Asymptotic Sum Rate Analysis Over Double Scattering Channels With MMSE Estimation and MRT Precoding

Thesis by
Jia Ye

In Partial Fulfillment of the Requirements
For the Degree of
Masters of Science

King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia

March, 2020
EXAMINATION COMMITTEE PAGE

The thesis of Jia Ye is approved by the examination committee

Committee Chairperson: Prof. Mohamed-Slim Alouini
Committee Members: Prof. Boon S. Ooi, Dr. Abla Kammoun
ABSTRACT

Asymptotic Sum Rate Analysis Over Double Scattering Channels With MMSE Estimation and MRT Precoding
Jia Ye

This thesis investigates the performance of a multi-user multiple-input single-output (MISO) system considering maximum ratio transmission (MRT) downlink precoding. The transmitted signal from the base station (BS) to each user is assumed to experience the double scattering channel. We adopt the minimum-mean-square-error (MMSE) channel estimator for the proposed model. Within this setting, we are interested in deriving tight approximations of the ergodic rate assuming the number of BS antennas, users, and scatterers grow large with the same pace. Under the special multi-keyhole channels, these deterministic equivalents are expressed in more simplified closed-form expressions. The simplified expressions reveal that unlike the standard Rayleigh channel in which the SINR grows as \( \mathcal{O}\left(\frac{N}{K}\right) \), the SINR associated with a multi-keyhole channel scales as \( \mathcal{O}\left(\frac{S}{K}\right) \). This particularly shows that the reaped gains of the large-scale MIMO over double scattering channels do not linearly increase with the number of antennas and are limited by the number of scatterers. We further show that the derived asymptotic results match the simulation results closely under moderate system dimensions and provide some useful insights into the interplay between \( N, K \) and \( S \).
ACKNOWLEDGEMENTS

I adore and cherish my two years master journey at KAUST very much, which has been fulled by various experience.

I would like to express my heartfelt appreciation to my graduate study supervisor: Prof. Mohamed-Slim Alouini for his approval of my Ms/PhD application at KAUST. I have grown a lot in learning and quality under his unreserved guidance in right research direction and useful courses. The rewarding study experience at KAUST would not happen without his guidance, encouragement, and unconditional support.

The thesis presented is a result of the incredible supervision and guidance I received from Dr. Abla Kammoun, a research scientist at KAUST. She taught me all the useful tools and answered all questions I have. I am very appreciate her unreserved instruction and help and learnt a lot from her thinking style and dedication in research.

I also would like to give special thanks to Dr. Qurrat-Ul-Ain Nadeem, Prof. Ahmed Sultan Salem, Dr. Gaofeng Pan, Prof. Shuaishuai Guo, and my research group members in the Communication Theory Lab for their help both in research and courses.

Finally, I would like to express my profound appreciation to my family members.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination Committee Page</td>
<td>2</td>
</tr>
<tr>
<td>Copyright</td>
<td>3</td>
</tr>
<tr>
<td>Abstract</td>
<td>4</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>5</td>
</tr>
<tr>
<td>List of Figures</td>
<td>8</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>9</td>
</tr>
<tr>
<td>1.1 LS-MIMO Technologies</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Double Scattering Channel</td>
<td>10</td>
</tr>
<tr>
<td>1.3 MRT Precoding Scheme</td>
<td>12</td>
</tr>
<tr>
<td>1.4 Motivation and Main Contributions</td>
<td>13</td>
</tr>
<tr>
<td>1.5 Outline and Notation</td>
<td>13</td>
</tr>
<tr>
<td><strong>2 System Model</strong></td>
<td>15</td>
</tr>
<tr>
<td>2.1 Channel Estimation</td>
<td>17</td>
</tr>
<tr>
<td>2.2 Precoding Scheme and Achievable Rate</td>
<td>18</td>
</tr>
<tr>
<td><strong>3 Main Results</strong></td>
<td>21</td>
</tr>
<tr>
<td>3.1 Assumptions</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Useful Results</td>
<td>23</td>
</tr>
<tr>
<td>3.3 Deterministic Approximation of the SINR</td>
<td>24</td>
</tr>
<tr>
<td>3.4 Rayleigh Product Channel</td>
<td>26</td>
</tr>
<tr>
<td>3.4.1 Standard Rayleigh Channel</td>
<td>27</td>
</tr>
<tr>
<td>3.4.2 Multi-keyhole Channel with Limited Scatterers</td>
<td>29</td>
</tr>
<tr>
<td><strong>4 Simulation Results</strong></td>
<td>31</td>
</tr>
<tr>
<td><strong>5 Concluding Remarks</strong></td>
<td>36</td>
</tr>
<tr>
<td>5.1 Summary</td>
<td>36</td>
</tr>
</tbody>
</table>
5.2 Future Research Work .................................................. 37
References ..................................................................... 43
Appendices .................................................................... 47
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The geometric model of the double scattering channel between the BS and the $k$-th user in group $g$</td>
<td>16</td>
</tr>
<tr>
<td>4.1</td>
<td>Sum rate versus SNR.</td>
<td>32</td>
</tr>
<tr>
<td>4.2</td>
<td>Sum rate versus SNR for a multi-keyhole channel.</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Sum rate versus $N$ with $K = 480$.</td>
<td>34</td>
</tr>
<tr>
<td>4.4</td>
<td>Sum rate versus $N$ with $K = 40$.</td>
<td>35</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In this thesis, the asymptotic ergodic rate of maximum ratio transmission (MRT) precoding in Large-scale multiple-input multiple-output (LS-MIMO) is studied. Double scattering channel is assumed in this thesis to model the channel fading between the base station (BS) and all users.

1.1 LS-MIMO Technologies

With the evolution of networks and invention of smart devices, the capacity of wireless communications networks as well as the Quality of Service (QoS) requirements of user are increasing exponentially. LS-MIMO has emerged as a key technology to significantly increase the spectral efficiency of wireless communication systems and ensure massive connectivity [1]. LS-MIMO is also known as massive MIMO scheme, large-scale antenna systems or very large multi-user-MIMO etc. is proposed by Marzetta in [2].

Specifically, LS-MIMO achieves large multiplexing gains by using hundreds of antennas to serve tens of users [3] simultaneously. The use of more antennas also helps focus energy into smaller regions thereby mitigating the inter-user interference [1, 4, 5]. The authors in [1] also showed that individual element failure of the antenna array does not play a key role on the performance of an LS-MIMO network. The excessive Degrees of Freedom in LS-MIMO systems can enhance transmission reliability [6]. Several works have shown that under full rank Rayleigh channels, LS-MIMO systems
provide large capacity gains that linearly scale with the number of antennas \[7\].

Some experiments also have been conducted to verify the feasibility of LS-MIMO technologies. The authors in \[8\] has tested that a LS-MIMO system equipped with $12 \times 12$ antennas can achieve the spectral efficiency of 50 bps/Hz with the transmission rate of 4.92 Gbps over a 100 MHz channel. The authors in \[9\] also conducted a LS-MIMO trial in an outdoor test field located in China, which investigated the feasibility and performance of LS-MIMO systems using different precoding schemes. Some LS-MIMO trials have been undertaken at the BT Labs in Adatastral Park, Suffolk \[10\]. The research team showed that LS-MIMO could offer spectrum efficiency achieving or even beyond 100 bits/s/Hz and gaining 10 times higher capacity than long term evolution systems.

1.2 Double Scattering Channel

In practice, it has been observed that full-rank Rayleigh channels represent a too optimistic model that ignores the presence of poor scattering conditions and spatial correlation. These two factors were shown in \[11\] to govern the performance of MIMO systems which is at a much lower level than the performances predicted over full-rank Rayleigh channels \[12, 13\]. In light of this observation, Gesbert et al. propose a new channel model coined ”double scattering channel model” which accounts for the channel rank deficiency, the spatial correlation as well as the limited scattering condition. Unlike the standard Rayleigh channel, the double scattering channel is product of Rayleigh distribution, which is non-Gaussian, making its analysis rather challenging.

The works about the analysis of the performance of MIMO systems over double scattering channels is relatively scarce and is mainly represented by \[14, 15, 16, 17, 18, 19, 20, 21, 22, 23\]. Specifically, the performance of point-to-point MIMO system under double scattering channels has been studied in \[14, 15, 16, 17, 18\], where the
authors in [14] showed that the required number of scatterers in the channel to achieve full diversity is the product of the number of transmit (Tx) antennas and receive (Rx) antennas. In [15], the optimal capacity-achieving beamforming directions under double scattering channels were studied in [15] and were proven to correspond to the eigenvectors of the transmit spatial correlation matrix, while a tight upper bound on the ergodic capacity was provided [16]. Interestingly, the results reveal that the keyhole channels with the application of multiple antennas do not allow for spatial multiplexity gains, but only for diversity gains. In this line, closed-form expressions for the diversity-multiplexing trade-off of double scattering MIMO channels were obtained in [17], while the diversity gain and array gain associated with the symbol error rate (SER) were studied in [18].

It was shown that the array gain of double scattering channel varies with the signal-to-noise ratio (SNR) rather than being a constant as in conventional Rayleigh and Rician channels. Closed-form upper bounds on the sum capacity of a MIMO multiple access system under double scattering channel were provided in [24].

Only few papers have investigated thus far the LS-MIMO models under double scattering fading [19, 20, 21, 22, 23]. The authors in [19] studied large antenna arrays assuming minimum-mean-square-error (MMSE) detector without considering the correlation between antennas of Tx and Rx sides. They derived implicit expressions for the signal-to-interference-plus-noise ratio (SINR) and the channel capacity. The authors in [20] derived deterministic approximation of the mutual information in a MIMO multiple-access system under double scattering fading by leveraging tools from random matrix theory while the spectral efficiency of a multi-cell MIMO system under double scattering channels was assessed through Monte Carlo simulations in [21]. Recently, the performance of LS multi-user MISO systems with regularized zero-forcing (RZF) precoder under double scattering channels has been studied in [22] where tight deterministic approximations of SINR and ergodic rate have been derived.
by taking advantage of the results from random matrix theory. The analysis in \cite{22} has allowed to shed light on the limitations of double scattering channels, however, the results are extremely involved and could not be interpreted easily.

