**QUEEN'S
UNIVERSITY
BELFAST**

Hardening the Channels by Precoder Design in Massive MIMO with Multiple-Antenna Users

Sutton, J., Ngo, H-Q., & Matthaiou, M. (2021). Hardening the Channels by Precoder Design in Massive MIMO with Multiple-Antenna Users. *IEEE Transactions on Vehicular Technology*.

Published in:

IEEE Transactions on Vehicular Technology

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

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

Publisher rights

Copyright 2021 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.

Hardening the Channels by Precoder Design in Massive MIMO with Multiple-Antenna Users

James A. C. Sutton, Hien Quoc Ngo, *Senior Member, IEEE*, and Michail Matthaiou, *Senior Member, IEEE*

Abstract—In this paper, we investigate the potential of a novel precoder for a massive multiple-input multiple-output (MIMO) system with multiple-antenna users under the influence of both Rayleigh and keyhole channel models. The proposed precoder is based on maximum-ratio processing and aims to enhance the channel hardening effect at the users. This is achieved through dividing each channel vector by its norm square. Closed-form expressions for the achievable spectral efficiency of the conventional and proposed precoders are derived. If modeled with Rayleigh fading, due to the physical hardening property of this channel, a moderate gain in performance is produced by the proposed precoding method, where the highest spectral efficiency gain occurs for a small number of users and user antennas. Under Rayleigh fading, the results also demonstrate that similar performance is returned by increasing the users or user antennas. By comparison, the proposed precoder is shown to produce significant performance gains in the keyhole channels under the condition that the system is limited to single-antenna users. Corresponding to this, the results showcase how the performance of massive MIMO under keyhole channels is heavily dependent on increasing the number of served users over user antennas.

Index Terms—Keyhole channel, massive MIMO, maximum-ratio processing, multiple-antenna users.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) wireless communication systems have displayed breakthrough advantages in communication systems and, as such, have become a core technology for the fifth generation (5G) wireless communication systems and beyond [2], [3]. Key improvements in throughput and reliability are achieved from massive MIMO through equipping the base station (BS) with a large number of antennas compared to the number of served users.

Due to these advantages, significant progress has been made on the research of massive MIMO [4]–[6]. Despite this, the research progress of multiple-antenna user massive MIMO is relatively limited. The biggest body of the related literature has focused on single-antenna user systems since multiple-antenna users are size and cost limited. Recent work, such as [7] and [8], continues to develop single-antenna user massive MIMO systems. However, modern wireless systems are capable of supporting multiple-antenna users, leaving it

an unexplored area of investigation. Multiple-antenna users are more desirable since additional data streams have the potential to extend massive MIMO systems' benefits even further due to additional multiplexing gains. It is for these reasons that the multiple-antenna user massive MIMO model is both more practical and desirable for future investigation; hence, this observation has motivated us to pursue research on this topic.

In this space, we have identified the following relevant works: The authors of [9] demonstrated how beamforming techniques could provide further performance improvements to a system with multiple-antenna users, through the application of a hybrid architecture at the BS. In [10], a system model of multi-user channels with multiple-antenna users was investigated for both spatially correlated Rayleigh and geometry-based channels. The results demonstrated that, if properly modeled in a geometry-based channel, the inter-user interference present in the Rayleigh channel would be removed. Utilizing maximum-ratio (MR) precoding, deterministic equivalent derivations were presented in [11] for the downlink spectral efficiency (SE) of massive MIMO systems with multiple-antenna users. From the derived expressions, the paper was able to showcase the performance benefit of multiple-antenna users in massive MIMO; more specifically, the performance gain becomes more significant in systems serving a small number of users.

Conventional MR linear signal processing schemes are well documented in previous research of massive MIMO and have even been investigated in the specific case of massive MIMO with multiple-antenna users [11]. More recent works, such as [7] and [8], have produced papers for massive MIMO under Rician fading. The work of [7] introduced a new pilot allocation algorithm to decrease the complexity of achieving optimal pilot allocation in Rician channels. Moreover, [8] analyzed the downlink performance of multicell massive MIMO over Rician fading, with and without beamforming training. Nevertheless, neither of these papers consider the multiple-antenna user model or the weak channel hardening in non-trivial channel models, such as keyhole channels. We now recall that MR processing relies on the channel hardening property. The work of [12] stated that under Rayleigh fading channels, the channel hardening property of massive MIMO will occur as the BS antennas increases infinitely. However, when the number of BS antennas is moderate or small, the level of channel hardening reduces [12], [13]. In addition, both single and multiple-antenna user massive MIMO have been modeled using Rayleigh fading models in past research [2], [4], [5], [14]. Rayleigh fading channel models are commonly used in massive MIMO research since they often simplify the

The authors are with the 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)

The work of H. Q. Ngo was supported by the U.K. Research and Innovation Future Leaders Fellowships under Grant MR/S017666/1. The work of M. Matthaiou was supported by the EPSRC, U.K., under Grant EP/P000673/1 and by a research grant from the Department for the Economy Northern Ireland under the US-Ireland R&D Partnership Programme. Parts of this work were presented at the IEEE PIMRC, Sept. 2019 [1].

analysis, regularly occur in practical deployments, whilst from a mathematical point of view, they offer channel hardening and favorable propagation [12]. As previously mentioned, massive MIMO has also been investigated under Rician fading in [7], [8] and [15] among others. Consequently, the analysis of massive MIMO under alternative channel models remains insufficiently researched. More importantly, other channel models, such as keyhole channels, do not offer channel hardening [11].

Motivated by the aboved discussion, we exploit the performance of multiple-antenna user massive MIMO under both Rayleigh and keyhole channels which occur in rich scattering environments linked via low rank connections [12]. The work of [12] demonstrated examples of how the keyhole channel can appear in practical scenarios. Some illustrative examples are when: the transmitter and receiver are separated by a large distance, tunnel or the receiver is located inside a building. Although the properties of keyhole channels have been documented in [16]–[19], the effects on massive MIMO have been vastly unexplored. The main contributions of this paper are summarized as follows:

- We propose a novel precoder which avails of the MR technique and aims to enhance the channel hardening effect in massive MIMO systems. We point out that enhancing the channel hardening effect in systems with limited hardening capabilities will ultimately boost the performance of massive MIMO.
- We derive closed-form expressions for the conventional and proposed MR precoders for multiple-antenna user massive MIMO under both Rayleigh fading and keyhole channels. These closed-form results hold for any finite number of BS antennas making them useful for future development of massive MIMO. The derived expressions for the keyhole model provide general closed-form equations, allowing for single-keyhole and rich scattering Rayleigh channels to be represented as special case scenarios.
- Using the closed-form SE expressions, the potential performance gains produced due to the proposed precoder are exploited across a range of operating conditions. Additionally, the behavior of different massive MIMO setups with multiple-antenna users modeled using the keyhole channel model is investigated.

Notation: Column vectors and matrices are represented by bold lower case and upper case letters, respectively. The superscripts $()^*$ and $()^H$ express the conjugate and conjugate-transpose, while the trace and determinant of a matrix are denoted as $\text{tr}()$ and $\det(\cdot)$, respectively. The expectation, absolute value and the Euclidean norm are listed as $\mathbb{E}\{\cdot\}$, $|\cdot|$ and $\|\cdot\|$, respectively. Also, the operator $\mathbb{V}\text{ar}\{\cdot\}$ denotes the variance of a given term. 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 massive MIMO system containing a BS equipped with M antennas serving K users, where each user

consists of N antennas. Let $\mathbf{G}_k \in \mathbb{C}^{M \times N}$ be the channel response matrix between the BS and the k th user. Then,

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

where β_k is the large-scale fading component and matrix $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ models the small-scale fading between user k and the BS.

Time-division duplex (TDD) mode is considered. With TDD, each coherence is divided into three main parts: uplink training, uplink data transmission, and downlink data transmission. In our work, the uplink data transmission is ignored and we focus the analysis exclusively on the downlink transmission.¹ The following sections detail the uplink training and downlink payload data transmission phases, together producing each coherence interval.

1) *Uplink Training:* The uplink training phase involves the following. Pilot sequences of length τ_p are sent to the BS from their respective users. The pilot sequence transmitted from the n th antenna of the k th user is denoted by $\varphi_{k,n} \in \mathbb{C}^{\tau_p \times 1}$. It is assumed that all pilot sequences are pairwise orthogonal, i.e., $\varphi_{k,n}^H \varphi_{k',n'} = 0$ if $(k,n) \neq (k',n')$, which requires $\tau_p \geq KN$. Furthermore, we assume that $\|\varphi_{k,n}\|^2 = 1, \forall k = 1 \dots K, n = 1 \dots N$. The $M \times \tau_p$ pilot matrix received at the BS is

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

where $\mathbf{g}_{k,n}$ is the n th column of $\mathbf{G}_k$, which represents the channel vector between the n th antenna of the k th user and the M BS antennas, ρ_p is the normalized power of each pilot symbol transmitted and $\mathbf{N}_p$ is an $M \times \tau_p$ noise matrix whose elements are independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ random variables (RVs).

The pilot sequence $\mathbf{Y}_p$ received at the BS is projected onto $\varphi_{k,n}$, allowing for the channel vector for each antenna of a user to be estimated,

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

where $\check{\mathbf{n}}_{p,k,n} = \mathbf{N}_p \varphi_{k,n}$ whose components are $\mathcal{CN}(0, 1)$ RVs. Given $\check{\mathbf{y}}_{p,k,n}$, the linear minimum mean square error (MMSE) estimate of the channel response $\mathbf{g}_{k,n}$ is²

$$\hat{\mathbf{g}}_{k,n} = \mathbb{E}\{\mathbf{g}_{k,n} \check{\mathbf{y}}_{p,k,n}^H\} (\mathbb{E}\{\check{\mathbf{y}}_{p,k,n} \check{\mathbf{y}}_{p,k,n}^H\})^{-1} \check{\mathbf{y}}_{p,k,n}. \quad (4)$$

2) *Downlink Data Transmission:* After the training phase, channel estimates are employed to precode the symbols for each user before the downlink transmission.

Let $\mathbf{x}_k \in \mathbb{C}^{N \times 1}$ be the symbol vector intended for user k , where $\mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\} = \mathbf{I}_N$ and $\mathbb{E}\{\mathbf{x}_k \mathbf{x}_{k'}^H\} = \mathbf{0}$, for $k \neq k'$, are

¹Note that the downlink and uplink transmission are different. Thus, we cannot directly apply our downlink analysis to the uplink transmission. The uplink transmission just follows the same conventional transmission protocol as, for example, in [2]. Therefore, it is not investigated herein.

