**QUEEN'S
UNIVERSITY
BELFAST**

Performance of a Novel Maximum-Ratio Precoder in Massive MIMO with Multiple-Antenna Users

Sutton, J. A. C., Ngo, H. Q., & Matthaiou, M. (2019). Performance of a Novel Maximum-Ratio Precoder in Massive MIMO with Multiple-Antenna Users. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)* (IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC): Proceedings). Institute of Electrical and Electronics Engineers Inc..

Published in:

IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2019 IEEE.

This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Performance of a Novel Maximum-Ratio Precoder in Massive MIMO with Multiple-Antenna Users

James A. C. Sutton, Hien Quoc Ngo, and Michail Matthaiou
Institute of Electronics, Communications and Information Technology (ECIT),
Queen's University Belfast, BT3 9DT, Belfast, U.K.
Email: {jsutton03, hien.ngo, m.matthaiou}@qub.ac.uk

Abstract—In this paper, we analyze and compare the performance of conventional maximum-ratio (MR) precoding to a proposed alternative scheme, when applied to a massive multiple-input multiple-output (MIMO) system with multiple-antenna users. In the proposed MR precoding scheme, each channel vector is divided by its norm square to increase the degree of hardening of the effective channel gains at the users. We derive closed-form expressions for the achievable spectral efficiency (SE) of both the conventional and proposed precoding schemes. These closed-form expressions are very simple and useful for further system design. For instance, for both precoding schemes, our results justify the motivation for additional user antennas since significant performance improvements are observed. The proposed scheme produces greater performance when compared to the conventional one across a range of system set-ups. More specifically, the proposed scheme is most effective in systems where the numbers of users and user antennas are kept small.

Index Terms—Channel hardening, massive MIMO, maximum-ratio processing, multiple-antenna users.

I. INTRODUCTION

Massive multiple-output multiple-input (MIMO) has become a core technology for the fifth generation (5G) of wireless communication systems due to its ability to meet the requirements set by the growing demands from the connected world. Massive MIMO systems are cellular wireless communication systems which can serve multiple users inside a cell with exceptionally good data rates. Specifically, massive MIMO operates with hundreds or thousands of base station antennas serving tens or hundreds of users inside the respective cell. Since the concept of massive MIMO was first introduced in [1] it has gained significant attention from researchers [2]–[6].

Previous research on massive MIMO almost exclusively focused on single-antenna user design, since spatial multiplexing gains at the base station could still be exploited despite the single-antenna users. Current technology no longer adheres to previous cost or size limitations of wireless devices, and as such, practical modern devices will operate with several antennas. The addition of antennas at the users would introduce multiplexing gains at the users, hence, increasing the

spectral efficiency (SE) and reliability of the system even more. Surprisingly, the potential benefits of employing multiple-antennas at the users in massive MIMO has been a relatively unexplored topic of investigation.

Papers [7] and [8] investigated massive MIMO with multiple-antennas at the users. In [7], general closed-form expressions for the SE were derived for a model incorporating multiple-antenna users. Analysis of these expressions indicated that simple linear processing would be sufficient to use with multiple-antenna user systems in massive MIMO. The analysis further showcased that additional antennas at the users would increase SE, hence, validating the relevance of developing future massive MIMO systems with this setup. One other paper [8] investigated massive MIMO with multiple-antenna users under spatially correlated Rayleigh and geometry-based channels. The work of [8] discovered that in both channels, the favorable propagation in multiple-antenna user systems decays, compared to the single-antenna user systems.