1.3 MRT Precoding Scheme

In massive MIMO systems, base stations (BS) tend to adopt appropriate precoding schemes to ensure reliable transmission. By utilizing the channel state information (CSI), precoding is a significant signal processing procedure to maximize system performance \cite{25}. The most well-known linear precoders include MRT \cite{26}, zero-forcing (ZF) \cite{27} and RZF \cite{28}, which are simple and efficient methods to reduce the complexity of the MIMO receiver.

In contrast to ZF precoding, which attempts to eliminate inter-user interference with a significant received energy loss of the desired signal \cite{29}, the MRT is devised to maximize the signal gain at the user of interest and is optimal in the noise-limited low SNR regime. It has been also shown to not only present low computational complexity but also to be asymptotically optimal when the number of antennas is much larger than that of users \cite{30}, making it particularly suitable for LS-MIMO systems.

Although RZF precoding combines the advantages of MRT and ZF by introducing a factor controlling the amount of inter-user interference that is allowed in the cell \cite{31}, the RZF precoder is recognized for its prohibitively high computational complexity since it requires the inversion of a large Gram matrix of the joint channels of all active users. MRT precoding has important advantages of robustness, low computational complexity, and high asymptotic performance \cite{32}. In fact MRT has been shown to outperform RZF precoding when the number of base station (BS) antennas is much greater than the number of terminal users \cite{30}.

This motivates our work in which we aim to study the MRT precoding.
1.4 Motivation and Main Contributions

Motivated by the above works, we study the downlink SINR and sum rate in a single-cell multi-user MISO system under the double scattering channel with $S$ scatterers between the BS and the users. For simplicity, we assume that the users are divided into $G$ groups with common correlation matrices for users in the same group. We consider a time-division duplex (TDD) protocol where the BS equipped with $N$ Tx antennas estimates the downlink using the MMSE estimation technique based on uplink pilot signaling and employs MRT precoding in the downlink link to serve $K$ users.

Under this setting, we derive tight deterministic approximation of the SINR and ergodic rate relating to the increasing number of BS antennas $N$, scatterers $S$ and users $K$. Interestingly, we observe through numerical results that the obtained results are accurate for system under moderate dimensions. Moreover, we simplify the obtained expressions in the case of a multi-keyhole channel, which provides interesting insights into the impact of the number of scatterers on the system performance. Particularly, we show that the SINR with MRT precoding over a double scattering channel does not grow unboundedly with the number of antennas and scales as $\frac{S}{K}$ in contrast to that over Rayleigh fading channels which grows as $\frac{N}{K}$. This is to be compared with the result in [22] showing that the SINR with RZF precoding grows without bounds with the number of antennas if the number of scatterers is larger than that of users.

1.5 Outline and Notation

The remainder of the thesis is organized as follows. Chapter 2 describes the system model. In Chapter 3, the asymptotically tight approximations of the SINR with MRT precoding is derived firstly, then the system sum rate is derived accordingly. Chapter
Chapter 4 presents the simulation results, and Chapter 5 draws some conclusions.

In this thesis, $x$, $\mathbf{x}$, and $\mathbf{X}$ denote a scalar, a vector, and a matrix respectively. $(\cdot)^H$ is the conjugate transpose. $\text{diag}(\mathbf{x})$ denotes a diagonal matrix whose diagonal entries are from vector $\mathbf{x}$. $\text{tr}(\mathbf{X})$ denotes the trace of $\mathbf{X}$. $\mathbb{E}[\cdot]$ represents the expectation operation. $\|\mathbf{X}\|$ is the spectral norm of $\mathbf{X}$. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. $\mathcal{CN}(\mu, \Sigma)$ stands for the circularly symmetric complex Gaussian distribution with mean $\mu$ and covariance $\Sigma$. Big O notation represented as $a = \mathcal{O}(b)$ serves as a flexible abbreviation for $|a| \leq \beta b$, where $\beta$ is a generic constant. The symbol $\xrightarrow{\text{a.s.}} 0$ represents almost sure convergence.
Chapter 2

System Model

In this thesis, a single-cell multi-user MISO system was considered, where BS is assumed to be equipped with $N$ antennas to serve $K$ single-antenna users. The users are divided into $G$ groups with $K_g$, $g = 1, \cdots, G$ users in $g$-th group. The users located in the same group are assumed to experience similar propagation environment. The received complex baseband signal at $k$-th user in group $g$ is expressed as,

$$y_{k,g} = h_{k,g}^H x + n_{k,g},$$

(2.1)

where $x \in \mathbb{C}^{N \times 1}$ represents the Tx signal vector, $h_{k,g}^H \in \mathbb{C}^{1 \times N}$ represents the double scattering channel vector [22] from the BS to $k$-th user in group $g$, and $n_{k,g} \sim \mathcal{CN}(0, \sigma^2)$ is the additive Gaussian noise. The signal vector $x$ transmitted from Tx is shown as

$$x = \sum_{g=1}^{G} \sum_{k=1}^{K_g} \sqrt{p_{k,g}} g_{k,g} s_{k,g},$$

(2.2)

where $g_{k,g} \in \mathbb{C}^{N \times 1}$ represents the precoding vector, $p_{k,g} \geq 0$ represents the signal power, and $s_{k,g} \sim \mathcal{CN}(0, 1)$ is the data symbol for $k$-th user in group $g$ respectively. Moreover, the precoding vectors satisfy the following total power constraints:

$$\mathbb{E} \left[ \|x\|^2 \right] = \mathbb{E} \left[ \text{tr} \left( P G G^H \right) \right] \leq \bar{P},$$

(2.3)
Figure 2.1: The geometric model of the double scattering channel between the BS and the k-th user in group g.

where $\bar{P} > 0$ represents the average total Tx power, $\mathbf{P} = \text{diag}(p_{1,1}, \cdots, p_{K,1}, p_{1,2}, \cdots, p_{K,G}) \in \mathbb{R}^{K \times K}$, and $\mathbf{G} = [\mathbf{G}_1, \cdots, \mathbf{G}_G] \in \mathbb{C}^{N \times K}$ represents the precoding matrix with $\mathbf{G}_g = [\mathbf{g}_{1,g}, \cdots, \mathbf{g}_{K_g,g}] \in \mathbb{C}^{N \times K_g}$.

We assume that the signal transmitted from the BS to the k-th user in group g experience the double scattering channel fading proposed in [11]. An illustration of the double scattering channel model is provided in Fig. 2.1, where $\sigma_{t,g}$ and $\sigma_{s,g}$ denote the radiated signal’s angular spread from the BS array and the Tx scatterers respectively, and $\mu_{t,g}$ and $\mu_{s,g}$ represent the mean angle of departure of the radiated signal from the BS array and TX scatterers respectively, while $d_t$ and $d_s$ are the distance between adjacent BS antennas and the distance between adjacent scatterers.

The considered double scattering channel vector $\mathbf{h}_{k,g}$ is defined as [11]:

$$
\mathbf{h}_{k,g} = \sqrt{S_g} \left( \frac{1}{\sqrt{S_g}} \mathbf{R}_{BS_g}^{1/2} \mathbf{W}_g \tilde{\mathbf{w}}_{g,\mathbf{g}} \right),
$$

where $S_g$ denotes the number of the Tx and the Rx scatterers in g-th group, $\mathbf{R}_{BS_g} \in \mathbb{C}^{N \times N}$ denotes the correlation matrix between the BS antennas and the $S_g$ Tx scatterers, $\mathbf{W}_g \in \mathbb{C}^{N \times S_g}$ describes the small-scale fading between the BS and the scattering cluster at the Tx side following a standard complex Gaussian distribution,
\( \bar{S}_g \in \mathbb{C}^{S_g \times S_g} \) represents the correlation matrix between the \( S_g \) Tx and Rx scatterers and \( \bar{w}_{k,g} \sim \mathcal{CN}(0, \frac{1}{S_g} I_{S_g}) \in \mathbb{C}^{S_g \times 1} \) describes the small-scale channel fading between the \( k \)-th user in group \( g \) and the scattering cluster at the receiver side. In view of (2.4), the channel lies in a space with a rank that is determined by both the structure of scattering during the communication process and the spatial correlation between the antennas at the BS. Finally, it is worth mentioning that \( \bar{S}_g \) can be assumed to be diagonal. The assumption does not lead to any loss of statistics generality of the received signal, which can be easily proven by leveraging the bi-unitarily invariant property of standard Gaussian random matrices.

### 2.1 Channel Estimation

The acquisition of accurate CSI in a LS-MIMO system in a timely manner using downlink pilot signaling necessitates the number of downlink pilots to scale with the number of BS antennas, causing a large overhead. A common workaround to this issue is to operate in TDD mode in which downlink and uplink transmissions are sent over the same frequency, thus allowing the physical channel to be reciprocal between the downlink and the uplink links. Based on that, the uplink channel is estimated at the BS during a dedicated uplink phase prior to be used in a data transmission phase to implement precoding in the downlink.

More specifically, during the uplink training process, mutually orthogonal pilot sequences are transmitted by all users. Based on the received observation, \( y_{tr}^{k,g} \in \mathbb{C}^{N \times 1} \), the estimated channel vectors \( \hat{h}_{k,g} \) can be computed by the BS with MMSE estimation scheme. The received observation can be written as,

\[
y_{tr}^{k,g} = h_{k,g} + \frac{1}{\sqrt{\rho_{tr}}} n_{tr}^{k,g}, \quad (2.5)
\]

where \( n_{tr}^{k,g} \sim \mathcal{CN}(0, I_N) \), and \( \rho_{tr} > 0 \) denotes the given effective training SNR. Thus,
the MMSE estimate channel vector $\hat{h}_{k,g} \in \mathbb{C}^{N \times 1}$ of the non-Gaussian channel $h_{k,g}$ is obtained by \cite{22},

$$\hat{h}_{k,g} = d_g R_{BS_g} Q_g y_{k,g}^t, \quad (2.6)$$

where $Q_g = \left( d_g R_{BS_g} + \frac{1}{\rho_t} I_N \right)^{-1}$, and $d_g = \frac{1}{S_g^t} \text{tr} \left( \bar{S}_g \right)$.

