonal elements of $\boldsymbol{L}\boldsymbol{L}^H$ are 0 for $i = 1, ..., K-2$. By comparing the last two rows of $\boldsymbol{H}\boldsymbol{H}^H$ and $\boldsymbol{L}\boldsymbol{L}^H$, l_{K-1K-1}, l_{KK} and l_{KK-1} are found. This is done by introducing the following matrices:

$$\boldsymbol{A} = \begin{bmatrix} \|\boldsymbol{h}_{K-1}\|^2 & \|\boldsymbol{h}_{K-1}\|\|\boldsymbol{h}_K\|\rho \\ \|\boldsymbol{h}_{K-1}\|\|\boldsymbol{h}_K\|\rho^H & \|\boldsymbol{h}_K\|^2 \end{bmatrix}, \tag{16}$$

and

$$\boldsymbol{B} = \begin{bmatrix} l_{K-1K-1}^2 & l_{K-1K-1}l_{K-1K}^H \\ l_{K-1K-1}^H l_{K-1K} & |l_{K-1K}|^2 + l_{KK}^2 \end{bmatrix}. \tag{17}$$

By solving $\boldsymbol{B} = \boldsymbol{A}$ for l_{K-1K-1} and l_{KK}, the following solutions are found as:

$$l_{K-1K-1}^2 = \|\boldsymbol{h}_{K-1}\|^2, \;\; l_{KK}^2 = \|\boldsymbol{h}_K\|^2(1-|\rho|^2). \tag{18}$$

Thus, the diagonal elements of $\boldsymbol{L}\boldsymbol{L}^H$ are found as $l_{ii}^2 = \|\boldsymbol{h}_i\|^2, i = 1, ..., K-1$ and $l_{KK}^2 = \|\boldsymbol{h}_K\|^2(1-|\rho|^2)$. Next step is to solve $R_{\text{DU}} \geq R_{\text{max-min}}$ for $\rho_{\text{max-min}}$ using (9) and (10) for the sum-rates and using the derived l_{ii}^2. Therefore, the following inequality is solved for $\rho_{\text{max-min}}$:

$$R_{\text{DU}} > K \log_2\left(1 + \frac{P_{\text{tot}}}{N_0 \sum_{i=1}^{K} \frac{1}{l_{ii}^2}}\right). \tag{19}$$

By defining $a = R_{\text{DU}}/K$, we have:

$$(2^a - 1) > \frac{P_{\text{tot}}}{N_0 \sum_{i=1}^{K} \frac{1}{l_{ii}^2}} \Rightarrow \frac{P_{\text{tot}}/N_0}{(2^a-1)} < \sum_{i=1}^{K} \frac{1}{l_{ii}^2}. \tag{20}$$

By replacing l_{ii}^2, (20) is simplified to:

$$\begin{aligned} \frac{P_{\text{tot}}/N_0}{(2^a-1)} &< \frac{1}{\|\boldsymbol{h}_K\|^2(1-|\rho|^2)} + \sum_{i=1}^{K-1} \frac{1}{\|\boldsymbol{h}_i\|^2} \\ \frac{P_{\text{tot}}/N_0}{(2^a-1)} - \eta_1 &< \frac{\eta_2}{1-|\rho|^2}, \end{aligned} \tag{21}$$

where $\eta_1 = \sum_{i=1}^{K-1} \frac{1}{\|\boldsymbol{h}_i\|^2}$ and $\eta_2 = \frac{1}{\|\boldsymbol{h}_K\|^2}$. If we assume that $\frac{P_{\text{tot}}/N_0}{(2^a-1)} - \eta_1 > 0$, (21) is simplified to:

$$\frac{\eta_2}{\frac{P_{\text{tot}}/N_0}{(2^a-1)} - \eta_1} > 1 - |\rho|^2 \Rightarrow |\rho|^2 > 1 - \frac{\eta_2}{\frac{P_{\text{tot}}/N_0}{(2^a-1)} - \eta_1}. \tag{22}$$