²The MMSE estimation is a standard technique in massive MIMO [2], [5], [6], [11], [12], [20]. A fundamental property of MMSE allows the estimation error and the channel estimate to be uncorrelated or in the case of Rayleigh fading statistically independent. This property simplifies the analysis substantially and it is the main reason that we choose MMSE over other channel estimation techniques, such as the one in [21].

assumed.³ We consider the $M \times N$ matrix $\mathbf{W}_k$ to express the downlink precoding matrix for the k th user. Then, the signal transmitted from the BS 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.,

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

Alternatively, (6) can be expressed as

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

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 the noise vector at user k with i.i.d. $\mathcal{CN}(0, 1)$ RVs.

In this work, we consider MR processing due to its: (a) low computational complexity; (b) substantial performance gain as M increases; (c) ability to be implemented in a distributed manner [2]. One particular application of MR processing in distributed systems is cell-free massive MIMO [23], [24]. Other linear processing techniques, such as zero-forcing (ZF) or regularized ZF (RZF), may enhance the system performance. However, there exists conditions that favour the use of MR processing, whilst the potential performance gain from using these schemes must be balanced with the resulting increased requirements. Compared to the MR precoder, ZF or RZF schemes have higher computational complexity (due to the involved matrix inversion), power consumption and backhaul requirements, especially in distributed systems. Furthermore, at low signal-to-noise ratio (SNR) conditions, MR yields superior performance compared to ZF. Compared to MR precoding, ZF does not work well with low-resolution analog-to-digital converters (ADCs) [25] and its effectiveness is greatly reduced in scenarios with very few active users [26]. The work of [26] provides a more complete analysis of MR and ZF, and also proposes a criterion for choosing between the MRC or the ZF precoders. Thus, depending on the operating conditions, system limitations and performance requirements, MR or (regularized) ZF precoders can be deployed. Our work considers a system with the low complexity MR precoder. The performance comparison of the MR and RZF precoders is performed via simulations in Section VI.

³Note that we consider a system with different data streams to different antennas at user k because it has the advantage of simplifying the receiver. In general, the number of data streams (the dimension of $\mathbf{x}_k$) can be smaller than N . In this case, block-diagonal precoding [22] can be used. If block-diagonal precoding is applied, the users will need to know the channels. In massive MIMO, to acquire channel estimation information (CSI) at the users, renders the channel estimation overhead very high. Investigating massive MIMO performance with block-diagonal precoder, taking into account imperfect channel estimation at the users and channel estimation overhead, is interesting and left for future work.

III. CONVENTIONAL AND PROPOSED MR PRECODERS

In this section, we summarize the conventional MR precoder and propose a novel MR precoder which aims at increasing the level of channel hardening at the users.

A. Conventional MR Precoders

Conventionally, the precoding matrix $\mathbf{W}_k$ is given by⁴

$$\mathbf{W}_k = \frac{\hat{\mathbf{G}}_k}{\sqrt{K \mathbb{E} \{ \text{tr}(\hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H) \}}}, \quad (9)$$

where $\hat{\mathbf{G}}_k = [\hat{\mathbf{g}}_{k,1}, \dots, \hat{\mathbf{g}}_{k,N}] \in \mathbb{C}^{M \times N}$. The normalization factor (the denominator) of (9) is chosen to meet the transmit power constraint (7).

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

$$\mathbf{g}_{k,n}^H \mathbf{w}_{k,n} = \frac{1}{\sqrt{K \mathbb{E} \{ \text{tr}(\hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H) \}}} \mathbf{g}_{k,n}^H \hat{\mathbf{g}}_{k,n},$$

where $\mathbf{w}_{k,n}$ is the n th column of $\mathbf{W}_k$.

The channel hardening property of massive MIMO systems is essential for user k to detect its desired signal $x_{k,n}$. With sufficient channel hardening, the effective channel gain $\mathbf{g}_{k,n}^H \mathbf{w}_{k,n}$ will be very close to its mean value. The implication of this is that user k can use $\mathbb{E} \{ \mathbf{g}_{k,n}^H \mathbf{w}_{k,n} \}$ as the true effective channel gain to detect $x_{k,n}$. However, a system with an inadequate number of BS antennas will experience weak channel hardening. In addition, even when the number of BS antennas is large, some propagation environments (e.g. keyhole channels) do not offer sufficient channel hardening [12]. Therefore, in these cases, the use of $\mathbb{E} \{ \mathbf{g}_{k,n}^H \mathbf{w}_{k,n} \}$ as $\mathbf{g}_{k,n}^H \mathbf{w}_{k,n}$ is problematic for signal detection. This is precisely the main reason behind developing a new MR precoder: with the purpose of hardening the effective channel gain of the desired signals received at each user, regardless of the number of base station antennas and propagation environments.

B. Proposed MR Precoders

If the precoding vector $\mathbf{w}_{k,n}$ scales with $\mathbf{g}_{k,n}/\|\mathbf{g}_{k,n}\|^2$ (i.e. $\mathbf{w}_{k,n} = c_{k,n} \mathbf{g}_{k,n}/\|\mathbf{g}_{k,n}\|^2$, where $c_{k,n}$ is deterministic), then the effective channel gain becomes

$$\mathbf{g}_{k,n}^H \mathbf{w}_{k,n} = c_{k,n}, \quad (10)$$

which is inherently hardened. For this choice, $\mathbf{g}_{k,n}^H \mathbf{w}_{k,n} = c_{k,n}$ will occur regardless of the number of BS antennas or propagation environments. Using this important observation, and taking into account the fact that the BS knows only the estimates of the channels, we propose the following novel MR precoder:

$$\mathbf{W}_k = q_k \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k}, \quad (11)$$

⁴There are many types of MR precoders which are linked to the different normalizations as acknowledged in [1]. Due to its common usage in literature such as [2], [11] and allowing for a tractable form for the analysis, we have chosen (9) as the precoder for our work. The investigations for the other types of MR precoding are left for future research.

where

$$\mathbf{D}_{\hat{\mathbf{G}}_k} \triangleq \text{diag} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,1}\|^2}, \dots, \frac{1}{\|\hat{\mathbf{g}}_{k,N}\|^2} \right\}, \quad (12)$$

and q_k is a normalization constant chosen to satisfy the power constraint (7), given by

$$q_k = \frac{1}{\sqrt{K \mathbb{E} \left\{ \text{tr} \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \right)^H \right) \right\}}}. \quad (13)$$

Different from the previous normalized MR in the literature, where the channel vector is divided by its norm, our proposed precoder (11) is created by dividing the channel vector by its norm square. The aim of the previous normalized MR is to have a short-term power constraint [27], while our proposed scheme aims to harden the channels.

IV. ACHIEVABLE SPECTRAL EFFICIENCY

In this section, we derive closed-form expressions for the spectral efficiencies of the conventional and proposed MR precoders, under Rayleigh fading and keyhole channels.

From (8) and [11, Theorem 2], the achievable SE of the k th user in bit/s/Hz is,

$$R_k = \frac{T - \tau_p}{T} \log_2 \left(\det \left(\mathbf{I}_N + \bar{\mathbf{H}}_k^H \bar{\mathbf{\Xi}}_k \bar{\mathbf{H}}_k \right) \right), \quad (14)$$

where T is the coherence interval,

$$\bar{\mathbf{H}}_k \triangleq \sqrt{\rho_d} \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \right\}, \quad (15)$$

$$\bar{\mathbf{\Xi}}_k \triangleq \left(\bar{\mathbf{\Xi}}_k - \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H \right)^{-1}, \quad (16)$$

and

$$\bar{\mathbf{\Xi}}_k \triangleq \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k'=1}^K \left(\mathbf{W}_{k'} \mathbf{W}_{k'}^H \right) \mathbf{G}_k \right\} + \mathbf{I}_N. \quad (17)$$

The following sections detail the channel models for the Rayleigh fading and keyhole channels. For each type of channel, closed-form expressions for the achievable SE (14) are derived for both conventional and proposed MR precoders.

A. Rayleigh Fading Channels

For Rayleigh fading channels, the small-scale fading matrix $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ in (1) includes i.i.d. $\mathcal{CN}(0, 1)$ RVs. Therefore, the channel estimate $\hat{\mathbf{g}}_{k,n}$ in (4) becomes

$$\begin{aligned} \hat{\mathbf{g}}_{k,n} &= \frac{\sqrt{\tau_p \rho_p} \beta_k}{\tau_p \rho_p \beta_k + 1} \tilde{\mathbf{y}}_{p,k,n} \\ &= \sqrt{\tau_p \rho_p} \alpha_k \mathbf{g}_{k,n} + \alpha_k \tilde{\mathbf{n}}_{p,k,n}, \end{aligned} \quad (18)$$

where $\alpha_k = \frac{\sqrt{\tau_p \rho_p} \beta_k}{\tau_p \rho_p \beta_k + 1}$. The channel estimation error is denoted as $\boldsymbol{\epsilon}_{k,n}$, where

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

In the case of Rayleigh fading, then due to the MMSE estimation property $\hat{\mathbf{g}}_{k,n}$ and $\boldsymbol{\epsilon}_{k,n}$ are independent. Moreover, 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. $\mathcal{CN}(0, \beta_k - \gamma_k)$.

1) *Conventional MR*: By substituting (18) into (9), the precoding matrix corresponding to the conventional MR scheme is given by

$$\mathbf{W}_k = \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}}. \quad (19)$$

The closed-form expression for the SE of the conventional MR processing in Rayleigh fading channels is provided in the following proposition.

Proposition 1: With the conventional MR precoders, the achievable SE of the k th user under Rayleigh fading channels is

$$R_k^{\text{R,conv}} = \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{M\gamma_k \rho_d / (KN)}{\rho_d \beta_k + 1} \right). \quad (20)$$

In (20), the superscripts “R” and “conv” stand for “Rayleigh fading” and “conventional”, respectively.

Proof: See Appendix A. $\square$

Remark 1: The above result is a generalization of the result with single-antenna users in massive MIMO literature. More precisely, if we set $N = 1$, the SE in (20) becomes the one in [2, Eq. (3.63)]. The SE of the conventional MR precoders for the systems with multiple antennas at the users was studied in [11]. However, [11] derived the deterministic equivalent result, when the numbers of BS antennas and users go to infinity with their ratio fixed. On the other hand, our result is obtained for finite numbers of BS antennas M and users K .

2) *Proposed MR*: By substituting (18) into (11), the precoding matrix corresponding to the proposed MR schemes is

$$\mathbf{W}_k = q_k \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k}, \quad (21)$$

where $q_k = \sqrt{\frac{(M-1)\gamma_k}{KN}}$. Then, we have the following closed-form expression for the SE.

Proposition 2: With the proposed MR precoders, the achievable SE of the k th user under Rayleigh faded channels is

$$\begin{aligned} R_k^{\text{R,prop}} &= \frac{T - \tau_p}{T} N \\ &\times \log_2 \left(1 + \frac{(M-1)\gamma_k \rho_d / (KN)}{\rho_d (\beta_k - \gamma_k / (KN)) + 1} \right). \end{aligned} \quad (22)$$

In (22), the superscript “prop” stands for “proposed”.