### 2.2 Precoding Scheme and Achievable Rate

As explained earlier, MMSE estimate in \eqref{eq:mmse} is utilized at the BS to implement the MRT precoding which consists in precoding the symbol of each user by the Hermitian of its channel vector. Compared to RZF/ZF precoding which involves the inversion of the Gram matrix of joint users’ channel matrix, the MRT precoding presents a much lower computational complexity, thus making it attractive from a practical standpoint. More formally, the MRT precoding vector is given as $g_{k,g} = \zeta \hat{h}_{k,g}$ while the precoding matrix writes as ,

$$G^H = \zeta \hat{H}, \quad (2.7)$$

where $\hat{H} = \left[ \hat{H}_1^H, \hat{H}_2^H, \ldots, \hat{H}_G^H \right]^H \in \mathbb{C}^{K \times N}$, with $\hat{H}_g = \left[ \hat{h}_{1,g}, \hat{h}_{2,g}, \ldots, \hat{h}_{K_g,g} \right]^H \in \mathbb{C}^{K_g \times N}$, and $\zeta$ set in such a way that it ensures the power constraint described in \eqref{eq:power_constraint}, which implies

$$\zeta^2 = \frac{\bar{P}}{\mathbb{E} \left[ \text{tr} \left( P \hat{H} \hat{H}^H \right) \right]} = \frac{\bar{P}}{\bar{\Theta}}, \quad (2.8)$$

where $\Theta = \mathbb{E} \left[ \text{tr} \left( P \hat{H} \hat{H}^H \right) \right]$.

For the users to decode the transmitted signals at the downlink, the CSI which is estimated in uplink link is needed. These can be exchanged by the BS but this
would cause a large overhead given the high number of transmitting antennas. To
 circumvent this issue, the channel hardening property of LS-MIMO systems according
to which the effective useful channel $h^H_{k,g}g_{k,g}$ of $k$-th user in group $g$ converges to its
expectation value when the number of antennas and scatters increase to large is
invoked. Hence, it suffices for each user to acquire the knowledge of the statistical
CSI for the computation of $\mathbb{E}[h^H_{k,g}g_{k,g}]$. For simplicity, we assume that $\mathbb{E}[h^H_{k,g}g_{k,g}]$
is perfectly known. Decomposing $y_{k,g}$ as

$$y_{k,g} = \sqrt{p_{k,g}} \mathbb{E}[h^H_{k,g}g_{k,g}] s_{k,g} + \sum_{(k',g') \neq (k,g)} \sqrt{p_{k',g'}} h^H_{k',g'} g_{k',g'} s_{k',g'} + \sqrt{p_{k,g}} (h^H_{k,g}g_{k,g} - \mathbb{E}[h^H_{k,g}g_{k,g}]) s_{k,g} + n_{k,g}. \quad (2.9)$$

By treating the interference term as a noise, the ergodic achievable rate and the
ergodic achievable sum rate with MRT precoding can be written respectively as

$$R_{k,gMRT} = \log (1 + \gamma_{k,gMRT}) \quad (2.10)$$

and

$$R_{sumMRT} = \sum_{g=1}^{G} \sum_{K_g} R_{k,gMRT}, \quad (2.11)$$

where $\gamma_{k,gMRT}$ is the effective SINR of the $k$-th user in group $g$, defined as the ratio
of the effective signal power to the sum of interference power and noise power is given
as

$$\gamma_{k,g}\text{MRT} = \frac{p_{k,g} \left| \mathbb{E} \left[ h_{k,g}^H \hat{h}_{k,g} \right] \right|^2}{\mathbb{E} \left[ h_{k,g}^H \hat{H}_{[k,g]} \hat{H}_{[k,g]}^H \right] + p_{k,g} \text{var} \left[ h_{k,g}^H \hat{h}_{k,g} \right] + \frac{\Theta}{\rho}},$$

(2.12)

where $\rho = \frac{\bar{P}}{\sigma^2}$, $\hat{H}_{[k,g]} = \left[ \hat{H}_{1}^H, \cdots, \hat{H}_{g-1}^H, \hat{h}_{1,g}, \cdots, \hat{h}_{k-1,g}, \cdots, \hat{h}_{K,g}, \cdots, \hat{H}_{G}^H \right]^H$ $\in \mathbb{C}^{K-1 \times N}$ and $P_{[k,g]} = \text{diag} \left( p_{1,1}, \cdots, p_{K-1,g-1}, p_{1,g}, \cdots, p_{k-1,g}, p_{k+1,g}, \cdots, p_{K,g}, \cdots, p_{K,G} \right) \in \mathbb{C}^{K-1 \times K-1}$. 
Chapter 3

Main Results

In this chapter, we will leverage tools from random matrix theory to derive the deterministic approximations of the SINR under MRT precoding and assuming the channel to follow a double scattering fading model, and then derive the system sum rate accordingly.

3.1 Assumptions

Analyzing the expression of the SINR in (2.12) for fixed dimensions over double scattering channels is not an easy problem. The main idea in this work is to assume that all system dimensions $N$, $K_g$, and $S_g$ for each group $g$ tend to infinity and to resort to asymptotic tools based on random matrix theory results to approximate the ergodic user rates. Such an approach has already been pursued in previous works [3], [33] but mainly for Rayleigh channels. The approximation is expected to be tight for large dimensions, an assumption that is not restrictive given that with the advent of 5G, BSs are equipped with a high number of antennas to serve a large number of users. Prior to stating the main results, we shall first formalize the growth rate regime in the following assumption:

Assumption 1. In large $(S_g, K_g, N)$ regime, $S_g$, $K_g$ and $N$ tend to infinity such that

$$0 < \lim \inf \frac{S_g}{N} \leq \lim \sup \frac{S_g}{N} < \infty,$$  \hspace{1cm} (3.1)
\[ 0 < \lim \inf \frac{K_g}{N} \leq \lim \sup \frac{K_g}{N} < \infty. \quad (3.2) \]

In the sequel, the above assumption will be represented as \( N \to \infty \). Moreover, the deterministic sequence \( X_N^o \) is denoted as the deterministic equivalent of random variable \( X_N \). \( X_N^o \) approximates \( X_N \) with \( X_N - X_N^o \overset{a.s.}{\to} 0 \). Accordingly, the asymptotic approximation \( \gamma_{k,g}^{oMRT} \) of the SINR \( \gamma_{k,g}^{MRT} \) with MRT precoding satisfies,

\[ \gamma_{k,g}^{MRT} - \gamma_{k,g}^{oMRT} \overset{a.s.}{\to} 0. \quad (3.3) \]

Moreover, the fixed diagonal matrix \( P \) whose diagonal elements are power allocation weights for each user should satisfy the following assumption.

**Assumption 2.** The diagonal values \( p_{k,g} \) in \( P = \text{diag} (p_{1,1}, \ldots, p_{K,1}, p_{1,2}, \ldots, p_{K,G}) \) are positive and of order \( O \left( \frac{1}{K} \right) \).

Additionally, we require the following two assumptions:

**Assumption 3.** For all groups, \( \lim \sup_N \| R_{BS_g} \| < \infty \) and \( \lim \sup_s \| S_g \| < \infty \).

**Assumption 4.** For all groups, \( \lim \inf_N \frac{1}{N} \text{tr} R_{BS_g} > 0 \) and \( \lim \inf_s \frac{1}{N} \text{tr} S_g > 0 \).

The analysis starts by expressing the double-scattering channel in (2.4) as,

\[ h_{k,g} = \sqrt{S_g} Z_g \tilde{w}_{k,g}, \quad (3.4) \]

where \( Z_g = \frac{1}{\sqrt{S_g}} R_{BS_g}^{1/2} W_g S_g^{1/2} \). We will make use of the Fubini theorem [22] to first assume \( Z_g \) to be deterministic. Under the assumption, the estimation of the double scattering channel in (2.6) can be expressed as

\[ \hat{h}_{k,g} = \Phi_g^{1/2} q_{k,g}, \quad (3.5) \]
where $\bar{q}_{k,g} \sim \mathcal{CN}(0, I_N)$ and the covariance matrix of the estimated channel $\hat{\Phi}_g$ is given as,

$$
\Phi_g = d_g^2 R_{BSg} Q_g \left( Z_g Z_g^H + \frac{1}{\rho_{tr}} I_N \right) Q_g^H R_{BSg}^H.
$$

(3.6)

During the analysis process, we assume $Z_g$ to be fixed firstly, and then allow $Z_g$ to be random and compute the final deterministic equivalent of the SINR.

### 3.2 Useful Results

From (2.12), we can see that the deterministic equivalent of the SINR depends on the effective signal term, interference term and noise term. Therefore, we first derive the deterministic equivalents of these quantities in the following lemmas.

**Lemma 1.** Under the setting of Assumption 1-4, the deterministic equivalent of the effective signal power is

$$
\frac{p_{k,g}}{K} \left| \mathbb{E} \left[ h_{k,g}^H h_{k,g} \right] \right|^2 - \frac{p_{k,g}}{K} \left[ d_g^2 \text{tr} \left( R_{BSg} R_{BSg} Q_g \right) \right]^2 \xrightarrow{N \to \infty} 0.
$$

(3.7)

**Proof.** See Appendix A for the proof of Lemma 1. \qed

**Lemma 2.** The deterministic equivalent of the interference term under Assumption 1-4 is calculated as

$$
\frac{1}{K} \mathbb{E} \left[ h_{k,g}^H \hat{h}_{k,g} P_{[k,g]} \hat{h}_{[k,g]} h_{k,g} \right] - \sum_{i=1}^{G} \sum_{l=1}^{K_i} p_{l,i} (\gamma_{i,g}^o + \psi_{i,g}^o) \xrightarrow{N \to \infty} 0,
$$