This past research on massive MIMO with multiple-antenna users considered conventional maximum-ratio (MR) linear signal processing schemes and relied on the channel hardening property at the users for data detection. As stated in [9], [10], the channel hardening property of massive MIMO occurs under Rayleigh fading channels as the base station antennas grows to infinity. Therefore, if the number of base station antennas is moderate, a conventional MR scheme will produce a system with limited hardening effectiveness at the user. In this work, we develop and analyze an alternative to conventional MR processing scheme, that can produce improved performance across a wide range of practical massive MIMO setups. Specifically, the proposed precoder is investigated with the intention of improving the channel hardening at the users of a massive MIMO model during its weakest operating conditions, without compromising performance elsewhere. Building off the work in [7], closed-form expressions for the achievable SEs of both the conventional and proposed MR precoders are derived. Note that, different with [7]

where an asymptotic downlink SE expression (when the numbers base station antennas and users grow to infinity) with MR precoder was derived, in our work, the closed-form expression for the achievable SE hold for any finite numbers of base station antennas and users. The performance differences of the two precoders are compared in order to investigate the characteristics of the proposed precoder and the benefits of using multiple antennas at the users.

Notation: Column vectors and matrices are represented by bold lower case and upper case letters, respectively. Superscript $()^H$ stands for the conjugate-transpose. The trace, Euclidean norm and expectation operators are expressed by $\text{tr}()$, $\|\cdot\|$ and $\mathbb{E}\{\cdot\}$ respectively. Finally, a circularly symmetric complex Gaussian random variable (RV) z with zero mean and variance σ^2 is denoted by $z \sim \mathcal{CN}(0, \sigma^2)$.

II. SYSTEM MODEL

Consider a single-cell massive MIMO system comprised of a base station containing M antennas that serves K users, where each user is equipped with N antennas. We assume that the systems is operating in time-division duplex (TDD) mode. Let $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ model the small-scale fading matrix between user k and the base station. The elements of $\mathbf{H}_k$ are assumed to be independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ RVs. Thus, the $M \times N$ channel response matrix between user k and the base station is modeled with

$$\mathbf{G}_k = \sqrt{\beta_k} \mathbf{H}_k, \quad (1)$$

where β_k represents the large-scale fading.

1) *Uplink Training:* During this phase, all users transmit their respective pilot sequences of length τ_p symbols to the base station. Let $\boldsymbol{\varphi}_{k,n} \in \mathbb{C}^{\tau_p \times 1}$ be the pilot sequence transmitted from the n th antenna of the k th user. Assume that all pilot sequences are pairwise orthogonal, and $\|\boldsymbol{\varphi}_{k,n}\|^2 = 1, \forall k = 1 \dots K, n = 1 \dots N$. This requires $\tau_p \geq KN$. The received pilot at the base station is

$$\mathbf{Y}_p = \sqrt{\tau_p \rho_p} \sum_{k=1}^K \sum_{n=1}^N \mathbf{g}_{k,n} \boldsymbol{\varphi}_{k,n}^H + \mathbf{N}_p, \quad (2)$$

where ρ_p is the normalised signal-to-noise ratio (SNR) of each transmitted pilot symbol, $\mathbf{g}_{k,n}$ is the n th column of the channel response matrix $\mathbf{G}_k$ (i.e. the channel vector between the all base station antennas and the n th antenna of the k th user) and $\mathbf{N}_p$ is an $M \times \tau_p$ noise matrix whose elements are i.i.d. $\mathcal{CN}(0, 1)$ RVs.

In order to estimate the channel response for each antenna of each user, the received pilot sequence $\mathbf{Y}_p$ is first projected onto $\boldsymbol{\varphi}_{k,n}$,

$$\begin{aligned} \check{\mathbf{y}}_{p,k,n} &\triangleq \mathbf{Y}_p \boldsymbol{\varphi}_{k,n} \\ &= \sqrt{\tau_p \rho_p} \mathbf{g}_{k,n} + \mathbf{n}_{p,k,n}, \end{aligned} \quad (3)$$

where $\mathbf{n}_{p,k,n} = \mathbf{N}_p \boldsymbol{\varphi}_{k,n}$ whose components are $\mathcal{CN}(0, 1)$. The MMSE estimate of the channel response $\mathbf{g}_{k,n}$ is