Proof: See Appendix B. $\square$

B. Keyhole Channels

Keyhole channels appear in scenarios where there are low-rank connections between the two rich scattering environments around the BS and users. Following the keyhole channel model of [28], the $M \times N$ channel matrix between the k th user and the BS is given by

$$\mathbf{G}_k = \sqrt{\beta_k} \sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} \mathbf{a}_j^{(k)H}, \quad (23)$$

where n_k is the number of effective keyholes, $\mathbf{a}_j^{(k)} \in \mathbb{C}^{N \times 1}$ is the random channel vector from the k th user to the j th keyhole, $\mathbf{b}_j^{(k)} \in \mathbb{C}^{M \times 1}$ is the random channel vector between the j th keyhole associated with the k th user and the BS,

and $c_j^{(k)}$ is the deterministic complex gain of the j th keyhole associated with the k th user. The elements of $\mathbf{b}_j^{(k)}$ and $\mathbf{a}_j^{(k)}$ are i.i.d. $\mathcal{CN}(0, 1)$ RVs. The gains $c_j^{(k)}$ are normalized such that $\mathbb{E} \left\{ |g_{k,n}^m|^2 \right\} = \beta_k$; $g_{k,n}^m$ is the channel response between the m th BS antenna and the k th user's n th antenna, hence,

$$\sum_{j=1}^{n_k} |c_j^{(k)}|^2 = 1. \quad (24)$$

The n th column of $\mathbf{G}_k$ is the $M \times 1$ channel vector between the BS and the n th antenna of user k , and can be constructed as

$$\mathbf{g}_{k,n} = \sqrt{\beta_k} \sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*}. \quad (25)$$

Substituting the keyhole channel model matrix into (4) allows the MMSE estimate for keyhole channels to be found as in the following lemma.

Lemma 1: Under keyhole channels, the MMSE estimate of $\mathbf{g}_{k,n}$ is

$$\begin{aligned} \hat{\mathbf{g}}_{k,n} &= \frac{\sqrt{\tau_p \rho_p} \beta_k}{\tau_p \rho_p \beta_k + 1} \tilde{\mathbf{y}}_{p,k,n} \\ &= \sqrt{\tau_p \rho_p} \alpha_k \mathbf{g}_{k,n} + \alpha_k \tilde{\mathbf{n}}_{p,k,n}, \end{aligned} \quad (26)$$

where $\alpha_k = \frac{\sqrt{\tau_p \rho_p} \beta_k}{\tau_p \rho_p \beta_k + 1}$.

Proof: See Appendix C. $\square$

Different from the Rayleigh fading cases, where the channel estimate and the channel estimation error are independent, in keyhole channels, $\hat{\mathbf{g}}_{k,n}$ and $\boldsymbol{\epsilon}_{k,n}$ are uncorrelated but not independent. This is because the keyhole channels are not Gaussian distributed. The elements of $\hat{\mathbf{g}}_{k,n}$ are zero mean RVs with variance γ_k , where $\gamma_k \triangleq \frac{\tau_p \rho_p \beta_k^2}{\tau_p \rho_p \beta_k + 1}$, while the elements of $\boldsymbol{\epsilon}_{k,n}$ are zero mean RVs with variance $(\beta_k - \gamma_k)$.

1) *Conventional MR:* The following precoding matrix is produced through the application of the keyhole channel to (9)

$$\mathbf{W}_k = \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}}. \quad (27)$$

Proposition 3: With the conventional MR processing scheme, the achievable SE of the k th user under keyhole channels is

$$\begin{aligned} R_k^{\text{K,conv}} &= \frac{T - \tau_p}{T} N \\ &\times \log_2 \left(1 + \frac{M\gamma_k \rho_d / (KN)}{(MN + 1)\gamma_k \frac{\rho_d}{KN} \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + \rho_d \beta_k + 1} \right). \end{aligned} \quad (28)$$

In (28), the superscript ‘‘K’’ stands for ‘‘Keyhole channels’’.

Proof: See Appendix D. $\square$

Remark 2: Let $|c_j^{(k)}|^2 = 1/n_k$, for all $j = 1, \dots, n_k$. Then, from (25), as $n_k \rightarrow \infty$, the keyhole channel approaches the Rayleigh fading channel. With this choice of $c_j^{(k)}$, we have

$$\sum_{j=1}^{n_k} |c_j^{(k)}|^4 = \frac{1}{n_k} \rightarrow 0, \text{ as } n_k \rightarrow \infty, \quad (29)$$

and hence, R_k in (28) becomes

$$\begin{aligned} R_k^{\text{K,conv}} &\rightarrow \frac{T - \tau_p}{T} N \\ &\times \log_2 \left(1 + \frac{M\gamma_k \rho_d / (KN)}{\rho_d \beta_k + 1} \right), \text{ as } n_k \rightarrow \infty \end{aligned} \quad (30)$$

which is identical to the SE expression for Rayleigh fading channels given by (20).

2) *Proposed MR:* With the proposed MR scheme, the normalization factor q_k in (11) is

$$q_k = \frac{1}{\sqrt{K \sum_{n=1}^N \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,n}\|^2} \right\}}} = \frac{1}{\sqrt{KNL_k}}, \quad (31)$$

where

$$L_k = \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,n}\|^2} \right\}. \quad (32)$$

For single-keyhole channels (i.e. $n_k = 1$), and perfect channel estimation, L_k does not converge. Thus, our proposed scheme is applied for the cases where $n_k > 1$. The exact closed-form expression for L_k is very hard, if not impossible, to obtain. Therefore, we apply a Taylor series expansion to produce the following approximation:

$$\begin{aligned} L_k &\approx \frac{1}{\mathbb{E} \{ \|\hat{\mathbf{g}}_{k,n}\|^2 \}} + \frac{1}{(\mathbb{E} \{ \|\hat{\mathbf{g}}_{k,n}\|^2 \})^3} \text{Var} \{ \|\hat{\mathbf{g}}_{k,n}\|^2 \} \\ &= \frac{M+1}{M^2} \frac{\gamma_k}{\beta_k^2} \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + \frac{M+1}{M^2} \frac{1}{\gamma_k}, \end{aligned} \quad (33)$$

where the detailed derivation of (33) is provided in Appendix E. Note that the approximation using the Taylor series technique is widely used in the massive MIMO literature, thanks to its accuracy, as articulated in [15]. The tightness of the above approximation will be also verified in the numerical results.

Proposition 4: With the proposed processing scheme, the achievable SE of the k th user under keyhole channel (for $n_k > 1$) is given by (34), shown at the top of the next page.

Proof: See Appendix F. $\square$

V. DISCUSSIONS

In this section, we provide some important insights obtained from the results in Section IV.

A. Conventional Scheme Vs. Proposed Scheme

- For Rayleigh fading channels, by comparing the results obtained in Propositions 1 and 2, we can see that the key difference between the spectral efficiencies of the conventional precoder and the proposed precoder is the gain uncertainties which are $\rho_d \beta_k$ and $\rho_d (\beta_k - \gamma_k / (KN))$ for the former and the latter, respectively. These gain uncertainties are produced due to the varying channel hardening effectiveness of either scheme. Note that, with a low number of users K and user antennas N , the gain uncertainty produced by the proposed scheme is much less compared to the conventional scheme, therefore improving system performance.

$$R_k^{\text{K,prop}} \approx \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{\frac{M^2}{M+1} \gamma_k \rho_d / (NK)}{\left(\frac{\gamma_k^2}{\beta_k^2} \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \left(\rho_d \frac{K-1}{K} \beta_k + \rho_d \frac{N-1}{KN} \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) + 1 \right)} \right). \quad (34)$$

$$R_k^{\text{K,prop}} = \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{1}{(N-1) \beta_k \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) \left(\frac{1}{\beta_k^2} \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) + 1 \right)} \right). \quad (36)$$

- For keyhole channels, with the conventional scheme, from Proposition 3, we can see that for a fixed n_k , as $M \rightarrow \infty$,

$$R_k^{\text{K,conv}} = \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{1}{N \sum_{j=1}^{n_k} |c_j^{(k)}|^4} \right). \quad (35)$$

The above result implies that even when the number of BS antennas goes to infinity the SE is bounded. Additionally, the upper bound depends only on the number of antennas at the user, and the keyhole gains. This is due to the fact that in keyhole channels, the channels do not harden, so the use of conventional MR scheme causes a large gain uncertainty, which scales as fast as the desired signal power does when M increases. Similarly, with the proposed scheme under keyhole channel models, we can see that for a fixed n_k as $M \rightarrow \infty$, Proposition 4 produces (36). Once again the SE is bounded even if the number of BS antennas increases to infinity. However, if the system is limited to single-antenna users, then when the number of BS increases to infinity the corresponding SE of Proposition 4 will increase without bound. This is due to the fact that the proposed MR scheme aims at hardening the channels, and hence, reduces the gain uncertainty significantly. This shows the fundamental benefit of our proposed scheme compared to the conventional scheme.

B. Keyhole Effect

The effect of the keyhole channel appears in the denominators of the SINRs in (28) for the conventional scheme, and in (34) for the proposed scheme which is

$$\kappa_k = \sum_{j=1}^{n_k} |c_j^{(k)}|^4.$$

We have that

$$\frac{1}{n_k} \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^2 \right)^2 \leq \kappa_k \leq \sum_{j=1}^{n_k} |c_j^{(k)}|^2, \quad (37)$$

where the upper bound is obtained from the fact that $|c_j^{(k)}|^2 \leq 1$, and the lower bound follows the Cauchy-Schwarz inequality. Using (24), we can rewrite (37) as

$$\frac{1}{n_k} \leq \kappa_k \leq 1. \quad (38)$$

The upper bound corresponds to the case of single-keyhole (degenerate keyhole) channels and is achieved when only one $|c_j^{(k)}|^2 = 1$ and the others equal 0, whereas the lower bound is achieved when $|c_j^{(k)}|^2 = 1/n_k$ for all j . From these results, we can obtain the following insights: (i) equal keyhole gain for a given number of keyholes n_k (i.e., $|c_j^{(k)}|^2 = 1/n_k$), serves as the best case scenario in the sense that it maximizes the SE; and (ii) a single-keyhole channel is the worst case which yields the smallest SE.

C. Effect of the Number of Base Station Antennas

For keyhole channels, if the proposed scheme is used and the system is limited to single-antenna users, then as $M \rightarrow \infty$, $R_k^{\text{K,prop}} \rightarrow \infty$. This means when the number of BS antennas grows without bound, the interuser interference and noise disappear. The reason is that although keyhole channels make the channel matrix become rank deficient, they still offer favorable propagation [12].