(3.8)
where \( \Upsilon_{i,g}^o = \frac{d^2_d}{K \rho_{tr}} \text{tr} \left( \mathbf{R}_{BS_g} \mathbf{R}_{BS_i} \mathbf{Q}_i \mathbf{Q}_i^H \mathbf{R}_{BS_i}^H \right) \) and \( \Psi_{i,g}^o \) is shown as

\[
\Psi_{i,g}^o = \begin{cases} 
\frac{d^2_d}{K} \text{tr} \left( \mathbf{R}_{BS_g} \mathbf{R}_{BS_i} \mathbf{Q}_i \mathbf{R}_{BS_i}^H \mathbf{R}_{BS_i}^H \right), & \text{if } i \neq g \\
\frac{d^2_d}{K} \sum_{j=1}^{S_g} \frac{S_j}{S_g} \text{tr} \left( \mathbf{R}_{BS_g} \mathbf{Q}_g^H \mathbf{R}_{BS_g}^H \right) \text{tr} \left( \mathbf{R}_{BS_g} \mathbf{R}_{BS_i} \mathbf{Q}_i \right) + \frac{d^2_d}{K} \sum_{j=1}^{S_g} \frac{S_j}{S_g} \sum_{n=1}^{S_g} \frac{s_{g,n}}{S_g} \text{tr} \left( \mathbf{R}_{BS_g} \mathbf{R}_{BS_i} \mathbf{Q}_g \mathbf{R}_{BS_i} \mathbf{Q}_g^H \mathbf{R}_{BS_i}^H \right), & \text{if } i = g
\end{cases}
\]
Theorem 1. Under the setting of Assumption 1-4, the downlink effective SINR of $k$-th user in group $g$ defined in (2.12) converges almost surely to $\gamma_{k,g}^{o}$, which can be expressed as,

$$\gamma_{k,g}\text{MRT} - \gamma_{k,g}^{o}\text{MRT} \xrightarrow{\text{a.s.}} 0, \quad N \to \infty$$

where $\gamma_{k,g}^{o}$ is shown as

$$\gamma_{k,g}^{o} = \frac{p_{g,k}}{K} \left[ d_{g}^{2} \text{tr} \left( R_{BS} R_{BS} Q_{g} \right) \right]^{2} \sum_{i=1}^{G} \sum_{l=1}^{K_{i}} p_{l,i} \left( \Upsilon_{i,g}^{o} + \Psi_{i,g}^{o} \right) + \Lambda_{g}^{o}$$

Proof. Substituting the results of Lemma 1, Lemma 2 and Lemma 3 into (2.12) completes the proof.

Corollary 1. The individual downlink ergodic rate $R_{k,g}$ of $k$-th user in group $g$ converges as,

$$R_{k,g} - R_{k,g}^{o} \xrightarrow{\text{a.s.}} 0, \quad N \to \infty$$

where

$$R_{k,g}^{o} = \log \left( 1 + \gamma_{k,g}^{o} \right),$$

where $\gamma_{k,g}^{o}$ is given by (3.12).

Proof. The proof is completed based on the almost sure convergence of $\gamma_{k,g}\text{MRT}$ in (3.11) and applying the continuous mapping theorem [34] to the logarithm function.
An approximation of $R_{sumMRT}$ can be derived by replacing $R_{k,g}$ in (2.11) with its asymptotic approximation $R^o_{k,g}$ as follows

$$R^o_{sumMRT} = \sum_{g=1}^{G} \sum_{k=1}^{K_g} \log \left( 1 + \gamma^o_{k,gMRT} \right),$$

(3.15)

such that $\frac{1}{K} (R_{sumMRT} - R^o_{sumMRT}) \to 0$ almost surely. The asymptotic approximation of $R_{sumMRT}$ will be shown to be tight using simulations.

We can see that the deterministic equivalents only depend on the covariance matrices $R_{BSg}$ and $\bar{S}_g$, which vary slowly, instead of the fast varying instantaneous channels. Although these correlation matrices have huge size in the large $N$, $S_g$ regime, they can be computed at the BS only with the statistics properties of the large-scale channel.

Since the expressions derived above do not offer direct insights, we simplify them for the Rayleigh product channel and obtain scaling laws for standard Rayleigh fading and multi-keyhole fading conditions. We also compare them with the results in [22] for the RZF precoder.

### 3.4 Rayleigh Product Channel

The Rayleigh product channel is treated as the multi-keyhole channel when it does not account for any correlation [35]. Under this setting, **Theorem 1** can be simplified as in the following corollaries.

**Corollary 2.** For single group $G = 1$, let $S_1 = S$, $K_1 = K$, $R_{BS,1} = R_{BS}$, $Q_1 = Q$, and assume $\bar{S}_1 = I_S$. Then $\gamma^o_{k,gMRT}$ in **Theorem 1** can be simplified as,

$$\gamma^o_{kMRT} = \frac{p_k}{P (\Upsilon + \Psi) + \frac{P}{K} \left[ \frac{1}{P} \text{tr} \left( R_{BS} Q Q^H R_{BS}^H \right) + \text{tr} \left( R_{BS} Q^H R_{BS}^H R_{BS} Q \right) \right]},$$

(3.16)
where \( P = \sum_{k=1}^{K} p_k \) denotes the total Tx power at BS, \( \Upsilon = \frac{1}{K_{\rho\tau}} \text{tr} \left( R_{BS} R_{BS} Q_Q R_{BS}^H R_{BS}^H \right) \) and 
\[
\Psi = \frac{s-1}{SK} \text{tr} \left( R_{BS} R_{BS} QR_{BS} Q^H R_{BS}^H \right) + \frac{1}{SK} \text{tr} \left( R_{BS} Q^H R_{BS}^H \right) \text{tr} \left( R_{BS} R_{BS} Q \right). 
\]

\[\text{Proof.}\] For a single group \( G = 1 \), and \( \bar{S}_1 = I_S, \ d_1 = \frac{1}{S_1} \text{tr} \left( S_1 \right) = 1. \) Substituting the value of \( d_1 \) into (3.12) and reducing (3.12) to a single group completes the proof. \( \square \)

\textbf{Corollary 3.} For single group \( G = 1 \), assuming \( S_1 = S, K_1 = K, S_1 = I_S, \) and \( R_{BS_1} = I_N \). Then the closed-form of \( \gamma_{k,g_{MRT}}^o \) shown in Theorem 1 can be given as,

\[
\gamma_{k,MRT}^o = \frac{p_k N}{P \left( \frac{1}{\rho \tau} + 1 + \frac{N-1}{S} \right) + \frac{P}{\rho} \left( 1 + \frac{1}{\rho \tau} \right)}. 
\]  
(3.17)

\[\text{Proof.}\] See Appendix D for the proof of Corollary 3. \( \square \)

\subsection{3.4.1 Standard Rayleigh Channel}

If the number of scatterers \( S \) is very large as compared to \( N \) and \( K \), the Rayleigh product channel approaches the Rayleigh channel, the behavior of which can be predicted by assuming that \( \frac{S}{N} \to \infty \) and \( \frac{S}{K} \to \infty \),

\textbf{Corollary 4.} Based on the assumption of Corollary 3 with \( \frac{S}{N}, \frac{S}{K} \to \infty \), the expression of \( \gamma_{k,MRT}^o \) defined in Theorem 1 approaches the bound \( \gamma_{k,MRT, \frac{S}{N}, \frac{S}{K} \to \infty} \) which can be written as

\[
\gamma_{k,MRT, \frac{S}{N}, \frac{S}{K} \to \infty}^o = \frac{p_k N}{P \left( \frac{1}{\rho \tau} + 1 \right) + \frac{P}{\rho} \left( 1 + \frac{1}{\rho \tau} \right)}. 
\]  
(3.18)

As expected, we retrieve the asymptotic SINR expression of an uncorrelated Rayleigh fading channel obtained in [22].

\textbf{Remark 1.} Based on Assumption 2, \( p_k \) is of order \( O \left( \frac{1}{K} \right) \), we can conclude that the SINR under a Rayleigh fading channel scales as \( \frac{N}{K} \). This proves that the SINR increases linearly with the number of transmit antennas. As the Rayleigh channel
model does not account for the number of scatterers, their effect is naturally not reflected in the asymptotic SINR expression.

In the denominator of (3.18), the term $\frac{P}{\rho_{tr}} + \frac{P}{\rho_{tr}}$ represents the loss incurred due to an imperfect CSI. The perfect CSI case can be studied by taking the limit of the asymptotic SINR expression as $\rho_{tr} \to \infty$. In doing so, we obtain the following result:

**Corollary 5.** Based on the assumption of **Corollary 4**, let $\rho_{tr} \to \infty$. Then the SINR for the perfect CSI case is given by

$$\gamma_{k,MRT}^o, \frac{S}{N}, \frac{S}{K} \to \infty = \frac{p_k N}{P + \frac{P}{\rho}}.$$  \hfill (3.19)

Let us now compare the performance of the MRT precoding with RZF precoding in the perfect CSI case in the case $\frac{S}{N} \to \infty$ and $\frac{S}{K} \to \infty$. In [22], the $\gamma_{k,RZF,F}^o, \frac{S}{N}, \frac{S}{K} \to \infty$ under RZF precoding and perfect CSI for Rayleigh product channel is given as,

$$\gamma_{k,RZF}^o, \frac{S}{N}, \frac{S}{K} \to \infty = \frac{p_k N}{P + \frac{P}{\rho}}.$$  \hfill (3.20)

where $\alpha$ is the regularization parameter and $\bar{m} = \frac{1}{N} - a + \sqrt{(a + \frac{N}{K} - 1)^2 + 4\alpha}$. Since it is difficult to compare these two term, we re-express $\bar{m}$ through binomial series $(1 + x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \mathcal{O}(x^3)$. By only taking the term with first order $\mathcal{O}(\frac{K}{N})$, we can approximate $\bar{m}$ as $\bar{m} \approx \alpha \frac{K}{N}$. Substituting the approximated $\bar{m}$ into (3.20) and only keeping the terms of order $\frac{N}{K}$ in the numerator and the denominator, we can approximate the SINR with RZF precoding under Rayleigh fading channel as,

$$\gamma_{k,RZF}^o, \frac{S}{N}, \frac{S}{K} \to \infty = \frac{p_k N}{P + \frac{P}{\rho}}.$$  \hfill (3.21)

**Remark 2.** Comparing $\gamma_{k,RZF}^o, \frac{S}{N}, \frac{S}{K} \to \infty$ with $\gamma_{k,MRT}^o, \frac{S}{N}, \frac{S}{K} \to \infty$, we can clearly see that the performance with MRT precoding is close to the performance with RZF precoding.
under Rayleigh fading channel in the low SNR regime, i.e. small $\rho$. For high SNR regime, the SINR under RZF precoding is much better than SINR under MRT precoding because MRT introduces more interference power $P$ than RZF precoding. In the particular case when $p_k = \frac{1}{K}$, $\gamma_{k\text{RZF},\frac{S}{N},\frac{S}{K} \to \infty}^o = C_{\text{RZF}} \frac{N}{K}$, where $C_{\text{RZF}} = \frac{\rho}{P}$ and $\gamma_{k\text{MRT},\frac{S}{N},\frac{S}{K} \to \infty}^o = C_{\text{MRT}} \frac{N}{K}$, where $C_{\text{MRT}} = \frac{\rho}{P} + \frac{P}{P+P}$. It can be clearly seen that the multiplicative constant of RZF is greater than that of MRT but become very close to it in the low SNR regime.