$$\begin{aligned} \hat{\mathbf{g}}_{k,n} &= \mathbb{E} \{ \mathbf{g}_{k,n} \check{\mathbf{y}}_{p,k,n}^H \} (\mathbb{E} \{ \check{\mathbf{y}}_{p,k,n} \check{\mathbf{y}}_{p,k,n}^H \})^{-1} \check{\mathbf{y}}_{p,k,n} \\ &= \frac{\sqrt{\tau_p \rho_p} \beta_k}{\tau_p \rho_p \beta_k + 1} \check{\mathbf{y}}_{p,k,n}. \end{aligned} \quad (4)$$

Let $\boldsymbol{\epsilon}_{k,n}$ be the channel estimation error, i.e.

$$\boldsymbol{\epsilon}_{k,n} = \mathbf{g}_{k,n} - \hat{\mathbf{g}}_{k,n}.$$

Then, due to the MMSE estimation property, $\hat{\mathbf{g}}_{k,n}$ and $\boldsymbol{\epsilon}_{k,n}$ are independent. Furthermore, we have that the elements of $\hat{\mathbf{g}}_{k,n}$ are i.i.d. $\mathcal{CN}(0, \gamma_k)$, where $\gamma_k \triangleq \frac{\tau_p \rho_p \beta_k^2}{\tau_p \rho_p \beta_k + 1}$, and the elements of $\boldsymbol{\epsilon}_{k,n}$ are i.i.d. $\mathcal{CN}(0, \beta_k - \gamma_k)$.

2) *Downlink Data Transmission:* Once the uplink training phase is completed and the base station has received adequate channel information, it will utilize these channel estimates to precode the symbols intended for K users, and broadcast the precoded signals to them.

Denote $\mathbf{x}_k \in \mathbb{C}^{N \times 1}$ as the signal intended for user k , where we assume that $\mathbb{E} \{ \mathbf{x}_k \mathbf{x}_k^H \} = \mathbf{I}_N$ and $\mathbb{E} \{ \mathbf{x}_k \mathbf{x}_{k'}^H \} = \mathbf{0}$. Let $\mathbf{W}_k$ be a $M \times N$ downlink precoding matrix associated with the k th user. Then, the transmitted signal at the base station can then be constructed as

$$\mathbf{s} = \sqrt{\rho_d} \sum_{k=1}^K \mathbf{W}_k \mathbf{x}_k, \quad (5)$$

where ρ_d is the normalized transmit power (normalized by the noise power), i.e.

$$\mathbb{E} \{ \|\mathbf{s}\|^2 \} = \rho_d \quad (6)$$

which is equivalent to

$$\sum_{k=1}^K \text{tr}(\mathbf{W}_k \mathbf{W}_k^H) = 1. \quad (7)$$

The k th user receives

$$\begin{aligned} \mathbf{y}_k &= \mathbf{G}_k^H \mathbf{s} + \mathbf{n}_k \\ &= \sqrt{\rho_d} \mathbf{G}_k^H \mathbf{W}_k \mathbf{x}_k + \sqrt{\rho_d} \mathbf{G}_k^H \sum_{k' \neq k}^K \mathbf{W}_{k'} \mathbf{x}_{k'} + \mathbf{n}_k, \end{aligned} \quad (8)$$

where $\mathbf{n}_k \in \mathbb{C}^{N \times 1}$ is noise vector with i.i.d. $\mathcal{CN}(0, 1)$ RVs. Then $\mathbf{x}_k$ will be detected from $\mathbf{y}_k$.