Furthermore, assume that $\rho_d = \rho_p = \rho/\sqrt{M}$, where ρ is fixed regardless of M . Then, from Propositions 1–4, as $M \rightarrow \infty$ we have

$$R_k^{\text{R,conv}} \rightarrow \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{\tau_p \rho^2 \beta_k^2}{KN} \right), \quad (39)$$

$$R_k^{\text{R,prop}} \rightarrow \frac{T - \tau_p}{T} N \log_2 \left(1 + \frac{\tau_p \rho^2 \beta_k^2}{KN} \right), \quad (40)$$

$$R_k^{\text{K,conv}} \rightarrow \frac{T - \tau_p}{T} N \times \log_2 \left(1 + \frac{\tau_p \rho^2 \beta_k^2}{\tau_p \rho^2 \beta_k^2 N \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + KN} \right), \quad (41)$$

$$R_k^{\text{K,prop}} \rightarrow \frac{T - \tau_p}{T} \log_2 \left(1 + \frac{\tau_p \rho^2 \beta_k^2}{K} \right). \quad (42)$$

The above results imply that, for all cases, when M is large, the transmit power at the BS can be made inversely

proportional to $\sqrt{M}$ with no reduction in SE. Power scaling laws in massive MIMO still hold under keyhole channels. Interestingly, with our proposed MR scheme, the SE of keyhole channels converges to the one of Rayleigh fading channels. This implies that under low transmit power, the effect of keyhole disappears when a large number of antennas is used.

D. Effect of the Number of Users/User Antennas

As discussed in Section V-C, if $M \gg NK$, then the interference will be cancelled out. In practice, M is large but is not much greater than NK . In this case, interference will have detrimental impact on the system performance. Thus, in this subsection, we investigate the effect of the number of users/user antennas. Let $\text{SINR}_k^{\text{R,conv}}$, $\text{SINR}_k^{\text{R,prop}}$, $\text{SINR}_k^{\text{K,conv}}$, and $\text{SINR}_k^{\text{K,prop}}$ be the effective SINRs corresponding to the SEs given by (20), (22), (28), and (34), respectively. Assume that $\eta_k = M/(NK)$ is fixed, as $M \rightarrow \infty$ (and hence, $NK \rightarrow \infty$), we have

$$\text{SINR}_k^{\text{R,conv}} \rightarrow \frac{\eta_k \gamma_k \rho_d}{\rho_d \beta_k + 1}, \quad (43)$$

$$\text{SINR}_k^{\text{R,prop}} \rightarrow \frac{\eta_k \gamma_k \rho_d}{\rho_d \beta_k + 1}, \quad (44)$$

$$\text{SINR}_k^{\text{K,conv}} \rightarrow \begin{cases} \frac{\eta_k N \gamma_k \rho_d}{\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + \rho_d \beta_k + 1}, & \text{if } N \text{ fixed} \\ 0, & \text{if } K \text{ fixed,} \end{cases} \quad (45)$$

$$\text{SINR}_k^{\text{K,prop}} \rightarrow \begin{cases} \frac{\eta_k \gamma_k \rho_d / \left(\frac{\gamma_k^2}{\beta_k^2} \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right)}{\rho_d \beta_k (N-1) \eta_k \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) + \rho_d \beta_k + 1}, & \text{if } N \text{ fixed} \\ 0, & \text{if } K \text{ fixed.} \end{cases} \quad (46)$$

We can see that, under Rayleigh fading channels, the SINRs of the conventional and proposed MR schemes converge to the same result. However, the asymptotic results of the conventional and proposed MR are different under keyhole channels. Interestingly, results (44) and (46) imply that, under keyhole channels, when the number of users is fixed, by increasing the number of antennas per user, the SINR will converge to 0.

The sum SE is affected by the channel estimation overhead $\frac{T-\tau_p}{T}$ (since $\tau_p \geq NK$), the degrees of freedom at the user side NK , and the interference. When NK increases, the degrees of freedom increase, but at the same time, channel estimation overhead and the interference also increase. Therefore, there is a fundamental performance tradeoff between these factors when we change the number of users/user antennas. Similar observations were made in [29, Myth 6].

VI. NUMERICAL RESULTS

Using the derived analytical expressions for the conventional and proposed precoders, numerical results are generated to analyze the performance of a massive MIMO with multiple-antenna users under both Rayleigh and keyhole channel models. For all examples, the pilot transmission duration is set equal to the number of data streams, i.e. $\tau_p = KN$ and a period of $T = 200$ is chosen.

Fig. 1. Simulated and analytical achievable sum SE versus N , for different K at $M = 100$ under Rayleigh fading channel models.

Fig. 2. Simulated and analytical achievable sum SE versus N , for different K at $M = 100$ and $n_k = 2$ under keyhole channel models.

Fig. 1 is produced to compare the analytical results formed using the derived closed-form equations (20) and (22) and simulation Monte-Carlo results using (14). For simplicity, we choose $\beta_k = 1$ for all k , $\rho_d = 10$ dB, and $\rho_p = 0$ dB. The results of Fig. 1 verify the correctness of our derived Propositions 1 and 2 to be verified. In addition, the result shows that when N increases, the sum SE first increases, reaches the maximum point, and then decreases. This behavior is as expected, and is explained in Section V-D.

Similar results are now produced in Fig. 2 for a keyhole channel model with 2 active keyholes to verify Propositions 3 and 4. As with Fig. 1, the closed-form expressions are compared to simulated results, where large-scale fading is ignored and the same transmission powers as Fig. 1 are used. The analytical results are verified as an accurate representation

Fig. 3. Cumulative distribution of the sum SEs, for different K and N , at $M = 50$ under Rayleigh fading channels.

Fig. 4. Cumulative distribution of the sum SEs, for different K and N , at $M = 50$ and $n_k = 2$ under keyhole channels.

of the system, since the approximation in Proposition 4 is shown to be tight to the simulation. Even when $K = 10$ (the case with the largest gap between the simulation and analysis), the simulation results are just about 10% larger than the analytical results. Fig. 2 demonstrates this relationship inside a keyhole channel. Counter to the previous figure, in Fig. 2 we can notice that as the number of user antennas is increased the performance of the proposed scheme deteriorates significantly. However, for very low values of user antennas, we see the proposed scheme produce huge performance gains; this effect is most exaggerated with single-antenna users. Therefore, under keyhole channel models the proposed scheme is significantly more desirable in single-antenna systems.⁵

A more practical model incorporating large-scale fading is now considered. A large-scale fading component accounting for both path-loss and shadowing is generated using a similar method as in [12]. Maximum and minimum distances of the users from the BS of 1000 m and 100 m are used, a path loss exponent of 3.8 is chosen and the median cell edge SNR is set at 10 dB. Also, normalized transmit powers of $\rho_d = 1W/N_0$ and $\rho_p = 0.2W/N_0$ are selected, where $N_0 = 6.36 \times 10^{-13}W$ is the noise power. These values are used for all future results producing cumulative distributions.

Fig. 3 presents the cumulative distributions of the sum achievable SEs for both conventional and proposed MR precoders using Rayleigh channel models. The distributions are

generated from 10,000 random realizations of user locations for different K and N , at $M = 50$. From Fig. 3 the following observations can be made: the results validate the SE gains produced through increasing user antennas. The advantages of the proposed processing scheme are also displayed, since when compared to the conventional scheme greater achievable SEs are produced under the same conditions of N and K . The performance difference in the two schemes is shown to decrease as the number of transmitted data streams (NK) grows. The results showcase how the proposed precoder will produce improved performance over a conventional MR scheme under all conditions. In particular, for Rayleigh fading channels the proposed precoder will be best utilized in systems with limited number of users and user antennas.

The results of Fig. 3 are repeated for a keyhole model with $n_k = 2$ and are displayed in Fig. 4. The proposed scheme can be seen to yield improved performance across single-antenna user scenarios. However, as with Fig. 2, increasing the user antennas results in extreme sum SE losses. Despite this feature, the proposed scheme remains superior. For instance, we can compare the performance of the proposed precoder with $K = 5$ and $N = 1$ to either conventional or proposed methods using $K = 5$ and $N = 2$ in Fig. 3. The results demonstrate that despite the additional user antennas, the proposed scheme under single-antenna operation yields significantly better performance. Therefore, under keyhole modeled channels, the proposed scheme with single-antenna users is an advantageous option. Comparing the single-antenna scenarios for Fig. 3 and Fig. 4, the performance gains produced by the proposed scheme are seen to be much larger in the keyhole channel model compared to Rayleigh fading. This result was expected since the enhancement of channel hardening will be exploited better by the keyhole channel. Also notice that in Fig. 4 as the number of served users increases, the percentage gain of the proposed precoder over the conventional grows. This means that in keyhole channels the proposed scheme will

⁵For simplicity, as in Fig. 2, we assume that $\beta_k = 1$, $c_j^{(k)} = 1/\sqrt{n_k}$, and we further assume that $\gamma_k = 1$ (perfect channel estimation). By equating (28) to (34), we obtain the cross-point N of Fig. 2 between the conventional and proposed under keyhole channels as

$$N^* = \frac{(r_k + M)\rho_d}{K n_k \left(\rho_d \left(\frac{K-1}{K} - r_k \right) + \frac{M \rho_d (1-r_k)}{K n_k} + 1 - r_k \right)},$$

where $r_k = \frac{M n_k}{(M+1)(n_k+1)}$. When $M \rightarrow \infty$, $N^* \rightarrow n_k + 1$. In other words, we can see that when M is large, the cross-point depends only on the number of keyholes, and that this cross-point increases when n_k increases.

Fig. 5. Cumulative distribution of the sum SEs, for different K and N , at $M = 100$.

Fig. 6. Cumulative distribution of the sum SEs, for different K and N , at $M = 100$, $n_k = 2$ under keyhole channels.

perform best in systems with a large number of single-antenna users. This conclusion contrasts with the findings for Rayleigh fading where low numbers of data streams are beneficial.

Fig. 5 demonstrates the difference in performance gain under Rayleigh fading of ten single antenna users and one user equipped with ten antennas, for both precoders. From the results we can conclude that under Rayleigh fading the performance difference between the two scenarios is very small. Therefore, there is no preferred setup between increased served users or user antennas for either precoders. The results also reflect how the difference between the two precoding methods is reduced as K and N increase.

Fig. 6 compares the performances of both precoders in keyhole channels. Once again a significant performance advantage can be noticed with the proposed precoder if single-

Fig. 7. Cumulative distribution of the sum SEs, for different number of N and K , where $M = 50$ and $n_k = 10$.

antenna users are present. Comparing the differences between utilizing ten single-antenna users with a single user equipped with ten antennas for each precoder returns a result similar to that of Fig. 4. Both schemes produce significantly worst performance with a single multiple-antenna user compared to many single-antenna users. However, the performance of the proposed scheme with single-antenna users is considerably better, making it preferable over all other operational setups.