3.4.2 Multi-keyhole Channel with Limited Scatterers

The number of scatterers is a parameter that depends on the propagation environment and over which the designer cannot have any control, while the number of antennas can be increased by equipping the BS with many of them. This motivates us to study whether an increase in the number of antennas can always result in an enhancement of the performances when a multi-keyhole channel with a limited number of scatterers is considered. Such a situation is modeled by the assumption $\frac{N}{S} \to \infty$ and $\frac{N}{K} \to \infty$.

Corollary 6. Based on the assumption of Corollary 3, let $\frac{N}{S}, \frac{N}{K} \to \infty$, and $p_k = \frac{P}{K}$. Then $\gamma_{k,g\text{MRT}}^o$ presented in Theorem 1 approaches the limit $\gamma_{k\text{MRT},\frac{N}{S},\frac{N}{K} \to \infty}^o$ which is obtained as,

$$\gamma_{k\text{MRT},\frac{N}{S},\frac{N}{K} \to \infty}^o = \frac{S}{K}.$$  \hfill (3.22)

Remark 3. Corollary 6 implies that for $\frac{N}{S}, \frac{N}{K} \to \infty$, $\gamma_{k\text{MRT},\frac{N}{S},\frac{N}{K} \to \infty}^o$ saturates at a fixed value $S/K$. Different from the result of Corollary 4 under Rayleigh fading, the multi-keyhole channel provides the SINR scaling with $O\left(\frac{S}{K}\right)$, which is independent of the number of transmit antennas at the BS. This does not only proves that the achievable SINR is determined by the number of scatterers in realistic environments but also that improving the number of transmit antennas at the BS is not a feasible
scheme to improve the SINR in LS-MIMO system under multi-keyhole channels.

With limited number of scatterers during the wireless communication process, the performance of a LS-MIMO system with MRT precoding is bounded. This is also the case of RZF precoding studied in [22]. As a matter of fact, according to the corollary of [22], $\gamma_{kRZF,N,S,N \to \infty}^{0}$ with $K > S$ is given by

$$\gamma_{kRZF,N,S,N \to \infty}^{0} = \frac{S}{K - S},$$

while it grows unboundedly with $N$ when $S < K$. Comparing the performance of RZF precoding with MRT precoding, we can see that MRT and RZF have similar performances when $K \gg S$, proving that in a very limited rank channel, there is no much gain to reap in going for the RZF precoding. However, if $K < S$, RZF precoding benefit from an increase of the number of antennas, similarly to what happens if a standard Rayleigh channel is considered. On the other hand, MRT precoding could not gain in this situation from an increase in the number of transmit antennas.
Chapter 4

Simulation Results

In this thesis, we consider a double scattering channel fading with correlation matrices \( \bar{S}_g = F(\mu_{s,g}, \sigma_{s,g}, d_{s,g}, S_g) \) and \( R_{BS_g} = F(\mu_{t,g}, \sigma_{t,g}, d_{t}, S_g) \), where \( F(\mu, \sigma, d, n) \) is given as [20],

\[
[F(\mu, \sigma, d, n)]_{k,l} = \frac{1}{n} \sum_{j=\frac{1}{2}}^{n-1} \exp \left[ -i 2\pi d(k-l) \cos \left( \frac{\pi}{2} + \frac{j\sigma}{n-1} + \mu \right) \right]. \tag{4.1}
\]

The main parameters adopting in simulation are set as \( K = 128, G = 4, S_g = \{130, 140, 160, 150\}, d_{s,g} = 2, d_{t} = 0.5, \sigma_{t,g} = \{\pi/5, \pi/6, \pi/5, \pi/7\}, \mu_{s,g} = \mu_{t,g} = \{-\pi/3, -\pi/9, \pi/9, \pi/3\}, \) and \( \sigma_{s,g} = \{\pi/6, \pi/6, \pi/6, \pi/6\} \). For simplicity, we assume power allocation \( P = \frac{1}{K} I_K \) with \( K_g = K/G \). Fig. 4.1 compares the downlink system ergodic sum rate using the deterministic equivalent provided in (3.15) using (3.12) to 1000 Monto-Carlo realizations of the SINR shown in (2.12). Under moderate system dimensions, the results show that the deterministic results yield a very close approximation. Since \( \gamma_{k,gMRT} \) converges to its deterministic approximation slowly, the mismatch between the Monte-Carlo results and the theoretical starts to grow in high SNR regime.

The comparison between RZF precoding and MRT precoding is also shown in Fig. 4.1. Even though MRT experiences performance loss as compared to RZF since it does not minimize the interference directly, the performance of MRT precoding can be improved by using a higher number of Tx antennas as done in the plotted figure. The MRT precoder with only 1 dB pilot training power and 128 antennas achieves
higher sum rate than RZF precoding with 10 dB pilot training power and 64 antennas in the low SNR regime. Moreover, MRT precoding is less affected from the decrease in training SNR in high SNR regime as compared to RZF. Thus the performance of MRT can be improved to and beyond the performance of RZF by using a higher number of Tx antennas while saving significantly on the computational complexity.

Fig. 4.2 investigates the effect of the number of scatterers on the single group system ergodic sum rate under multi-keyhole channel, i.e. $G = 1$, $S_g = I_S$, $R_{BSg} = I_N$, $N = 32$, $K = 200$ and $\rho_{tr} = 10$dB. We can see that the plotted approximation of the sum rate using (3.17) match the Monte-Carlo simulation closely when there is only few scatterers. When $S < N$, the spatial multiplexing gains increase linearly with $S$, while they decrease when $S > N$. The reason is that the number of antennas at the BS limits the degrees of freedom. The limiting sum rate as $\frac{S}{N}, \frac{S}{K} \to \infty$ is also plotted using the SINR in (3.18). The performance is shown to approach to the performance of a Rayleigh fading channel as the number of scatterers grows large. Moreover, the sum-rate with RZF precoding when $\frac{S}{N}, \frac{S}{K} \to \infty$ is shown to be close to the sum-rate
with MRT precoding in the low SNR regime, which verifies the correctness of Remark 2. As the SNR increases, the gap between the sum-rate with RZF precoder and that with MRT precoder grows since MRT brings in more interference power as shown in (3.19) and (3.21). In conclusion, MRT precoding approaches the performance of RZF in the low SNR regime with much lower computational complexity.

As the number of antennas at the BS antennas increases, the single group system performance under the Rayleigh product channel is shown in Fig. 4.3 and Fig. 4.4. It can be seen that the performance saturates at a limiting sum rate given by the result of Corollary 6. This result confirms the fact that deploying more antennas is useless when there are limited scatterers in the propagation environment. As shown in Fig. 4.3, the sum-rate with RZF precoder is also bounded by a fixed value given by substituting (3.23) in the sum-rate, when $K > S$. The sum-rate with MRT precoding is close to the one with RZF precoding when $K \gg S$ while the gap between the sum-rate under RZF and MRT becomes larger and larger as $S$ grows. Such an observation is an agreement with what we deduced earlier from the comparison between (3.23)
and (3.22). However, the performance of MRT precoder is bounded when $S > K$ to the limit $S/K$, while the performance in RZF precoder grows unboundedly with $N$ when $S > K$, as shown in Fig. 4.4. Moreover, a higher number of $N$ is needed to reach the bound as the value of $S$ increases. In our case, the performance of the Rayleigh product channel grows unboundedly only when $\frac{S}{K} \rightarrow \infty$, which is in accordance with the results of Corollary 6. In conclusion, the spatial multiplexing gains in realistic implementations will always be limited by the number of scatterers when using MRT precoder. Moreover, MRT precoding could provide near-optimal performance with much lower computational complexity when there is a limited number of scatterers and many active users.
Figure 4.4: Sum rate versus N with $K = 40$. 
Chapter 5

Concluding Remarks

5.1 Summary

In this thesis, we derived the deterministic equivalents of the effective SINR as well as ergodic rates adopting MRT precoding scheme for a LS multi-user MISO system. The transmitted signal from BS to users is assumed to experience the double scattering channel fading. Unlike the standard Rayleigh fading channel, this model accounts for keyhole channel effect, a practical situation encountered in case of limited number of scatterers. Although being more realistic, the study of this model has been remained scarce, thus motivating the present work. Even though derived under the assumption of large system dimensions, these deterministic equivalents where shown to be close to Monte-Carlo results for moderate system dimensions as well. In an effort to draw meaningful insights into the performance of MRT precoding over the double scattering channels, a simplified multi-keyhole channel is considered. Particularly, it has been shown that contrary to a Rayleigh fading channel which presents a performance that improves with the number of Tx antennas, the performance of MRT precoding over multi-keyhole channels saturates at a limit that depends on the number of scatterers. A comparison with the RZF precoding reveals that MRT results in a comparable performance with RZF at low SNR or when the number of scatterers is limited and a large number of users is served. In such a situation, it becomes an interesting alternative since it allows to avoid the prohibitively high computational complexity of RZF.
5.2 Future Research Work

We will extend the double scattering model into the multi-cell Massive MIMO system, where the effective average SINR as well as the sum rate will be investigated to see the influence of the number of double scatterers, the number of antennas, and the number of users with MMSE channel estimator. Moreover, we will illustrate the pilot contamination by considering a scenario where every cell has the same number of users and uses the same pilot sequence. It means that BS $l$ estimated the channel of UE $k$ in cell $l$, while UE $k$ in other cells transmit the same pilots. In this case, the mutual interference is introduced during pilot transmission, which leads to correlated and reduced channel estimation. Some details of the considered system model are now presented:

In the future work, the system sum rate of a LS-MIMO system with $L$ cells is studied. The BS in each cell is assumed to be equipped with $N$ antennas and serve $K$ single-antenna users sharing the same time-frequency resource. We assume channel reciprocity for the forward and reverse links, i.e., the propagation condition is same for both forward and reverse links, and block fading. Moreover, the channel fading remains constant for a duration of $T$ symbols, which means a constant and flat-fading channel $h_{l,k}^i \in \mathbb{C}^M$ between user $k$ in cell $l$ and BS in cell $i$ (for any $k,l,i$). During coherence intervals, the channel realizations are assumed to be independent. Also, $\tau_p$ symbols are utilized for uplink pilot signaling. The received complex baseband signal vector at $k$-th users in the $l$-th cell is

$$y_{l,k} = \sum_{i=1}^{L} (h_{l,k}^i)^H x_i^l + n_{l,k}, \quad (5.1)$$

where $(h_{l,k}^i)^H \in \mathbb{C}^{1 \times N}$ is the channel matrix form the BS of $i$-th cell to users in $l$-th
cell, $x_i^t$ is the Tx signal vector at BS of $i$-th cell, which is given as

$$x_i^t = \sum_{t=1}^{K} \sqrt{p_{i,t}^i} g_{i,t}^i s_{i,t}^i,$$

(5.2)

where $p_{i,t}^i \geq 0$ denotes the signal power, $g_{i,t}^i \in \mathbb{C}^{N \times 1}$ is the precoding vector, and $s_{i,t}^i \sim \mathcal{CN}(0,1)$ represents the data symbol at BS of $i$-th cell for user $t$ of cell $i$ respectively. The transmitted signal vectors satisfy

$$\mathbb{E} [\|x_i^t\|^2] = \mathbb{E} \left[ \text{tr} \left( P_i^i (G_i^i)^H G_i^i \right) \right] \leq \overline{P},$$

(5.3)

where $\overline{P} > 0$ is the average total Tx power at every BS, $P_i^i = \text{diag}(p_{i,1}^i, \cdots, p_{i,K}^i) \in \mathbb{R}^{K \times K}$ and $G_i^i = [g_{i,1}^i, \cdots, g_{i,K}^i] \in \mathbb{C}^{N \times K}$ is the precoding matrix of cell $i$.

We assume that the channel between the BS and user $k$ in group $g$ follows the double scattering channel model proposed in [11]. The double scattering channel vector $h_{i,k}^l$ is modeled as [11]:

$$h_{i,k}^l = \sqrt{S_{i,k}^l} \left[ \frac{1}{\sqrt{S_{i,k}^l}} (R_{i,k}^l)^{1/2} W_{i,k}^l \left( \overline{S}_{i,k}^l \right)^{1/2} \right] \tilde{w}_{i,k}^l,$$

(5.4)

where $S_{i,k}^l$ is the number of scatters at the $l$-th Tx and $k$-th Rx sides in cell $i$, $R_{i,k}^l \in \mathbb{C}^{M \times M}$ is the correlation matrix between the $l$-th BS antennas and the $S_{i,k}^l$ Tx scatters, $\overline{S}_{i,k}^l \in \mathbb{C}^{S_{i,k}^l \times S_{i,k}^l}$ is the correlation matrix between the $S_{i,k}^l$ Tx and Rx scatters. The small-scale fading between the $l$-th BS and the Tx side scattering cluster is described as $W_{i,k}^l \in \mathbb{C}^{M \times S_{i,k}^l}$, which is a standard complex Gaussian matrix, while the small-scale channel fading between the $k$-th user in cell $i$ and the cluster at the $k$-th Rx side is described as $\tilde{w}_{i,k}^l \sim \mathcal{CN}(0, S_{i,k}^l I_{S_{i,k}^l}) \in \mathbb{C}^{S_{i,k}^l \times 1}$. $S_{i,k}^l$ can be assumed to be diagonal without any loss of the received signal’s statistics generality.

All users are assumed to transmit $\tau_p = fK$ pilot sequences at the same time for
channel estimation purposes. $f$ is a positive integer, which denotes the pilot reuse factor. Although users in each cell use the orthogonal pilots, users locate in different cells might use the same pilot. It happens when the the overall number of users in the system is larger than the length of the pilot sequences, i.e., $LK > \tau_p$. The reason is that the pilot sequences are reused in the system when $f < L$. The received pilot signal $Y^l \in \mathbb{C}^{M \times \tau_p}$ at BS $l$ is expressed

$$Y^l = \sum_{i=1}^{L} H_i^l \Phi_i^H + \frac{1}{\sqrt{\rho_{tr}}} N^l,$$  \hspace{1cm} (5.5)$$

where $H_i^l = [h_{i,1}^l, \ldots, h_{i,K}^l] \in \mathbb{C}^{M \times K}$, the orthogonality of pilot sequences in a cell implies that the $\tau_p \times K$ pilot matrix used in cell $i$, $\Phi_i = [\phi_{i,1}, \ldots, \phi_{i,K}]$, satisfies $\Phi_i^H \Phi_i = \tau_p I_K$. The pilots are divided into $f$ distinct groups and each cell belongs to one such group. If cell $i$ and cell $l$ use the same pilot sequences, we also have $\Phi_i^H \Phi_l = \tau_p I_K$. Otherwise, $\Phi_i^H \Phi_l = 0$ meaning that these cells use different orthogonal pilot sequences. Let us denote by $\mathcal{P}_l \subset \{1, \ldots, L\}$ the indices of the cells employing the same pilot sequence as cell $l$. Additionally, we let $\rho_{tr} > 0$ denote the effective training SNR of each user. Finally, each entries of noise matrix $N^l \in \mathbb{C}^{M \times \tau_p}$ follow the Gaussian distribution $\mathcal{CN}(0, \sigma^2)$.

Then, $l$-th BS can adopt MMSE estimation to obtain the channel estimate $\hat{h}_{l,k}^l$ based on the received pilot sequence, which is expressed as

$$y_{l,k}^l = Y_l^l \phi_{l,k} = \sum_{i \in \mathcal{P}_l} \tau_p h_{i,k}^l + \frac{1}{\sqrt{\rho_{tr}}} \tilde{n}_{l,k}^l,$$  \hspace{1cm} (5.6)$$

where $\tilde{n}_{l,k}^l = N_l^l \phi_{l,k} \sim \mathcal{CN}(0, \tau_p \sigma^2 I_M)$. Note that the channels from users that send the same uplink pilot sequence $\phi_{l,k}$ are contained in the summation $\sum_{i \in \mathcal{P}_l} h_{i,k}^l$. The
MMSE estimator $\hat{h}_{l,k}^i$ is

$$\hat{h}_{l,k}^i = B_{l,k}^i y_{l,k}^i = B_{l,k}^i \left( \sum_{i \in P_l} \tau_h h_{i,k}^l + \frac{1}{\sqrt{\rho_{tr}}} \tilde{u}_{l,k}^i \right), \quad (5.7)$$

where $B_{l,k}^i = C_{h_{l,k}^i} C_{y_{l,k}^i}^{-1}$.

Next, we provide SINR expression for downlink transmission. With the above communication scheme, the BSs have channel estimates while the users do not have any channel estimate. From (5.1) and (5.2), the signal received by the $k$-th user in cell $l$ can be re-expressed as

$$y_{l,k} = \sum_{i=1}^L \sum_{t=1}^K \sqrt{p_{i,t}^l} (h_{i,k}^l)^H g_{i,t}^i s_{i,t}^l + n_{l,k}$$

$$= \sqrt{p_{l,k}^l} (h_{l,k}^l)^H g_{l,k}^l s_{l,k}^l + \sum_{(i,t) \neq (l,k)} \sqrt{p_{i,t}^l} (h_{i,k}^l)^H g_{i,t}^i s_{i,t}^l + n_{l,k}. \quad (5.8)$$

The BS then utilizes the MMSE estimate shown in (5.7) to conduct digital precoding without CSI. However, the user cannot compute the effective downlink SINR without CSI. Since the useful channel $(h_{l,k}^l)^H g_{l,k}^l$ converges to its expectation value when $M, S_{l,k}^l$ increase with a well-known channel hardening feature of LS-MIMO, it is sufficient for each user to ignore the resulting performance loss in the large system with the statistical CSI. Using this idea, $y_{l,k}$ can be decomposed as

$$y_{l,k} = \sqrt{p_{l,k}^l} \mathbb{E} \left[ (h_{l,k}^l)^H g_{l,k}^l \right] s_{l,k}^l + \sqrt{p_{l,k}^l} \left( (h_{l,k}^l)^H g_{l,k}^l - \mathbb{E} \left[ (h_{l,k}^l)^H g_{l,k}^l \right] \right) s_{l,k}^l$$

$$+ \sum_{(i,t) \neq (l,k)} \sqrt{p_{i,t}^l} (h_{i,k}^l)^H g_{i,t}^i s_{i,t}^l + n_{l,k}. \quad (5.9)$$

Assuming that $\mathbb{E} \left[ (h_{l,k}^l)^H g_{l,k}^l \right]$ is perfectly known at the user. Thus, the user could
achieves SINR as

$$\gamma_{l,k} = \frac{p_{l,k}^l \left| \mathbb{E} \left[ (h_{l,k}^l)^H g_{l,k}^l \right] \right|^2}{p_{l,k}^l \text{var} \left[ (h_{l,k}^l)^H g_{l,k}^l \right] + \sum_{(i,t) \neq (l,k)} p_{i,t}^i \mathbb{E} \left[ (h_{i,k}^i)^H g_{i,t}^i \right]^2 + \sigma^2}$$  (5.10)

This work also considers MF precoding with simple closed-form expression given as

$$g_{l,t}^i = \xi_l h_{l,t}^i$$  (5.11)

and the precoding matrix is

$$G_l^i = \xi_l \hat{H}_l^i,$$  (5.12)

where $\hat{H}_l^i = [h_{l,1}^i, \ldots, h_{l,K}^i] \in \mathbb{C}^{M \times K}$ and $\xi_l$ ensures the power constraint shown in (5.3). Thus, we can derive that

$$\mathbb{E} \left[ \text{tr} \left( P_l^l (G_l^i)^H G_l^i \right) \right] = \mathbb{E} \left[ \text{tr} \left( P_l^l \xi_l (\hat{H}_l^i)^H \xi_l \hat{H}_l^i \right) \right] = \xi_l^2 \mathbb{E} \left[ \text{tr} \left( P_l^l (\hat{H}_l^i)^H \hat{H}_l^i \right) \right] \leq \mathbb{P}.$$  (5.13)

Thus, we could express $\xi_l$ as

$$\xi_l^2 \leq \frac{\mathbb{P}}{\mathbb{E} \left[ \text{tr} \left( P_l^l (\hat{H}_l^i)^H \hat{H}_l^i \right) \right]} = \frac{\mathbb{P}}{\Theta_l},$$  (5.14)

where $\Theta_l = \mathbb{E} \left[ \text{tr} \left( P_l^l (\hat{H}_l^i)^H \hat{H}_l^i \right) \right]$.