In this work, we consider MR processing since it is simple and works satisfactorily when M grows large, as in massive MIMO. Furthermore, it can be implemented in a distributed manner [11]. In the following, we will first present the conventional MR processing, and then provide our proposed scheme.

a) *Conventional MR Processing Scheme:* With conventional MR processing, the precoding matrix $\mathbf{W}_k$ is given by [7]

$$\begin{aligned}\mathbf{W}_k &= \frac{\hat{\mathbf{G}}_k}{\sqrt{K \mathbb{E} \left\{ \text{tr} \left(\hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \right) \right\}}} \\ &= \frac{\hat{\mathbf{G}}_k}{\sqrt{KM N \gamma_k}},\end{aligned}\quad (9)$$

where $\hat{\mathbf{G}}_k = [\hat{\mathbf{g}}_{k,1}, \dots, \hat{\mathbf{g}}_{k,N}] \in C^{M \times N}$.

From (8), the effective channel gain corresponding to the desired signal $x_{k,n}$ is

$$\mathbf{g}_{k,n}^H \mathbf{w}_{k,n} = \frac{1}{\sqrt{KM N \gamma_k}} \mathbf{g}_{k,n}^H \hat{\mathbf{g}}_{k,n}.$$

To detect $x_{k,n}$, user k relies on the channel hardening property of massive MIMO systems. More precisely, when the number of base station antennas is large, the effective channel gain $\mathbf{g}_{k,n}^H \mathbf{w}_{k,n}$ is very close to its mean value, and hence, user k uses $\mathbb{E} \left\{ \mathbf{g}_{k,n}^H \mathbf{w}_{k,n} \right\}$ as the true effective channel gain to detect the desired signal. However when the number of base station antennas is moderately large, the channel is less hardened and hence, the use of $\mathbb{E} \left\{ \mathbf{g}_{k,n}^H \mathbf{w}_{k,n} \right\}$ for signal detection is not good enough. Therefore, in the next section, we propose a new MR processing scheme which aims at hardening the effective channel gain regardless of the number of base station antennas.

b) *Proposed MR Processing Scheme:* To harden the effective channel gain $\mathbf{g}_{k,n}^H \mathbf{w}_{k,n}$, $\mathbf{w}_{k,n}$ should scale as $\mathbf{g}_{k,n} / \|\mathbf{g}_{k,n}\|^2$. This is due to the fact that $\mathbf{g}_{k,n}^H \mathbf{g}_{k,n} / \|\mathbf{g}_{k,n}\|^2 = 1$ regardless of the number of base station antennas. This requires that the base station have perfect channel state information. However the base station knows only the estimates of the channels. Therefore, we propose to use the following MR precoder:

$$\mathbf{W}_k = \alpha_k \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k}, \quad (10)$$

where

$$\mathbf{D}_{\hat{\mathbf{G}}_k} \triangleq \text{diag} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,1}\|^2}, \dots, \frac{1}{\|\hat{\mathbf{g}}_{k,N}\|^2} \right\},$$

and α_k is a normalization constant chosen to satisfy the power constraints (7). Assuming the large-scale fading is known at the users, then from (5), one particular choice of α_k to satisfy (7) is

$$\alpha_k^2 = \frac{(M-1) \gamma_k}{KN}. \quad (11)$$

III. ACHIEVABLE SPECTRAL EFFICIENCY

In this section, we derive the achievable downlink SE for the conventional and proposed MR techniques. To obtain these achievable SEs, we use the techniques of [7]. Using [7, Theo. 2], the achievable SE of the k th user in bit/s/Hz is

$$R_k = \frac{T - \tau_p}{T} \log_2 |\mathbf{I}_N + \bar{\mathbf{H}}_k^H \bar{\mathbf{\Xi}}_k \bar{\mathbf{H}}_k|, \quad (12)$$

where T is the coherence interval and

$$\bar{\mathbf{H}}_k \triangleq \sqrt{\rho_d} \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \right\}, \quad (13)$$

$$\bar{\mathbf{\Xi}}_k \triangleq (\mathbf{\Xi}_k - \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H)^{-1}, \quad (14)$$

and

$$\mathbf{\Xi}_k \triangleq \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k'=1}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} + \mathbf{I}_N. \quad (15)$$

In what follows, we derive closed-form expressions for the achievable SE (12) for the conventional and proposed MR processing, with finite M . Note that the achievable SE of the conventional MR processing was derived in [7]. But in [7], the authors derived a deterministic equivalent of the SE assuming that M and K go to infinity with their ratio bounded.