In Fig. 7, the gains in performance from increasing the number of user antennas and users are compared when utilizing conventional precoding in both Rayleigh and keyhole channels. We can observe an improved performance after increasing either the number of users or user antennas. More specifically, comparing the performances of the Rayleigh fading channel when using ten single-antenna users or one user equipped with ten antennas, the SEs are seen to be similar. However, when utilizing a keyhole channel, multiple single-antenna users are seen to produce enhanced performance when compared to a single user with multiple-antennas. If the number of keyholes increases the advantage of increasing the users as opposed to user antennas will diminish. At that point, the keyhole channel behaves equivalently with a Rayleigh faded channel where similar performance is produced through higher numbers of users or user antennas.

Fig. 8 presents the similarities between Rayleigh and keyhole channels under a constraint of single-antenna users and using the proposed scheme. The performance of the proposed scheme in both channels can be seen to be comparable to one another, even when repeating the results with less keyholes. This is in contrast to the previous results using single-antenna users in Fig. 7, which returned greater performance losses with conventional precoding in keyhole channels. The similarities in performance of the two channels displays the effectiveness of the new precoder's channel hardening effect, particularly in the poorly hardened keyhole channel. In Fig. 8, for $K = 1$ and $N = 1$, the keyhole channel is seen to produce greater SE compared to the Rayleigh fading model. This phenomenon is

Fig. 8. Cumulative distribution of the sum SEs, for different number of N and K , where $M = 50$ and $n_k = 10$.

Fig. 9. Achievable sum SE versus the number of BS antennas M , for the conventional/proposed MR and RZF precoders, at $K = 5$ and $N = 1$, under keyhole channel models.

only present while $K = 1$ and $N = 1$ but does remain for any other number of keyholes.

Finally, we compare the performance of our proposed scheme with the RZF (as the benchmark) via simulations, as shown in Fig. 9. Here, the parameter setup is the same as the one in Fig. 2. The RZF is chosen as in [30], where the regularization parameter equals $\frac{1}{\rho_d}$ (which is optimal under perfect CSI). We can see from Fig. 9 that when M increases, the relative performance gap between the RZF and the proposed MR scheme reduces significantly. The reason is that this gap comes from two main factors: channel hardening and inter-user interference. When M increases, the inter-user interference reduces due to the favorable propagation property of massive MIMO, and hence, the main factor affecting the above gap is the channel hardening. Our proposed scheme hardens the channels, thereby reducing the above performance

gap significantly.

VII. CONCLUSIONS

We have conducted an analysis for a proposed MR precoding scheme using both Rayleigh fading and keyhole channel models, operating within a massive MIMO system with multiple-antenna users. With a Rayleigh channel model, the proposed scheme demonstrated a moderate performance enhancement. However, under a keyhole channel model the proposed scheme yields significantly more SE, provided the system is limited to single-antenna users. The analysis of massive MIMO system with multiple-antenna users revealed that under Rayleigh fading, similar performance gains are possible when increasing the number of users or user antennas. However, keyhole channels perform better with a higher number of users rather than with a higher number of user antennas. Note that there are practical constraints that were not considered in this paper. For example, in a practical system, the SE of massive MIMO will be limited by pilot contamination [31]. Additionally, massive MIMO systems are expected to make use of large numbers of low quality components [32], resulting in multiple hardware imperfections, such as the quantization noise caused by low resolution ADCs [33]. Therefore, future work on the proposed precoder should investigate the SE behavior against these constraints. It is finally worth mentioning the similarities shared between the low-rank channel model for a Reconfigurable Intelligent Surfaces (RIS) based system [34] and the keyhole channel model. Therefore, investigating the performance of our proposed scheme in RIS-based systems is interesting, and will be part of our future work.

APPENDIX

A. Proof of Proposition 1

By following the formula (14) for the SE, closed-form expressions for $\bar{\mathbf{H}}_k$ and Ξ_k are required. Substituting (9) into (15),

$$\begin{aligned}\bar{\mathbf{H}}_k &= \sqrt{\rho_d} \mathbb{E} \left\{ \mathbf{G}_k \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \right\} \\ &= \frac{\sqrt{\rho_d M \gamma_k}}{\sqrt{KN}} \mathbf{I}_N,\end{aligned}\quad (47)$$

where $\mathbf{E}_k = [\epsilon_{k,1}, \dots, \epsilon_{k,N}]$ and using the fact that $\hat{\mathbf{G}}_k$ is independent of $\mathbf{E}_k$, (47) is obtained via the decomposition of $\mathbf{G}_k$ as $\hat{\mathbf{G}}_k + \mathbf{E}_k$.

The derivation of Ξ_k is now required. We have

$$\begin{aligned}\Xi_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k'=1}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} + \mathbf{I}_N \\ &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} \\ &\quad + \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \right\} + \mathbf{I}_N.\end{aligned}\quad (48)$$

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

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

Closed-form expressions for $\mathbf{A}_k$ and $\mathbf{B}_k$ will now be produced. From (21)

$$\begin{aligned} \mathbf{A}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{\hat{\mathbf{G}}_{k'}}{\sqrt{KMN\gamma_{k'}}} \frac{\hat{\mathbf{G}}_{k'}^H}{\sqrt{KMN\gamma_{k'}}} \right) \mathbf{G}_k \right\} \\ &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{1}{KMN\gamma_{k'}} N\gamma_{k'} \mathbf{I}_M \right) \mathbf{G}_k \right\} \\ &= \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N. \end{aligned} \quad (49)$$

Similarly,

$$\begin{aligned} \mathbf{B}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \frac{\hat{\mathbf{G}}_k^H}{\sqrt{KMN\gamma_k}} \mathbf{G}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right)^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right) \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k + \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right. \\ &\quad \left. + \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k + \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \\ &\quad + \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \mathbf{E}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} + \frac{\rho_d}{K} (\beta_k - \gamma_k) \mathbf{I}_N. \end{aligned} \quad (50)$$

The fact that $\hat{\mathbf{G}}_k$ and $\mathbf{E}_k$ are independent, allows the fourth and fifth equalities from (50) to be achieved. The non-diagonal elements of $\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\}$ can be seen to be zero, while its n th diagonal element is equal to

$$\begin{aligned} &\mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= \mathbb{E} \left\{ \|\hat{\mathbf{g}}_{k,n}\|^4 \right\} + \mathbb{E} \left\{ \hat{\mathbf{g}}_{k,n}^H \left(\sum_{n' \neq n}^N \hat{\mathbf{g}}_{k,n'} \hat{\mathbf{g}}_{k,n'}^H \right) \hat{\mathbf{g}}_{k,n} \right\} \\ &= \gamma_k^2 M(M+1) + \gamma_k^2 M(N-1) \\ &= \gamma_k^2 M(M+N), \end{aligned} \quad (51)$$

Here, the second equality of (51) followed from [35, Lemma 2.9]. Therefore, $\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} = \gamma_k^2 M(M+N) \mathbf{I}_N$. From this we can obtain,

$$\mathbf{B}_k = \frac{\rho_d}{KN} (M\gamma_k + N\beta_k) \mathbf{I}_N. \quad (52)$$

Through substituting (49) and (52) into (48), we obtain

$$\mathbf{\Xi}_k = \left(\rho_d \beta_k + \frac{M\rho_d\gamma_k}{KN} + 1 \right) \mathbf{I}_N. \quad (53)$$

Substituting (47) and (53) into (14) yields the desired result in Proposition 1.

B. Proof of Proposition 2

Closed-form equations for both $\bar{\mathbf{H}}_k$ and $\mathbf{\Xi}_k$ are required in order to achieve the desired expression for the achievable SE with the proposed MR pre-coder. First, we determine $\bar{\mathbf{H}}_k$ by substituting (11) into (15),

$$\begin{aligned} \bar{\mathbf{H}}_k &= \sqrt{\rho_d} q_k \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_k \right\} \\ &= \sqrt{\rho_d} q_k \mathbf{I}_N. \end{aligned} \quad (54)$$

Next, we compute $\mathbf{\Xi}_k$. We have

$$\begin{aligned} \mathbf{\Xi}_k &= \rho_d \mathbb{E} \left\{ \underbrace{\mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k}_{\mathbf{C}_k} \right\} \\ &\quad + \underbrace{\rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \right\}}_{\mathbf{D}_k} + \mathbf{I}_N. \end{aligned} \quad (55)$$

- Compute $\mathbf{C}_k$:

Since $\mathbf{G}_k$ is independent of $\hat{\mathbf{G}}_{k'}$ for $k' \neq k$, then by using (11) we have

$$\begin{aligned} \mathbf{C}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(q_{k'}^2 \hat{\mathbf{G}}_{k'} \mathbf{D}_{\hat{\mathbf{G}}_{k'}} \mathbf{D}_{\hat{\mathbf{G}}_{k'}}^H \hat{\mathbf{G}}_{k'}^H \right) \mathbf{G}_k \right\} \\ &= \rho_d \sum_{k' \neq k}^K q_{k'}^2 \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{n=1}^N \frac{\hat{\mathbf{g}}_{k',n} \hat{\mathbf{g}}_{k',n}^H}{\|\hat{\mathbf{g}}_{k',n}\|^4} \mathbf{G}_k \right\}. \end{aligned} \quad (56)$$

The (i, j) th element of $\mathbf{C}_k$ can be presented as

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

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

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

when $i = j$. Therefore,

$$\mathbf{C}_k = \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N. \quad (57)$$

- Compute $\mathbf{D}_k$:

We have

$$\begin{aligned} \mathbf{D}_k &= \rho_d q_k^2 \mathbb{E} \left\{ \mathbf{G}_k^H \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \right) \mathbf{G}_k \right\} \\ &= \rho_d q_k^2 \mathbb{E} \left\{ \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right)^H \left(\hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \right) \left(\hat{\mathbf{G}}_k + \mathbf{E}_k \right) \right\} \\ &= \rho_d q_k^2 \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \right. \\ &\quad \left. + \mathbb{E} \left\{ \mathbf{E}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \mathbf{E}_k \right\} \right) \\ &= \rho_d q_k^2 \mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} \\ &\quad + \rho_d q_k^2 \frac{N(\beta_k - \gamma_k)}{(M-1)\gamma_k} \mathbf{I}_N, \end{aligned} \quad (58)$$

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

$$\begin{aligned} & \mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= 1 + \mathbb{E} \left\{ \sum_{n' \neq n}^N \hat{\mathbf{g}}_{k,n}^H \frac{\hat{\mathbf{g}}_{k,n'} \hat{\mathbf{g}}_{k,n'}^H}{\|\hat{\mathbf{g}}_{k,n'}\|^4} \hat{\mathbf{g}}_{k,n} \right\} \\ &= 1 + \mathbb{E} \left\{ \sum_{n' \neq n}^N \left| \frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|} \right|^2 \frac{1}{\|\hat{\mathbf{g}}_{k,n'}\|^2} \right\}. \end{aligned} \quad (59)$$