To show the assumption is always true, we need to show:

$$\frac{P_{\text{tot}}/N_0}{(2^a-1)} - \eta_1 > 0 \Rightarrow \frac{P_{\text{tot}}}{\eta_1 N_0} > (2^a - 1). \tag{23}$$

We need to simplify 2^a:

$$2^a = 2^{\frac{(K-1)}{K}\log_2\left(1+\frac{P_{\text{tot}}}{N_0\eta_1}\right)} = \left(1 + \frac{P_{\text{tot}}}{N_0\eta_1}\right)^{\frac{(K-1)}{K}}. \tag{24}$$

Therefore, using (24), (23) is simplified to:

$$1 + \frac{P_{\text{tot}}}{\eta_1 N_0} > \left(1 + \frac{P_{\text{tot}}}{N_0\eta_1}\right)^{\frac{(K-1)}{K}}, \tag{25}$$

which is always true. Thus, by replacing η_1 and η_2 in (22), (12) is achieved.

For the equal received power control, $\rho_{\text{equ.}}$ is found by solving $R_{\text{DU}} > R_{\text{equ.}}$ as follows:

$$R_{\text{DU}} > \sum_{i=1}^{K} \log_2\left(1 + \frac{P_{\text{tot}} l_{ii}^2}{N_0\left(\sum_{j=1}^{K} \frac{\|\boldsymbol{h}_i\|^2}{\|\boldsymbol{h}_j\|^2}\right)}\right). \tag{26}$$

By using the derived l_{ii}, (26) is simplified to:

$$R_{\text{DU}} > \log_2\left(1 + \frac{P_{\text{tot}}(1-|\rho|^2)}{N_0\eta}\right) + \sum_{i=1}^{K-1} \log_2\left(1 + \frac{P_{\text{tot}}}{N_0\eta}\right), \tag{27}$$

where $\eta = \sum_{i=1}^{K} \frac{1}{\|\boldsymbol{h}_i\|^2}$. Furthermore, by defining:

$$b = R_{\text{DU}} - (K-1)\log_2\left(1 + \frac{P_{\text{tot}}}{N_0\eta}\right), \tag{28}$$

(27) is simplified to:

$$\begin{aligned} 2^b &> \left(1 + \frac{P_{\text{tot}}(1-|\rho|^2)}{N_0\eta}\right) \\ \frac{N_0\eta(2^b-1)}{P_{\text{tot}}} &> 1 - |\rho|^2 \Rightarrow |\rho|^2 > 1 - \frac{N_0\eta(2^b-1)}{P_{\text{tot}}}, \end{aligned} \tag{29}$$

which results in (13). □

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, Sep. 2014, pp. 76–80.
[3] H. Yang and T. L. Marzetta, "Massive MIMO with max-min power control in line-of-sight propagation environment," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4685–4693, Nov. 2017.
[4] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Linear pre-coding performance in measured very-large MIMO channels," in *2011 IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2011, pp. 1–5.
[5] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.
[6] T. Arai, A. Ohta, S. Kurosaki, K. Maruta, T. Iwakuni, and M. Iizuka, "Correlation-based user scheduling and multi-planar parallelogram array for massive antenna systems," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2016, pp. 1–6.
[7] S. Sigdel and W. A. Krzymien, "Simplified fair scheduling and antenna selection algorithms for multiuser MIMO orthogonal space-division multiplexing downlink," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1329–1344, Mar. 2009.
[8] A. Farsaei, A. Alvarado, F. M. J. Willems, and U. Gustavsson, "An improved dropping algorithm for line-of-sight massive MIMO With max-min power control," *IEEE Communications Letters*, vol. 23, no. 6, pp. 1109–1112, Jun. 2019.
[9] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-time transmission using Tomlinson-Harashima precoding," in *Proc. ITG Fachbericht*, Jan. 2002, pp. 139–148.
[10] C. Windpassinger, "Detection and precoding for multiple input multiple output channels," Ph.D. dissertation, Univ. Erlangen-Nurnberg, Erlangen, Germany, 2004.
[11] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Transactions on Communications*, vol. 53, no. 7, pp. 1220–1230, Jul. 2005.
[12] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.
[13] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, May 2005.