Theorem 1: With the conventional MR processing scheme, the achievable SE of the k th user is

$$R_k^{\text{conv}} = \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{\rho_d / (KN) M \gamma_k}{\rho_d \beta_k + 1} \right). \quad (16)$$

Proof: See Appendix A. $\square$

Theorem 2: With the proposed MR processing scheme, the achievable SE of the k th user is

$$\begin{aligned}R_k^{\text{prop}} &= \frac{T - \tau_p}{T} N \\ &\times \log_2 \left(1 + \frac{\rho_d / (KN) (M-1) \gamma_k}{\rho_d (\beta_k - \gamma_k / (KN)) + 1} \right).\end{aligned} \quad (17)$$

Proof: See Appendix B. $\square$

Remark 1: From Theorems 1 and 2, the main difference between two precoders is the gain uncertainty which comes from the channel hardening effect. More precisely, the gain uncertainties are $\rho_d \beta_k$ and $\rho_d (\beta_k - \gamma_k / (KN))$ for the conventional MR precoder and the proposed MR precoder, respectively. Clearly, if KN is small, the proposed scheme has much less gain uncertainty than the conventional one, and hence, does improve the system performance.

IV. NUMERICAL RESULTS

In this section, we provide numerical results to evaluate the performance of the proposed precoder compared with the performance to the conventional scheme. For all examples, we choose $T = 200$, and the pilot transmission duration is set equal to the number of data streams

Fig. 1. Simulated and analytical achievable sum SE versus M , for different K and N .

$\tau_p = KN$. Low values of K are chosen so that the differences of the precoders can be easily demonstrated.

We first verify the correctness of our closed-form expressions. For simplicity, we choose $\beta_k = 1$ for all k , $\rho_d = 10$ dB, and $\rho_p = 0$ dB. Fig. 1 presents the analytical results using the closed-form expressions (16) and (17) together with the simulation Monte-Carlo results using (12). It is shown that similar values are returned using the derived closed-form expressions to those of the simulation, therefore verifying the derived expressions and results as an accurate representation of the system. As explained, increasing the number of base station antennas will increase the channel hardening effect, therefore in Fig. 1 the relative difference between the two schemes is shown to decrease with increasing M . Furthermore, as expected, the performance improves noticeably by increasing the number of base station antennas for both MR schemes.

We now consider a more practical model for large-scale fading. More precisely, the large-scale fading accounts for path loss and shadowing and is generated by using the same method as in [9]. In this example, the maximum and minimum distances of the users from the base station of 1000 m and 100 m are selected, path loss exponent is set to 3.8 and a median cell edge SNR of 10 dB is chosen. Furthermore, we choose $\rho_d = 1W/N_0$, and $\rho_p = 0.2W/N_0$, where $N_0 = 6.36 \times 10^{-13}$ is the noise power.

The cumulative distributions of the sum achievable SEs for both the conventional and the proposed MR precoders for different K and N , at $M = 50$ are presented in Fig. 2. The cumulative distributions are generated from 10,000 random realizations of user locations. The figure validates the gains of using multiple antennas at each user, since the achievable SE is shown to increase with increasing N . The results showcase the profitability of the proposed precoding

Fig. 2. Cumulative distribution of the sum SEs, for different K and N , at $M = 50$.

scheme, since higher achievable SEs are produced in comparison to the conventional precoder for any N and K . In particular, when the total number of transmitted data streams NK increases, the performance difference between the proposed and conventional schemes reduces. This suggests that the proposed precoder will be best utilized in certain systems that contain small numbers of users and user antennas. Nevertheless, the proposed scheme systematically outperforms the conventional one even for higher number of K and N .