Conditioned on $\hat{\mathbf{g}}_{k,n'}$, $\frac{\hat{\mathbf{g}}_{k,n} \hat{\mathbf{g}}_{k,n'}^H}{\|\hat{\mathbf{g}}_{k,n'}\|}$ is a Gaussian RV with zero mean and variance γ_k which does not depend on $\hat{\mathbf{g}}_{k,n'}$. A Gaussian RV will depend only on its first and second moments, therefore, $\frac{\hat{\mathbf{g}}_{k,n} \hat{\mathbf{g}}_{k,n'}^H}{\|\hat{\mathbf{g}}_{k,n'}\|}$ is a Gaussian RV independent of $\hat{\mathbf{g}}_{k,n'}$. Consequently, (59) can be expressed as

$$\begin{aligned} & \mathbb{E} \left\{ \left[\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right]_{n,n} \right\} \\ &= 1 + \sum_{n' \neq n}^N \mathbb{E} \left\{ \left| \frac{\hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n'}}{\|\hat{\mathbf{g}}_{k,n'}\|} \right|^2 \right\} \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{g}}_{k,n'}\|^2} \right\} \\ &= 1 + \sum_{n' \neq n}^N \gamma_k \frac{1}{(M-1)\gamma_k} = 1 + \frac{N-1}{M-1}, \end{aligned} \quad (60)$$

where the second equality has made use of [35, Lemma 2.10]. Therefore,

$$\mathbb{E} \left\{ \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_{\hat{\mathbf{G}}_k} \mathbf{D}_{\hat{\mathbf{G}}_k}^H \hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \right\} = \left(1 + \frac{N-1}{M-1} \right) \mathbf{I}_N. \quad (61)$$

By substituting (61) into (58), we obtain

$$\begin{aligned} \mathbf{D}_k &= \rho_d q_k^2 \left(1 + \frac{N-1}{M-1} \right) \mathbf{I}_N \\ &\quad + \rho_d q_k^2 \frac{N(\beta_k - \gamma_k)}{(M-1)\gamma_k} \mathbf{I}_N. \end{aligned} \quad (62)$$

The substitution of (57) and (62) into (55) yields

$$\mathbf{\Xi}_k = \rho_d \beta_k \mathbf{I}_N + \rho_d \frac{M-2}{KN} \gamma_k \mathbf{I}_N + \mathbf{I}_N. \quad (63)$$

The desired result (22), can be obtained from the results of (14), (54), and (63).

C. Proof of Lemma 1

For the proof of Lemma 1, (4) requires closed-form equations of both $\mathbb{E} \left\{ \mathbf{g}_{k,n} \check{\mathbf{y}}_{p,k,n}^H \right\}$ and $\mathbb{E} \left\{ \check{\mathbf{y}}_{p,k,n} \check{\mathbf{y}}_{p,k,n}^H \right\}$. We have,

$$\begin{aligned} & \mathbb{E} \left\{ \mathbf{g}_{k,n} \check{\mathbf{y}}_{p,k,n}^H \right\} = \mathbb{E} \left\{ \mathbf{g}_{k,n} (\sqrt{\tau_p \rho_p} \mathbf{g}_{k,n}^H + \check{\mathbf{n}}_p^H) \right\} \\ &= \mathbb{E} \left\{ \sqrt{\tau_p \rho_p} \beta_k \sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*} \sum_{j=1}^{n_k} a_{j,n}^{(k)} \mathbf{b}_j^{(k)H} c_j^{(k)*} \right\} \\ &= \sqrt{\tau_p \rho_p} \beta_k \sum_{j=1}^{n_k} |c_j^{(k)}|^2 \mathbf{I}_M = \sqrt{\tau_p \rho_p} \beta_k \mathbf{I}_M. \end{aligned} \quad (64)$$

where in (64), we have used $\mathbb{E} \left\{ \mathbf{b}_j^{(k)} \mathbf{b}_j^{(k)H} \right\} = \mathbf{I}_M$, $\mathbb{E} \left\{ |a_{j,n}|^2 \right\} = 1$, and the condition (24). Similarly, we have

$$\begin{aligned} \mathbb{E} \left\{ \check{\mathbf{y}}_{p,k,n} \check{\mathbf{y}}_{p,k,n}^H \right\} &= \mathbb{E} \left\{ (\sqrt{\tau_p \rho_p} \mathbf{g}_{k,n} + \check{\mathbf{n}}_p) (\sqrt{\tau_p \rho_p} \mathbf{g}_{k,n}^H + \check{\mathbf{n}}_p^H) \right\} \\ &= \mathbb{E} \left\{ \tau_p \rho_p \mathbf{g}_{k,n} \mathbf{g}_{k,n}^H \right\} + \mathbb{E} \left\{ \sqrt{\tau_p \rho_p} \mathbf{g}_{k,n} \check{\mathbf{n}}_p^H \right\} \\ &\quad + \mathbb{E} \left\{ \sqrt{\tau_p \rho_p} \check{\mathbf{n}}_p \mathbf{g}_{k,n}^H \right\} + \mathbb{E} \left\{ \check{\mathbf{n}}_p \check{\mathbf{n}}_p^H \right\} \\ &= \tau_p \rho_p \beta_k \mathbf{I}_M + \mathbf{I}_M. \end{aligned} \quad (65)$$

Substitute (64) and (65) into (4) will return (18).

D. Proof of Proposition 3

Using (14)–(16), we can see that closed-form derivations for $\bar{\mathbf{H}}_k$ and $\mathbf{\Xi}_k$ are required to achieve the final result. Beginning with $\bar{\mathbf{H}}_k$, we substitute (27) into (15) to obtain

$$\begin{aligned} \bar{\mathbf{H}}_k &= \sqrt{\rho_d} \mathbb{E} \left\{ \mathbf{G}_k^H \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \right\} \\ &= \frac{\sqrt{\rho_d} M}{\sqrt{KMN\gamma_k}} \alpha_k^2 (\tau_p \rho_p \beta_k + 1) \mathbf{I}_N. \end{aligned} \quad (66)$$

This outcome (66) is achieved by substituting $\mathbf{G}_k$ with $\hat{\mathbf{G}}_k + \mathbf{E}_k$, where $\mathbf{E}_k = [\epsilon_{k,1}, \dots, \epsilon_{k,N}]$, and by utilizing the uncorrelated relationship of $\mathbf{G}_k$ and $\mathbf{E}_k$. An expression for $\mathbf{\Xi}_k$ is now derived. From (17) we can obtain

$$\begin{aligned} \mathbf{\Xi}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k'=1}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k \right\} + \mathbf{I}_N \\ &= \mathbf{\Upsilon}_k + \mathbf{\Omega}_k + \mathbf{I}_N, \end{aligned} \quad (67)$$

where

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

and

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

The final term $\mathbf{\Xi}_k$ will now require closed-form expressions for $\mathbf{\Upsilon}_k$ and $\mathbf{\Omega}_k$. Substituting (27) into $\mathbf{\Upsilon}_k$ will produce

$$\begin{aligned} \mathbf{\Upsilon}_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{\hat{\mathbf{G}}_{k'}}{\sqrt{KMN\gamma_{k'}}} \frac{\hat{\mathbf{G}}_{k'}^H}{\sqrt{KMN\gamma_{k'}}} \right) \mathbf{G}_k \right\} \\ &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \sum_{k' \neq k}^K \left(\frac{N \alpha_{k'}^2 (\tau_p \rho_p \beta_{k'} + 1)}{KMN\gamma_{k'}} \right) \mathbf{G}_k \right\} \\ &= \rho_d \sum_{k' \neq k}^K \frac{\alpha_{k'}^2 (\tau_p \rho_p \beta_{k'} + 1)}{K\gamma_{k'}} \beta_k \mathbf{I}_N = \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N. \end{aligned} \quad (68)$$

Similarly,

$$\begin{aligned}\Omega_k &= \rho_d \mathbb{E} \left\{ \mathbf{G}_k^H \frac{\hat{\mathbf{G}}_k}{\sqrt{KMN\gamma_k}} \frac{\hat{\mathbf{G}}_k^H}{\sqrt{KMN\gamma_k}} \mathbf{G}_k \right\} \\ &= \frac{\rho_d}{KMN\gamma_k} \mathbb{E} \left\{ \mathbf{G}_k^H \hat{\mathbf{G}}_k \hat{\mathbf{G}}_k^H \mathbf{G}_k \right\} \\ &= \frac{\rho_d \alpha_k^2}{KMN\gamma_k} (\tau_p \rho_p \mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k \}) \\ &\quad + \frac{\rho_d \alpha_k^2}{KMN\gamma_k} (\mathbb{E} \{ \mathbf{G}_k^H \tilde{\mathbf{N}}_{p,k} \tilde{\mathbf{N}}_{p,k}^H \mathbf{G}_k \}),\end{aligned}\quad (69)$$

where due to the fact that $\mathbf{G}_k$ and $\tilde{\mathbf{N}}_{p,k}$ are uncorrelated, the terms $\mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \tilde{\mathbf{N}}_{p,k}^H \mathbf{G}_k \}$ and $\mathbb{E} \{ \mathbf{G}_k^H \tilde{\mathbf{N}}_{p,k} \mathbf{G}_k^H \mathbf{G}_k \}$ can be removed. Closed-form expressions for $\mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k \}$ and $\mathbb{E} \{ \mathbf{G}_k^H \tilde{\mathbf{N}}_{p,k} \tilde{\mathbf{N}}_{p,k}^H \mathbf{G}_k \}$ must be determined.

- Compute $\mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k \}$:

The off-diagonal elements of $\mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k \}$ can be shown to be zero. However, its n th diagonal element is equal to

$$\mathbb{E} \{ [\mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k]_{n,n} \} = v_{k,n} + \sum_{n' \neq n}^N z_{k,n,n'}, \quad (70)$$

where $v_{k,n} = \mathbb{E} \{ \|\mathbf{g}_{k,n}\|^4 \}$ and $z_{k,n,n'} = \mathbb{E} \{ |\mathbf{g}_{k,n}^H \mathbf{g}_{k,n'}|^2 \}$.