Thus SINR for $k$-th user in $l$ cell could be re-expressed as

$$\gamma_{l,k,MF} = \frac{p_{l,k}^l \left| \mathbb{E} \left[ (h_{l,k}^l)^H \hat{h}_{l,k}^l \right] \right|^2}{p_{l,k}^l \text{var} \left[ (h_{l,k}^l)^H \hat{h}_{l,k}^l \right] + \sum_{(i,t) \neq (l,k)} p_{i,t}^i \xi_l^2 \mathbb{E} \left[ \left| (h_{i,k}^i)^H \hat{h}_{i,t}^i \right|^2 \right] + \frac{\Theta_l}{\rho}}.$$  (5.15)
where $\rho = \frac{P}{\sigma^2}$.

The user $k$ in cell $l$ could achieve downlink ergodic rate as

$$R_{l,k} = \log (1 + \gamma_{l,k,MF}).$$  \hspace{1cm} (5.16)

The ergodic achievable sum rate for cell $l$ is given as

$$R_{l,sum} = \sum_{t=1}^{K} R_{l,k},$$  \hspace{1cm} (5.17)

while the ergodic achievable sum rate for the whole network could be expressed as

$$R = \sum_{i=1}^{L} R_{i,sum}.$$  \hspace{1cm} (5.18)

In the future research, we are interested in deriving the tight approximations of SINR and the ergodic rate assuming the number of BS antennas, users, and scatterers grow large with the same space by adopting some random matrix tools.
REFERENCES


APPENDICES

A Proof of Lemma 1

As shown in \cite{36} (Appendix E), \(Z_g\) in (3.4) satisfies that \(\lim \sup_N \|Z_g Z_g^H\| < \infty\) almost surely. As \(Z_g\) is independent of \(\tilde{\mathbf{w}}_{k,g}\) for \(k = 1, \ldots, K_g\), we proceed in two steps. By assuming \(Z_g\) to be deterministic firstly, we use trace lemma (Lemma 3 from \cite{37}) and then leverage the random model of \(Z_g\) to find a deterministic equivalent of the SINR.

Substituting (3.4) into the MMSE estimate given in (2.6), we obtain:

\[
\frac{1}{K} \mathbf{h}_{k,g}^H \hat{\mathbf{h}}_{k,g} = \frac{d_g}{K} \mathbf{w}_{k,g}^H Z_g^H \mathbf{R}_{BS_g} Q_g Z_g \mathbf{w}_{k,g} + \frac{d_g}{K \sqrt{\rho_{tr}}} \mathbf{w}_{k,g}^H Z_g^H \mathbf{R}_{BS_g} Q_g n_{k,g}^{tr}, \tag{A.1}
\]

where \(\mathbf{w}_{k,g} = \sqrt{S_g} \tilde{\mathbf{w}}_{k,g} \sim \mathcal{CN}(0, \mathbf{I}_{S_g})\).

The second term in the right-hand side of the above equation converges to zero, while the second term converges to \(\frac{d_g}{K} \text{tr} (Z_g^H \mathbf{R}_{BS_g} Q_g)\), thus yielding:

\[
\frac{1}{K} \mathbf{h}_{k,g}^H \hat{\mathbf{h}}_{k,g} - \frac{d_g}{K} \text{tr} (Z_g^H \mathbf{R}_{BS_g} Q_g Z_g) \xrightarrow{a.s.} 0, \tag{A.2}
\]

where \(Z_g\)'s are random and modeled as \(\frac{1}{\sqrt{S_g}} \mathbf{R}_{BS_g}^{1/2} \mathbf{W}_g S_g^{1/2} \mathbf{Z}_g\) in the double scattering.
model. We thus obtain:

\[
\frac{d_g}{K} \text{tr} (Z_g^H R_{BS_g} Q_g Z_g) = \frac{d_g}{K} \sum_{j=1}^{s_g} z_{g,j}^H R_{BS_g} Q_g z_{g,j} = \frac{d_g}{K} \sum_{j=1}^{s_g} \overline{s}_{g,j} \sqrt{\frac{S_g}{S_g}} w_{g,j}^H R_{BS_g}^{1/2} R_{BS_g} Q_g R_{BS_g}^{1/2} w_{g,j}, 
\]

where \( z_{g,j} = \sqrt{\frac{s_g}{S_g}} R_{BS_g}^{1/2} w_{g,j} \) is the \( j \)-th column of \( Z_g \), \( \overline{s}_{g,j} \) is the \( j \)-th diagonal element of \( S_g \), \( w_{g,j} \) is the \( j \)-th column of \( W_g \), which is composed of standard complex Gaussian elements. Thus, we can have following convergence,

\[
\frac{d_g}{K} \text{tr} (Z_g^H R_{BS_g} Q_g Z_g) - \frac{d_g}{K} \text{tr} (R_{BS_g}^2 Q_g) \xrightarrow{a.s.} 0. 
\]  

(A.4)

Now combining (A.1) with (A.2) will yield,

\[
\frac{1}{K} \mathbf{h}_{k,g}^H \hat{\mathbf{h}}_{k,g} - \frac{d_g^2}{K} \text{tr} (R_{BS_g}^2 Q_g) \xrightarrow{a.s.} 0. 
\]

(A.5)

It should be noted that the above convergence also implies the convergence in mean as the sequence of random variables \( \frac{1}{K} \mathbf{h}_{k,g}^H \hat{\mathbf{h}}_{k,g} \) is uniformly integrable. Therefore, we can have

\[
\frac{1}{K} \mathbb{E} \left[ \mathbf{h}_{k,g}^H \hat{\mathbf{h}}_{k,g} \right]^2 - \frac{1}{K} \left[ \frac{d_g^2}{K} \text{tr} (R_{BS_g}^2 Q_g) \right]^2 \xrightarrow{a.s.} 0. 
\]

(A.6)

The proof of Lemma 1 is completed.
B Proof of Lemma 2

Similar to the proof of Lemma 1, we also consider $Z_g$ to be deterministic and use (3.4) to have

$$
\frac{1}{K} h_{k,g}^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} h_{k,g} = \frac{1}{K} \bar{w}_{k,g}^H Z_g^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} Z_g \bar{w}_{k,g}. \tag{B.1}
$$

Using the trace lemma, we have

$$
\frac{1}{K} h_{k,g}^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} h_{k,g} - \frac{1}{K} \mathrm{tr} \left( Z_g^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} Z_g \right) \xrightarrow{\mathrm{a.s.}} 0. \tag{B.2}
$$

As the difference between term $\mathrm{tr} \left( Z_g^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} Z_g \right) = \sum_{i=1}^G \sum_{l=1}^{K_i} p_{l,i} \hat{h}_{l,i}^H Z_g \hat{h}_{l,i}$ and the term $\mathrm{tr} \left( Z_g^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} Z_g \right) = \sum_{i=1}^G \sum_{l=1}^{K_i} \sum_{l,(i)
eq (k,g)} p_{l,i} \hat{h}_{l,i}^H Z_g \hat{h}_{l,i}$ is of order $O(1/K)$, we may substitute in $\mathrm{tr} \left( Z_g^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} Z_g \right) \hat{H}_{[k,g]}$ by $\hat{H}$ and $P_{[k,g]}$ by $P$. Using (3.5) and applying trace lemma, we have following convergence

$$
\frac{1}{K} h_{k,g}^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} h_{k,g} - \sum_{i=1}^G \sum_{l=1}^{K_i} p_{l,i} T_{i,g} \xrightarrow{\mathrm{a.s.}} 0, \tag{B.3}
$$

where $T_{i,g} = \frac{1}{K} \mathrm{tr} \left( \Phi_i Z_g Z_g^H \right)$.

Replacing $\Phi_i$ by (3.6), $T_{i,g}$ could be further expressed as

$$
T_{i,g} = \frac{d_i^2}{K \rho_{tr}} \mathrm{tr} \left( Z_g^H R_{BS,i} Q_i Q_i^H R_{BS,i}^H Z_g \right) + \frac{d_i^2}{K} \mathrm{tr} \left( Z_g^H R_{BS,i} Q_i Z_i Z_i^H Q_i^H R_{BS,i}^H Z_g \right)
$$

$$
= \Upsilon_{i,g} + \Psi_{i,g}. \tag{B.4}
$$

Note that $\Upsilon_{i,g}$ is still a function of $Z_g$ which is random. We treat it in a same
way as shown in (A.3) and use trace lemma to obtain the following deterministic approximation,

\[ \gamma_{i,g} = \frac{d_i^2}{K} \rho_{\text{tr}} \text{tr} \left( R_{BS_i} R_{BS_i}^H Q_i Q_i^H R_{BS_i}^H \right) \xrightarrow{\text{a.s., } N \to \infty} 0. \]  

(B.5)

In order to eliminate the dependence of \( \Psi_{i,g} \) on \( Z_g \) and \( Z_i \), we divide it into two cases: \( \bar{\Psi}_{i,g} \) for \( i \neq g \) and \( \Psi_{g,g} \) for \( i = g \). For the first case, \( Z_g \) and \( Z_i \) are independent. By assuming \( Z_g \) to be deterministic, we can obtain,

\[ \bar{\Psi}_{i,g} - \frac{d_i^3}{K} \text{tr} \left( R_{BS_i} Q_i^H R_{BS_i}^H Z_g Z_g^H R_{BS_i} R_{BS_i}^H \right) \xrightarrow{\text{a.s., } N \to \infty} 0. \]  