V. CONCLUSIONS

We have investigated the potential performance gains that a proposed MR precoding scheme could introduce into a massive MIMO system with multiple-antenna users. The analysis demonstrated that, regardless of the precoding scheme implemented, compelling performance gains were produced with additional user antennas. However, increasing the user antennas indefinitely would limit the performance due to contributing to the channel estimation overhead. Comparing the conventional and proposed precoding schemes, the latter was shown to produce better performance across all operating conditions. Specifically, it was concluded that the alternative proposed precoder works exceptionally well in a system that serves a low number of users, each with a small number of user antennas.

APPENDIX

A. Proof of Theorem 1

From the SE formula (12), we need to find the closed-form expressions for $\bar{\mathbf{H}}_k$ and Ξ_k . By substituting (9) into

(13),

$$\begin{aligned}\bar{\mathbf{H}}_k &= \sqrt{\rho_d} \mathbb{E} \left\{ \mathbf{G}_k \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \right\} \\ &= \frac{\sqrt{\rho_d M \gamma_k}}{\sqrt{KN}} \mathbf{I}_N,\end{aligned}\quad (18)$$

where (18) is obtained via the decomposition of $\mathbf{G}_k$ as $\hat{\mathbf{G}}_k + \mathbf{E}_k$, where $\mathbf{E}_k = [\epsilon_{k,1}, \dots, \epsilon_{k,N}]$, and using the fact that $\hat{\mathbf{G}}_k$ is independent of $\mathbf{E}_k$.

We next compute Ξ_k . We have

$$\begin{aligned}\Xi_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k'=1}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} + \mathbf{I}_N \\ &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} \\ &\quad + \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \right\} + \mathbf{I}_N.\end{aligned}\quad (19)$$

Denote by

$$\mathbf{A}_k \triangleq \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\}$$

and

$$\mathbf{B}_k \triangleq \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \right\}.$$

Now, we will find closed-form results for $\mathbf{A}_k$ and $\mathbf{B}_k$. Using (9), we have

$$\begin{aligned}\mathbf{A}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{\hat{\mathbf{G}}_{k'}}{\sqrt{KMN\gamma_{k'}}} \frac{\hat{\mathbf{G}}_{k'}^H}{\sqrt{KMN\gamma_{k'}}} \right) \mathbf{G}_k \right\} \\ &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{1}{KMN\gamma_{k'}} N \gamma_{k'} \mathbf{I}_M \right) \mathbf{G}_k \right\} \\ &= \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N.\end{aligned}\quad (20)$$

Similarly,

$$\begin{aligned}\mathbf{B}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \frac{\hat{\mathbf{G}}_k^H}{\sqrt{KMN\gamma_k}} \mathbf{G}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right)^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right) \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k + \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right. \\ &\quad \left. + \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k + \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \\ &\quad + \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} + \frac{\rho_d}{K} (\beta_k - \gamma_k) \mathbf{I}_N,\end{aligned}\quad (21)$$

where the fourth and fifth equalities are obtained by using the fact that $\hat{\mathbf{G}}_k$ and $\mathbf{E}_k$ are independent. We can see that the off-diagonal elements of $\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\}$ are zero, while its n th diagonal element is equal to

$$\begin{aligned}\mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= \mathbb{E} \left\{ \|\hat{\mathbf{g}}_{k,n}\|^4 \right\} + \mathbb{E} \left\{ \hat{\mathbf{g}}_{k,n}^H \left(\sum_{n' \neq n}^N \hat{\mathbf{g}}_{k,n'} \hat{\mathbf{g}}_{k,n'}^H \right) \hat{\mathbf{g}}_{k,n} \right\} \\ &= \gamma_k^2 M (M+1) + \gamma_k^2 M (N-1) \\ &= \gamma_k^2 M (M+N),\end{aligned}\quad (22)$$

where the second equality of (22) followed [12, Lemma 2.9]. Thus, $\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} = \gamma_k^2 M (M+N) \mathbf{I}_N$. As a consequence,

$$\mathbf{B}_k = \frac{\rho_d}{KN} (M\gamma_k + N\beta_k) \mathbf{I}_N. \quad (23)$$