Using (25) we can determine $v_{k,n}$:

$$\begin{aligned}v_{k,n} &= \mathbb{E} \{ \|\mathbf{g}_{k,n}\|^4 \} \\ &= \beta_k^2 \mathbb{E} \left\{ \left| \frac{\sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*}}{\sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2}} \right|^4 \right\}.\end{aligned}\quad (71)$$

Conditioned on $\{a_{j,n}^{(k)}\}$, $\frac{\sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*}}{\sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2}}$ is a Gaussian vector

with zero mean and covariance matrix $\mathbf{I}_M$; $\frac{\sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*}}{\sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2}}$

is a Gaussian vector which is independent of $\{a_{j,n}^{(k)}\}$, this is due to the covariance matrix's independence of $\{a_{j,n}^{(k)}\}$ and since a Gaussian RV will depend on only the first and second moments. As a result,

$$\begin{aligned}v_{k,n} &= \beta_k^2 \mathbb{E} \left\{ \left| \frac{\sum_{j=1}^{n_k} c_j^{(k)} \mathbf{b}_j^{(k)} a_{j,n}^{(k)*}}{\sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2}} \right|^4 \right\} \\ &\quad \times \mathbb{E} \left\{ \left| \sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2} \right|^4 \right\} \\ &= \beta_k^2 M (M+1) \mathbb{E} \left\{ \left| \sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2} \right|^4 \right\},\end{aligned}\quad (72)$$

where [35, Lemma 2.9] is used to derive the last equality. The expectation in (72) can be simplified as follows,

$$\begin{aligned}\mathbb{E} \left\{ \left| \sqrt{\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2} \right|^4 \right\} &= \mathbb{E} \left\{ \left(\sum_{j=1}^{n_k} |c_j^{(k)} a_{j,n}^{(k)*}|^2 \right)^2 \right\} \\ &= \sum_{j=1}^{n_k} \mathbb{E} \left\{ |c_j^{(k)} a_{j,n}^{(k)*}|^4 \right\} \\ &\quad + \sum_{j=1}^{n_k} \sum_{j \neq i}^{n_k} \mathbb{E} \left\{ |c_j^{(k)} a_{j,n}^{(k)*}|^2 \right\} \mathbb{E} \left\{ |c_i^{(k)} a_{i,n}^{(k)*}|^2 \right\} \\ &= 2 \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + \sum_{j=1}^{n_k} \sum_{j \neq i}^{n_k} |c_j^{(k)}|^2 |c_i^{(k)}|^2 \\ &= \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1.\end{aligned}\quad (73)$$

Substituting (73) into (72), we obtain

$$v_{k,n} = \beta_k^2 M (M+1) \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right). \quad (74)$$

Now, we compute $z_{k,n,n'}$. From (25), for $n' \neq n$, we have

$$\begin{aligned}z_{k,n,n'} &= \mathbb{E} \{ |\mathbf{g}_{k,n}^H \mathbf{g}_{k,n'}|^2 \} \\ &= \beta_k^2 \mathbb{E} \left\{ |(\mathbf{C}_k \mathbf{a}_{k,n})^H (\mathbf{C}_k \mathbf{a}_{k,n'})|^2 \right\} \\ &= \beta_k^2 \text{Tr} (\mathbb{E} \{ \mathbf{C}_k^H \mathbf{C}_k \mathbf{C}_k^H \mathbf{C}_k \}) \\ &= \beta_k^2 \text{Tr} \left(\mathbb{E} \left\{ \sum_{j=1}^{n_k} |c_j^{(k)} \mathbf{b}_j^{(k)}|^2 \right\} \right) \\ &= \beta_k^2 \left(M^2 \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + M \right),\end{aligned}\quad (75)$$

where $\mathbf{C}_k = [c_1^{(k)} \mathbf{b}_1^{(k)}, c_2^{(k)} \mathbf{b}_2^{(k)}, \dots, c_{n_k}^{(k)} \mathbf{b}_{n_k}^{(k)}]$ and $\mathbf{a}_{k,n} = [a_{1,n}^{(k)*}, a_{2,n}^{(k)*}, \dots, a_{n_k,n}^{(k)*}]^T$. Substitute (74) and (75) into (70),

$$\begin{aligned}\mathbb{E} \{ [\mathbf{G}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{G}_k]_{n,n} \} &= \beta_k^2 M (M+1) \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \\ &\quad + \beta_k^2 M (N-1) \left(M \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right).\end{aligned}\quad (76)$$

- Compute $\mathbb{E} \{ \mathbf{G}_k^H \tilde{\mathbf{N}}_k \tilde{\mathbf{N}}_k^H \mathbf{G}_k \}$:

We have,

$$\begin{aligned}\mathbb{E} \{ \mathbf{G}_k^H \tilde{\mathbf{N}}_k \tilde{\mathbf{N}}_k^H \mathbf{G}_k \} &= N \mathbb{E} \{ \mathbf{G}_k^H \mathbf{G}_k \} \\ &= N \begin{bmatrix} \mathbb{E} \{ \mathbf{g}_{k,1}^H \mathbf{g}_{k,1} \} & \dots & 0 \\ \vdots & \mathbb{E} \{ \mathbf{g}_{k,2}^H \mathbf{g}_{k,2} \} & \\ 0 & \ddots & \mathbb{E} \{ \hat{\mathbf{g}}_{k,n}^H \hat{\mathbf{g}}_{k,n} \} \end{bmatrix} \\ &= MN \beta_k \mathbf{I}_N.\end{aligned}\quad (77)$$

By substituting (76) and (77) into (69) the following result is returned

$$\begin{aligned}\mathbf{\Omega}_k &= \frac{\tau_p \rho_d \rho_p \alpha_k^2 \beta_k^2}{KN \gamma_k} \left((M+1) \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \right) \mathbf{I}_N \\ &+ \frac{\tau_p \rho_d \rho_p \alpha_k^2 \beta_k^2}{KN \gamma_k} \left((N-1) \left(M \sum_{j=1}^{n_k} \|c_j^{(k)}\|^4 + 1 \right) \right) \mathbf{I}_N \\ &+ \frac{\rho_d \alpha_k^2 \beta_k}{K \gamma_k} \mathbf{I}_N.\end{aligned}\quad (78)$$

Substituting (68) and (78) into (67), we obtain

$$\begin{aligned}\mathbf{\Xi}_k &= \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N \\ &+ \frac{\tau_p \rho_d \rho_p \alpha_k^2 \beta_k^2}{KN \gamma_k} \left((M+1) \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \right) \mathbf{I}_N \\ &+ \frac{\tau_p \rho_d \rho_p \alpha_k^2 \beta_k^2}{KN \gamma_k} \left((N-1) \left(M \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \right) \mathbf{I}_N \\ &+ \frac{\rho_d \alpha_k^2 \beta_k}{K \gamma_k} \mathbf{I}_N + \mathbf{I}_N.\end{aligned}\quad (79)$$

The substitution of (66) and (79) into (14) yields Proposition 3.

E. Derivation of Equation (33)

As stated, we make use of the Taylor series expansion to achieve an approximation for L_k

$$\begin{aligned}L_k &\approx \frac{1}{\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^2\}} + \frac{1}{(\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^2\})^3} \text{Var}\{\|\hat{\mathbf{g}}_{k,n}\|^2\} \\ &= \frac{1}{M \gamma_k} + \frac{1}{M^3 \gamma_k^3} (\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^4\} - M^2 \gamma_k^2).\end{aligned}\quad (80)$$

A closed-form derivation of $\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^4\}$ is found below,

$$\begin{aligned}\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^4\} &= \mathbb{E}\left\{\left\|\sqrt{\tau_p \rho_p} \alpha_k \mathbf{g}_{k,n} + \alpha_k \tilde{\mathbf{n}}_{p,k,n}\right\|^4\right\} \\ &= \tau_p^2 \rho_p^2 \alpha_k^4 \mathbb{E}\{\|\mathbf{g}_{k,n}\|^4\} + \alpha_k^4 \mathbb{E}\{\|\tilde{\mathbf{n}}_{p,k,n}\|^4\} \\ &\quad + 2\tau_p \rho_p \alpha_k^4 \mathbb{E}\{\|\mathbf{g}_{k,n}\|^2 \|\tilde{\mathbf{n}}_{p,k,n}\|^2\} \\ &\quad + 2\tau_p \rho_p \alpha_k^4 \mathbb{E}\{|\mathbf{g}_{k,n}^H \tilde{\mathbf{n}}_{p,k,n}|^2\} \\ &= \tau_p^2 \rho_p^2 \alpha_k^4 \beta_k^2 M (M+1) \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 + 1 \right) \\ &\quad + \alpha_k^4 M (M+1) + 2\tau_p \rho_p \alpha_k^4 M^2 \beta_k \\ &\quad + 2\tau_p \rho_p \alpha_k^4 M \beta_k,\end{aligned}\quad (81)$$

where the first term of the final result makes use of (74). By using $\alpha_k = \frac{\gamma_k}{\sqrt{\tau_p \rho_p \beta_k}}$, we obtain

$$\mathbb{E}\{\|\hat{\mathbf{g}}_{k,n}\|^4\} = M (M+1) \frac{\gamma_k^4}{\beta_k^2} \sum_{j=1}^{n_k} |c_j^{(k)}|^4 + M (M+1) \gamma_k^2.\quad (82)$$

Substituting (82) into (80) will produce the result seen in (33).

F. Proof of Proposition 4

As with the previous propositions derivations, Proposition 4 requires closed-form expressions for $\bar{\mathbf{H}}_k$ and $\mathbf{\Xi}_k$. We have

$$\begin{aligned}\bar{\mathbf{H}}_k &= \sqrt{\rho_d} \mathbb{E}\{\mathbf{G}_k^H \mathbf{W}_k\} \\ &= \sqrt{\rho_d} q_k \mathbb{E}\{\hat{\mathbf{G}}_k^H \hat{\mathbf{G}}_k \mathbf{D}_k\} \\ &= \sqrt{\rho_d} q_k \mathbf{I}_N.\end{aligned}\quad (83)$$

Compute $\mathbf{\Xi}_k$ according to

$$\begin{aligned}\mathbf{\Xi}_k &= \rho_d \mathbb{E}\left\{\mathbf{G}_k^H \sum_{k'=1}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k\right\} + \mathbf{I}_N \\ &= \rho_d \mathbb{E}\left\{\mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k\right\} \\ &\quad + \rho_d \mathbb{E}\{\mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k\} + \mathbf{I}_N.\end{aligned}\quad (84)$$

Let us set $\mathbf{R}_k = \rho_d \mathbb{E}\{\mathbf{G}_k^H \sum_{k' \neq k}^K (\mathbf{W}_{k'} \mathbf{W}_{k'}^H) \mathbf{G}_k\}$ and $\mathbf{S}_k = \rho_d \mathbb{E}\{\mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k\}$. Now have

$$\begin{aligned}\mathbf{R}_k &= \rho_d \sum_{k' \neq k}^K \mathbb{E}\{\mathbf{G}_k^H \mathbf{W}_{k'} \mathbf{W}_{k'}^H \mathbf{G}_k\} \\ &= \rho_d \frac{(K-1)}{K} \beta_k \mathbf{I}_N.\end{aligned}\quad (85)$$