(B.6)

Extending the analysis to random \( Z_g \) in a same way as done in (A.3), we have

\[ \Psi_{i,g} - \frac{d_i^3}{K} \text{tr} \left( R_{BS_i} Q_i^H R_{BS_i} Q_i R_{BS_i} Q_i^H R_{BS_i} \right) \xrightarrow{\text{a.s., } N \to \infty} 0. \]  

(B.7)

For the second case: \( i = g \), \( \Psi_{g,g} \) can be re-expressed as

\[ \Psi_{g,g} = \frac{d_g^2}{K} \sum_{j=1}^{S_g} z_{g,j}^H Q_g^H R_{BS_g}^H \left( Z_g z_g^H - z_{g,j} z_{g,j}^H \right) R_{BS_g}^H Q_g z_{g,j} \]

\[ + \frac{d_g^2}{K} \sum_{j=1}^{S_g} z_{g,j}^H Q_g R_{BS_g} z_{g,j} z_{g,j}^H R_{BS_g} Q_g z_{g,j} \]

\[ = Z_{1,g} + Z_{2,g}. \]  

(B.8)

Since we have removed the dependence of \( Z_{1,g} \) on the vector \( z_{g,j} \), we can apply trace lemma to have,

\[ Z_{1,g} - Z_{3,g} \xrightarrow{\text{a.s., } N \to \infty} 0, \]  

(B.9)
where $Z_{3,g}$ could be expressed as

$$Z_{3,g} = \frac{d_g^3}{K} \text{tr} \left( R_{BSg} Q_g^H R_{BSg}^H \left( Z_g Z_g^H - z_{g,j} z_{g,j}^H \right) R_{BSg} Q_g \right). \quad (B.10)$$

Next we re-express $Z_{3,g}$ as,

$$Z_{3,g} = \frac{d_g^2}{K} \sum_{j=1}^{S_g} \bar{s}_{g,j}^2 \sum_{n=1}^{S_g} \frac{\bar{s}_{g,n}}{S_g} \text{tr} \left( R_{BSg} R_{BSg}^H Q_g R_{BSg}^H Q_g^H R_{BSg}^H z_{g,n} \right). \quad (B.11)$$

Using the trace lemma, we obtain the following convergence result for $Z_{3,g}$,

$$Z_{3,g} - \frac{d_g^2}{K} \sum_{j=1}^{S_g} \bar{s}_{g,j}^2 \sum_{n=1}^{S_g} \frac{\bar{s}_{g,n}}{S_g} \text{tr} \left( R_{BSg} R_{BSg}^H Q_g R_{BSg}^H Q_g^H R_{BSg}^H \right) \xrightarrow{a.s.} 0 \quad (B.12)$$

Substituting the results of (B.12) into (B.9), the final deterministic equivalent of $Z_{1,g}$ is obtained. Moreover, using the trace lemma on $Z_{2,g}$, we obtain,

$$Z_{2,g} - \frac{d_g^2}{K} \sum_{j=1}^{S_g} \bar{s}_{g,j}^2 \sum_{n=1}^{S_g} \frac{\bar{s}_{g,n}}{S_g} \text{tr} \left( R_{BSg}^H R_{BSg} Q_g^H R_{BSg} Q_g^H \right) \xrightarrow{a.s.} 0 \quad (B.13)$$

Substituting the result of (B.13) and the result of (B.9), we obtain the deterministic approximation of $\Psi_{g,g}$. Combining this result with (B.6), we obtain the deterministic equivalent of $\Psi_{i,g}$, which is given by (3.9). Using (B.3) and (B.4), we finally obtain,

$$\frac{1}{K} h_{k,g}^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} h_{k,g} - \sum_{i=1}^{G} \sum_{l=1}^{K_i} \rho_{i,l} \left( \Upsilon_{i,g}^o + \Psi_{i,g}^o \right) \xrightarrow{a.s.} 0 \quad (B.14)$$

In a similar way to the previous analysis, based on the fact that the sequence of random variables $\frac{1}{K} h_{k,g}^H \hat{H}_{[k,g]}^H P_{[k,g]} \hat{H}_{[k,g]} h_{k,g}$ is uniformly integrable, the deterministic equivalent of Lemma 2 is obtained.
C Proof of Lemma 3

First, we compute the asymptotic approximation of the variance term. Using the expression from (A.1), \( \text{var} \left( \frac{1}{K} h_{k,g}^H \hat{h}_{k,g} \right) \) can be re-expressed as

\[
\text{var} \left( \frac{1}{K} h_{k,g}^H \hat{h}_{k,g} \right) = \text{var} \left( X + Y \right),
\]

where \( X = \frac{d_g}{K} \tilde{w}_{k,g}^H z_g^H R_{BS_g} Q_g Z_g \tilde{w}_{k,g} \) and \( Y = \frac{d_g}{K \sqrt{\text{tr}}} \tilde{w}_{k,g}^H z_g^H R_{BS_g} Q_g n_{tr}^r k_g \). It can be seen that \( \text{var} \left( X + Y \right) \) is bounded by \( 2 \text{var} \left( X \right) + 2 \text{var} \left( Y \right) \). The variance of \( Y \) can be bounded by \( \mathbb{E} \left[ Y^2 \right] \) which converges to zero. As for the variance of \( X \), using the trace Lemma and based on the fact that \( \sup_N \| Z_g Z_g^H \|_2 < \infty \) almost surely, we can prove that \( \text{var}(X) \) converges to zero.

Thus, we can conclude that the noise term can be approximated in the almost sure sense by \( \Theta_{\rho} \), which can be derived as follows.

Using the definition of \( h_{k,g} \) from (3.5), we have

\[
\frac{1}{K} h_{k,g}^H \hat{h}_{k,g} = \frac{1}{K} q_{k,g}^H \Phi_g q_{k,g}. \tag{C.2}
\]

Based on the trace lemma, we have

\[
\frac{1}{K} h_{k,g}^H \hat{h}_{k,g} - \frac{1}{K} \text{tr} (\Phi_g) \xrightarrow{a.s.} 0. \tag{C.3}
\]

Using the defined expression of \( \Phi_g \) from (3.6), we have

\[
\text{tr} (\Phi_g) = d_g^2 \text{tr} \left( Z_g^H Q_g^H R_{BS_g}^H R_{BS_g} Q_g Z_g \right) + \frac{d_g^2}{\rho_{tr}} \text{tr} \left( R_{BS_g} Q_g Q_g^H R_{BS_g}^H \right). \tag{C.4}
\]
Then, we can derive the deterministic equivalent of $\frac{1}{K} \text{tr} (\Phi_g)$ as

$$
\frac{1}{K} \text{tr} (\Phi_g) - \frac{1}{K} \left[ d^3_g \text{tr} \left( R_{BS_g} Q_g^H R_{BS_g}^H R_{BS_g} Q_g \right) + \frac{d^2_g}{\rho_{tr}} \text{tr} \left( R_{BS_g} Q_g Q_g^H R_{BS_g}^H \right) \right] \xrightarrow{a.s.} 0.
$$

(C.5)

Combining (C.5) and (C.3) yields the asymptotic approximation of $\frac{1}{K} \hat{h}_{k,g}^H \hat{h}_{k,g}$. We can see that $\frac{1}{K} \text{tr} \left( \hat{P} \hat{H} H^H \right) = \frac{1}{K} \sum_{g=1}^{G} \sum_{k=1}^{K_g} p_{k,g} \hat{h}_{k,g}^H \hat{h}_{k,g}$ leading to

$$
\Theta \frac{1}{K \rho} - \frac{1}{K \rho} \sum_{g=1}^{G} \sum_{k=1}^{K_g} p_{k,g} \left[ \frac{d^2_g}{\rho_{tr}} \text{tr} \left( R_{BS_g} Q_g Q_g^H R_{BS_g}^H \right) \right] \xrightarrow{a.s.} 0.
$$

(C.6)

Combining the deterministic equivalent of $\Theta \frac{1}{K \rho}$ with $\text{var} \left( \frac{1}{K} \hat{h}_{k,g}^H \hat{h}_{k,g} \right) \xrightarrow{a.s.} 0$, the proof of the Lemma 3 is completed.
D Proof of Corollary 3

For a single group $G = 1$, $\bar{S}_1 = I_S$, $R_{BS_1} = I_N$, $d_1 = \frac{1}{S_1} \text{tr} (S_1) = 1$, and $Q_g = \frac{\rho_{tr}}{1 + \rho_{tr}} I_N$. The expressions shown in (A.6) and (C.6) can be straightforwardly reduced as,

$$\frac{p_k}{K} \left\| \mathbb{E} \left[ h_k^H \hat{h}_k \right] \right\|^2 - \frac{p_{k,g} N^2 \rho_{tr}^2}{K (1 + \rho_{tr})^2} \xrightarrow{\text{a.s.}} 0$$ (D.1)

and,

$$\frac{\Theta}{K \rho} - \frac{P}{\rho} \left[ \frac{N \rho_{tr}^2}{K (1 + \rho_{tr})^2} + \frac{N \rho_{tr}^2}{K \rho_{tr} (1 + \rho_{tr})^2} \right] \xrightarrow{\text{a.s.}} 0.$$ (D.2)

Similarly, $\Upsilon$ and $\Psi$ can also be reduced for the case $i = g$ as,

$$\Upsilon = \frac{N \rho_{tr}^2}{K \rho_{tr} (1 + \rho_{tr})^2} \xrightarrow{\text{a.s.}} 0.$$ (D.3)

and,

$$\Psi = \left[ \frac{N (S - 1) \rho_{tr}^2}{K S (1 + \rho_{tr})^2} + \frac{N^2 \rho_{tr}^2}{K S (1 + \rho_{tr})^2} \right] \xrightarrow{\text{a.s.}} 0.$$ (D.4)

Substituting (D.3) and (D.4) into (B.4) considering single group, we can further obtain the simplified expression for the interference term. Finally, combining all these terms into (2.12) and dividing the same terms $\frac{N \rho_{tr}^2}{K (1 + \rho_{tr})^2}$ both on numerator and denominator, the proof of the Corollary 3 is completed.
E  Papers Submitted and Under Preparation