Substituting (20) and (23) into (19), we obtain

$$\Xi_k = \left(\rho_d \beta_k + \frac{M \rho_d \gamma_k}{KN} + 1 \right) \mathbf{I}_N. \quad (24)$$

The substitution of (18) and (24) into (12) yields the desired result in Theorem 1.

B. Proof of Theorem 2

In order to derive the closed-form expression for the achievable SE with the proposed MR pre-coder, equations for terms $\bar{\mathbf{H}}_k$ and Ξ_k are required. Begin by computing $\bar{\mathbf{H}}_k$ by substituting (10) into (13),

$$\begin{aligned}\bar{\mathbf{H}}_k &= \sqrt{\rho_d} \alpha_k \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_k \right\} \\ &= \sqrt{\rho_d} \alpha_k \mathbf{I}_N.\end{aligned}\quad (25)$$

We next compute Ξ_k . We have

$$\begin{aligned}\Xi_k &= \rho_d \mathbb{E} \left\{ \underbrace{\mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k}_{\mathbf{C}_k} \right\} \\ &\quad + \underbrace{\rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \right\}}_{\mathbf{D}_k} + \mathbf{I}_N.\end{aligned}\quad (26)$$

• Compute $\mathbf{C}_k$:

By using (10), and since $\mathbf{G}_k$ is independent of $\hat{\mathbf{G}}_{k'}$ for $k' \neq k$, we have

$$\begin{aligned}\mathbf{C}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\alpha_{k'}^2 \hat{\mathbf{G}}_{k'} \mathbf{D}_{\hat{\mathbf{G}}_{k'}} \mathbf{D}_{\hat{\mathbf{G}}_{k'}}^H \hat{\mathbf{G}}_{k'}^H \right) \mathbf{G}_k \right\} \\ &= \rho_d \sum_{k' \neq k}^K \alpha_{k'}^2 \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{n=1}^N \frac{\hat{\mathbf{g}}_{k',n} \hat{\mathbf{g}}_{k',n}^H}{\|\hat{\mathbf{g}}_{k',n}\|^4} \mathbf{G}_k \right\}.\end{aligned}\quad (27)$$

We can see that the (i, j) th element of $\mathbf{C}_k$ is

$$\rho_d \sum_{k' \neq k}^K \alpha_{k'}^2 \mathbb{E} \left\{ \mathbf{g}_{k,i}^H \sum_{n=1}^N \frac{\hat{\mathbf{g}}_{k',n} \hat{\mathbf{g}}_{k',n}^H}{\|\hat{\mathbf{g}}_{k',n}\|^4} \mathbf{g}_{k,j} \right\},$$

which is equal to 0 when $i \neq j$ and

$$\rho_d \sum_{k' \neq 1}^K \alpha_{k'}^2 \frac{N \beta_k}{(M-1) \gamma_{k'}}$$

when $i = j$. Therefore,

$$\mathbf{C}_k = \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N. \quad (28)$$

• Compute $\mathbf{D}_k$:

We have

$$\begin{aligned} \mathbf{D}_k &= \rho_d \alpha_k^2 \mathbb{E} \left\{ \mathbf{G}_k^H \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \right) \mathbf{G}_k \right\} \\ &= \rho_d \alpha_k^2 \mathbb{E} \left\{ \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right)^H \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \right) \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right) \right\} \\ &= \rho_d \alpha_k^2 \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \right. \\ &\quad \left. + \mathbb{E} \left\{ \mathbf{E}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \right) \\ &= \rho_d \alpha_k^2 \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \\ &\quad + \rho_d \alpha_k^2 \frac{N(\beta_k - \gamma_k)}{(M-1) \gamma_k} \mathbf{I}_N, \quad (29) \end{aligned}$$

where the last equality follows from the same method which returns $\mathbf{C}_k$. It is obvious to show that $\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\}$ is a diagonal matrix whose (n, n) th element is equal to

$$\begin{aligned} &\mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= 1 + \mathbb{E} \left\{ \sum_{n' \neq n}^N \frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|^4} \hat{\mathbf{g}}_{k,n} \right\} \\ &= 1 + \mathbb{E} \left\{ \sum_{n' \neq n}^N \left| \frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|} \right|^2 \frac{1}{\|\hat{\mathbf{g}}_{k,n'}\|^2} \right\}. \quad (30) \end{aligned}$$

Conditioned on $\hat{\mathbf{g}}_{k,n}$, $\frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|}$ is a Gaussian RV with zero mean and variance γ_k which does not depend on $\hat{\mathbf{g}}_{k,n'}$. Since Gaussian RV depends only on the first and second moments, $\frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|}$ is a Gaussian RV independent of $\hat{\mathbf{g}}_{k,n'}$. As a result, (30) can be rewritten as

$$\begin{aligned} &\mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= 1 + \sum_{n' \neq n}^N \mathbb{E} \left\{ \left| \frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|} \right|^2 \right\} \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,n'}\|^2} \right\} \\ &= 1 + \sum_{n' \neq n}^N \gamma_k \frac{1}{(M-1) \gamma_k} = 1 + \frac{N-1}{M-1}, \quad (31) \end{aligned}$$

where in the second equality, we have used [12, Lemma 2.10]. Therefore,

$$\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} = \left(1 + \frac{N-1}{M-1} \right) \mathbf{I}_N. \quad (32)$$

Substituting (32) into (29), we obtain

$$\begin{aligned} \mathbf{D}_k &= \rho_d \alpha_k^2 \left(1 + \frac{N-1}{M-1} \right) \mathbf{I}_N \\ &\quad + \rho_d \alpha_k^2 \frac{N(\beta_k - \gamma_k)}{(M-1) \gamma_k} \mathbf{I}_N. \quad (33) \end{aligned}$$

The substitution of (28) and (33) into (26) yields

$$\mathbf{\Xi}_k = \rho_d \beta_k \mathbf{I}_N + \rho_d \frac{M-2}{KN} \gamma_k \mathbf{I}_N + \mathbf{I}_N. \quad (34)$$

From (12), (25), and (34), we obtain the desired result (17).

ACKNOWLEDGMENT

The work of H. Q. Ngo was supported by the UK Research and Innovation Future Leaders Fellowships under Grant MR/S017666/1. The work of M. Matthaiou was supported by the RAEng/The Leverhulme Trust Senior Research Fellowship under Grant LTSRF1718\14\2 and by a research grant from the Department for the Economy Northern Ireland under the US-Ireland R&D Partnership Programme.

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, Nov. 2010.
- [2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [5] Q. Zhang, S. Jin, K. K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Select. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
- [6] X. Liu, Y. Li, L. Xiao, and J. Wang, "Performance analysis and power control for multi-antenna V2V underlay massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4374–4387, Jul. 2018.
- [7] X. Li, E. Björnson, S. Zhou, and J. Wang, "Massive MIMO with multi-antenna users: When are additional user antennas beneficial?" in *Proc. IEEE Int. Conf. Telecommun. (ICT)*, May 2016, pp. 1–6.
- [8] X. Wu and D. Liu, "Novel insight into multi-user channels with multi-antenna users," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 1961–1964, Sep. 2017.
- [9] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 16, no. 5, pp. 2921–2935, May 2017.
- [10] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3–4, pp. 154–655, Oct. 2017.
- [11] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, UK: Cambridge University Press, 2016.
- [12] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.