Here, (85) is obtained since the matrix $\mathbb{E}\{\mathbf{G}_k^H \mathbf{W}_{k'} \mathbf{W}_{k'}^H \mathbf{G}_k\}$ is diagonal, its diagonal elements can be reduced to $\frac{\beta_k}{K}$ using the power constraint $\mathbb{E}\{\text{tr}(\mathbf{W}_{k'} \mathbf{W}_{k'}^H)\} = \frac{1}{K}$. Next, we compute $\mathbf{S}_k$. We can easily show that $\mathbf{S}_k$ is a diagonal matrix whose n -th diagonal element is

$$\begin{aligned}[\mathbf{S}_k]_{n,n} &= \rho_d \mathbb{E}\{\mathbf{g}_{k,n}^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{g}_{k,n}\} \\ &= \rho_d \mathbb{E}\left\{\mathbf{g}_{k,n}^H q_k^2 \left(\sum_{m=1}^N \frac{\hat{\mathbf{g}}_{k,m} \hat{\mathbf{g}}_{k,m}^H}{\|\hat{\mathbf{g}}_{k,m}\|^4} \right) \mathbf{g}_{k,n}\right\} \\ &= \rho_d q_k^2 \mathbb{E}\left\{\mathbf{g}_{k,n}^H \frac{\hat{\mathbf{g}}_{k,n} \hat{\mathbf{g}}_{k,n}^H}{\|\hat{\mathbf{g}}_{k,n}\|^4} \mathbf{g}_{k,n}\right\} \\ &\quad + \rho_d q_k^2 \sum_{m \neq n}^N \mathbb{E}\left\{\mathbf{g}_{k,n}^H \frac{\hat{\mathbf{g}}_{k,m} \hat{\mathbf{g}}_{k,m}^H}{\|\hat{\mathbf{g}}_{k,m}\|^4} \mathbf{g}_{k,n}\right\} \\ &\approx \rho_d q_k^2 + \rho_d q_k^2 \sum_{m \neq n}^N \mathbb{E}\left\{\text{tr}\left(\mathbf{g}_{k,n} \mathbf{g}_{k,n}^H \frac{\hat{\mathbf{g}}_{k,m} \hat{\mathbf{g}}_{k,m}^H}{\|\hat{\mathbf{g}}_{k,m}\|^4}\right)\right\} \\ &\approx \rho_d q_k^2 + \rho_d q_k^2 \sum_{m \neq n}^N \beta_k M \sum_{j=1}^{n_k} |c_j^{(k)}|^4 \mathbb{E}\left\{\frac{1}{\|\hat{\mathbf{g}}_{k,m}\|^2}\right\} \\ &= \rho_d q_k^2 + \rho_d q_k^2 \sum_{m \neq n}^N \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) L_k \\ &= \rho_d q_k^2 + \rho_d \frac{N-1}{KN} \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right),\end{aligned}\quad (86)$$

where the last equality is obtained from (31). Therefore,

$$\mathbf{S}_k = \left(\rho_d q_k^2 + \rho_d \frac{N-1}{KN} \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) \right) \mathbf{I}_N.\quad (87)$$

Substituting (85) and (87) into (84), we obtain

$$\Xi_k = \rho_d \frac{K-1}{K} \beta_k \mathbf{I}_N + \rho_d q_k^2 \mathbf{I}_N + \rho_d \frac{N-1}{KN} \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) \mathbf{I}_N + \mathbf{I}_N. \quad (88)$$

The substitution of (83) and (88) into (14) yields

$$R_k^{\text{K,prop}} \approx \frac{T - \tau_p}{T} N \times \log_2 \left(1 + \frac{\rho_d / (NKL_k)}{\rho_d \frac{K-1}{K} \beta_k + \rho_d \frac{N-1}{KN} \beta_k M \left(\sum_{j=1}^{n_k} |c_j^{(k)}|^4 \right) + 1} \right). \quad (89)$$

From (33) and (89) we arrive at the desired result in Proposition 4.

REFERENCES

- [1] J. A. Sutton, H. Q. Ngo, and M. Matthaiou, "Performance of a novel maximum-ratio precoder in massive MIMO with multiple-antenna users," in *Proc. IEEE PIMRC*, Sept. 2019.
- [2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, UK: Cambridge University Press, 2016.
- [3] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.
- [4] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [6] 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.
- [7] N. Akbar, S. Yan, N. Yang, and J. Yuan, "Location-aware pilot allocation in multicell multiuser massive MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7774–7778, Aug. 2018.
- [8] S.-N. Jin, D.-W. Yue, and H. H. Nguyen, "Multicell massive MIMO: Downlink rate analysis with linear processing under Ricean fading," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3777–3791, Apr. 2019.
- [9] F. Rezaei and A. Tadaion, "Multi-layer beamforming in uplink/downlink massive MIMO systems with multi-antenna users," *Signal Processing*, vol. 164, pp. 58–66, Nov. 2019.
- [10] 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.
- [11] X. Li, E. Björnson, S. Zhou, and J. Wang, "Massive MIMO with multi-antenna users: When are additional user antennas beneficial?" in *Proc. IEEE ICT*, May 2016.
- [12] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.
- [13] R. Chopra, C. R. Murthy, H. A. Suraweera, and E. G. Larsson, "Blind channel estimation for downlink massive MIMO systems with imperfect channel reciprocity," *IEEE Trans. Signal Process.*, vol. 68, pp. 3132–3145, Apr. 2020.
- [14] J. Qiu, K. Xu, X. Xia, Z. Shen, and W. Xie, "Downlink power optimization for cell-free massive MIMO over spatially correlated Rayleigh fading channels," *IEEE Access*, vol. 8, pp. 56214–56227, Mar. 2020.
- [15] 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. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
- [16] C. Zhong, S. Jin, K.-K. Wong, and M. R. McKay, "Ergodic mutual information analysis for multi-keyhole MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1754–1763, Jun. 2011.
- [17] G. Levin and S. Loyka, "From multi-keyholes to measure of correlation and power imbalance in MIMO channels: Outage capacity analysis," *IEEE Trans. Inf. theory*, vol. 57, no. 6, pp. 3515–3529, Jun. 2011.
- [18] C. Zhong, S. Jin, K.-K. Wong, and M. R. McKay, "Outage analysis for optimal beamforming MIMO systems in multikeyhole channels," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1451–1456, Mar. 2010.
- [19] Q. Wang, D. Dong, and T. Zhang, "Secure degrees of freedom of rank-deficient MIMO interference channel," *IEEE Commun. Lett.*, vol. 23, no. 10, pp. 1740–1743, Oct. 2019.
- [20] T. C. Mai, H. Q. Ngo, and T. Q. Duong, "Cell-free massive MIMO systems with multi-antenna users," in *Proc. IEEE (GlobalSIP)*, Nov. 2018, pp. 828–832.
- [21] Y. Zhou and T. Ng, "Performance analysis on MIMO-OFCDM systems with multi-code transmission," *IEEE Wireless Commun.*, vol. 8, pp. 4426–4433, Sep. 2009.
- [22] Q. H. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [23] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [24] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Mar. 2018.
- [25] X. Zhao, J. Zhang, J. Zhang, F. Xiong, and B. Ai, "Efficient receiver for cell-free massive MIMO systems with low-resolution ADCs," in *Proc. IEEE ICC*, Jun. 2020, pp. 1–6.
- [26] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4849–4864, Dec. 2015.
- [27] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "On the performance of cell-free massive MIMO with short-term power constraints," in *Proc. IEEE CAMAD*, Dec. 2016, pp. 225–230.
- [28] X. Li, S. Jin, X. Gao, and M. R. McKay, "Capacity bounds and low complexity transceiver design for double-scattering MIMO multiple access channels," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2809–2822, May 2010.
- [29] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: 10 myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
- [30] Q.-U.-A. Nadeem, A. Kammoun, M. Debbah, and M. S. Alouini, "Asymptotic analysis of RZF over double scattering channels with MMSE estimation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2509–2526, May 2019.
- [31] B. Sun, Y. Zhou, J. Yuan, and J. Shi, "Interference cancellation based channel estimation for massive MIMO systems with time shifted pilots," *IEEE Trans. Commun.*, vol. 19, no. 10, pp. 6826–6843, Oct. 2020.
- [32] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Wireless Commun.*, vol. 14, pp. 4353–4368, Aug. 2015.
- [33] B. Sun, Y. Zhou, J. Yuan, Y. F. Liu, L. Wang, and L. Liu, "High order PSK modulation in massive MIMO systems with 1-bit ADCs," *IEEE Trans. Commun.*, pp. 1–1, Oct. 2020.
- [34] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Wireless Commun.*, vol. 18, pp. 4157–4170, Aug. 2019.
- [35] 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.

James A. C. Sutton received his MEng. in electrical and electronic engineering from Queen's University Belfast in 2018. He is in the process of completing his Ph.D within the Institute of Electronics, Communications and Information Technology (ECIT) at Queen's University Belfast. His research focuses on massive MIMO and cell-free massive MIMO, with particular emphasis on multiple-antenna users and precoding techniques.

Hien Quoc Ngo is currently a Lecturer with Queen's University Belfast, U.K. He has coauthored many research articles in wireless communications and coauthored the Cambridge University Press textbook *Fundamentals of Massive MIMO* (2016). His main research interests include massive (large-scale) MIMO systems, cell-free massive MIMO, physical layer security, and cooperative communications. He received the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015, the IEEE ComSoc Leonard G. Abraham Prize in 2017, and the Best

Ph.D. Award from EURASIP in 2018. He also received the IEEE Sweden VT-COM-IT Joint Chapter Best Student Journal Paper Award in 2015. He was awarded the UKRI Future Leaders Fellowship in 2019. He serves as the Editor for the IEEE Transactions on Wireless Communications, IEEE Wireless Communications Letters, the Digital Signal Processing, and the Physical Communication (Elsevier). He was a Guest Editor of IET Communications, and a Guest Editor of IEEE ACCESS in 2017.

Michail Matthaiou (S'05–M'08–SM'13) was born in Thessaloniki, Greece in 1981. He obtained the Diploma degree (5 years) in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece in 2004. He then received the M.Sc. (with distinction) in Communication Systems and Signal Processing from the University of Bristol, U.K. and Ph.D. degrees from the University of Edinburgh, U.K. in 2005 and 2008, respectively. From September 2008 through May 2010, he was with the Institute for Circuit Theory and Signal

Processing, Munich University of Technology (TUM), Germany working as a Postdoctoral Research Associate. He is currently a Professor of Communications Engineering and Signal Processing and Deputy Director of the Centre for Wireless Innovation (CWI) at Queens University Belfast, U.K. after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, beyond massive MIMO, intelligent reflecting surfaces, mm-wave/THz systems and deep learning for communications.

Dr. Matthaiou and his coauthors received the IEEE Communications Society (ComSoc) Leonard G. Abraham Prize in 2017. He currently holds the ERC Consolidator Grant BEATRICE (2021-2026) focused on the interface between information and electromagnetic theories. He was awarded the prestigious 2018/2019 Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship and also received the 2019 EURASIP Early Career Award. His team was also the Grand Winner of the 2019 Mobile World Congress Challenge. He was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He has co-authored papers that received best paper awards at the 2018 IEEE WCSP and 2014 IEEE ICC and was an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS for 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He is currently the Editor-in-Chief of Elsevier Physical Communication, a Senior Editor for IEEE WIRELESS COMMUNICATIONS LETTERS and an Associate Editor for the IEEE JSAC SERIES ON MACHINE LEARNING FOR COMMUNICATIONS AND NETWORKS.