On the spectral efficiency for massive MIMO systems with imperfect spacial covariance information

Atchutaram K. Kocharlakota, Student Member, IEEE, Karthik Upadhya, Member, IEEE, and Sergiy A. Vorobyov, Fellow, IEEE

Abstract

This paper studies the impact of imperfect channel covariance information on the uplink and downlink spectral efficiencies of a time division duplexed (TDD) massive multiple-input multiple-output (MIMO) system. Specifically, we derive analytical lower bounds on the uplink and downlink spectral efficiencies of a user with imperfect knowledge on channel vectors, as well as that of the channel covariance matrices, of the users. We consider a linear minimum mean squared estimator (LMMSE) and element-wise LMMSE channel estimation for the unknown channel in this paper. These analytical bounds enable us to choose the sample size for covariance matrix estimation to meet spectral efficiency requirements. The accurate agreement between derived bounds and simulated bounds based on random samples of channel vector and covariance matrices is shown.

Index Terms

Spectral efficiency, massive multiple-input multiple-output (MIMO), covariance estimation, channel estimation, pilot contamination.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) system comprises base stations (BSs) with a large number of antennas (hundreds) to serve multiple users (tens) within each cell. It is considered to be one of the key technologies for the fifth generation (5G) cellular systems due to the improved spectral efficiency (SE) through spatial multiplexing [1]–[5] using low computational complexity [1], [6], [7]. However, knowledge of channel state information (CSI) at the base station (BS) is essential to realize the benefits of a massive MIMO system. Since
the downlink (DL) channel estimation requires a prohibitively large number of pilot resources, the time-division duplexing (TDD), under the assumption of channel reciprocity, is preferred to frequency-division duplexing in massive MIMO, and DL precoding matrices are computed using the channel estimated in the uplink (UL) [2].

It is reasonable to assume that wireless channels are approximately constant for finite time and frequency, termed coherence time and coherence frequency, respectively. Therefore the number of orthogonal pilots that can be used for UL channel estimation is finite and the pilot sequences in UL should be reused by users across the cells, causing the pilot contamination problem [1], [8], [9], with which a matched filter combiner has been shown to impose a ceiling on the throughput [11]. Several pilot decontamination techniques have been studied to mitigate this problem [8], [10]–[14]. However, despite the presence of pilot contamination, the sum rate of the massive MIMO system has been recently proven to be unbounded under the assumption that the individual user spatial covariance matrices are available at the BS [15]. In practice, however these covariance matrices also have to be estimated in the presence of pilot contamination, which is challenging.

Methods for estimating the covariance matrices have been proposed recently in [15]–[17]. In [16], the spatial covariance matrix is decontaminated through supervised/unsupervised clustering, while [15] presents two methods that avoid contamination in the covariance matrices by allocating dedicated orthogonal pilots for each user for covariance estimation. In [17], a new pilot structure and a covariance matrix estimation method are developed that offer higher throughput and lower mean squared error (MSE) of the channel estimates than earlier methods.

In order to compare the performance of the aforementioned methods with large-antenna systems in the presence of pilot contamination, closed-form expressions for the achievable ergodic SE for an arbitrary number of antennas at the BS that also takes the channel estimation error into account prove valuable. Closed-form expressions for uplink (UL) ergodic achievable SE in single and multi-cell massive MIMO systems with various linear receive-combiners designed using the minimum mean-squared error (MMSE) channel estimate have been derived in [18] and [19], respectively. Similar expressions for the achievable SE in the DL have been derived in [8].

However, the closed-form expressions in the aforementioned articles have been derived under the assumption of imperfect CSI and perfect covariance information. While in [20], a lower bound on the ergodic capacity with matched filter receiver combining and with both imperfect
CSI and covariance information has been obtained numerically, closed-form expressions for this bound have not yet been derived, to the best of our knowledge. Such expressions would allow us to characterize the SE in terms of the number of samples required to estimate the covariance matrices.

In this paper, we derive closed-form expressions for the UL and DL spectral efficiencies in a massive MIMO system with imperfect CSI and covariance information at the BS. Since estimated covariance matrices are random, we assume that the code-word spans through multiple realizations of these random matrices, thereby allowing us to characterize the performance of a massive MIMO system with estimated covariance matrices. We consider matched filter receiver combining at the BS which employs either the linear minimum mean squared error (LMMSE) or the element-wise LMMSE channel estimate.

The following are the contributions of this paper

- We have derived closed-form expressions which lower bound the UL and DL spectral efficiency when the LMMSE channel estimate computed using estimated covariance matrices are used in a matched filter combiner.
- Similar expressions have also been derived for the case when the element-wise LMMSE channel estimate is used in a matched filter combiner.
- Expressions that lower bound the UL and DL SE when regularized covariance matrix estimates are used in the element-wise LMMSE channel estimates have also been derived.

The paper is organized as follows: In Section II, we describe the system model along with a detailed explanation on the channel estimation methods and covariance matrices estimation. Section III reports our main derivation in order to obtain closed-form expressions for the UL SE for three different combinations of channel estimation and covariance estimation techniques as well as to obtain closed-form expressions for the DL SE. Section IV provides the simulation results and their comparison with the main results obtained in Section III. We conclude this work in Section V. Technical proofs of lemmas and theorems in the paper appear in appendices at the end of the paper.

**Notation:** We use boldface capital letters for matrices, and boldface lowercase letters for vectors. The superscripts $(\cdot)^*$, $(\cdot)^\top$, and $(\cdot)^H$ denote element-wise conjugate, transpose, and Hermitian transpose operations, respectively. Moreover, $CN(m, R)$ denotes (circularly symmetric)

---

1Some preliminary results will be also reported in [21].
complex Gaussian random vector with mean vector $m$ and covariance matrix $R$, while $W(N, R)$ denotes Wishart random matrix with $N$ degrees of freedom and $R$ is the covariance matrix that corresponds to underlying Gaussian random vectors. The element in $i^{th}$ row and $j^{th}$ column of the matrix $A$ is denoted as $[A]_{ij}$, $I$ stands for an identity matrix (of appropriate size), $\text{diag}(A)$ is a diagonal matrix whose diagonal elements are same as the diagonal elements of the matrix $A$. We use $\text{tr}(\cdot)$ to denote trace of a matrix, $\|\cdot\|$ to denote $l_2$ norm of a vector or a matrix, i.e., Frobenius norm, and $\mathbb{E}\{\cdot\}$ stands for the mathematical expectation. Finally, the symbol $\delta_{ij}$ is the Kronecker delta such that $\delta_{ij} = 1$ if $i = j$, and 0 otherwise.

II. SYSTEM MODEL

A massive MIMO system with $L$ cells each serving $K$ users is considered. Each BS is assumed to have $M$ antennas and each user has a single antenna.

The UL channel between user $(l, k)$ ($k^{th}$ user in $l^{th}$ cell) and BS $j$ is denoted as $h_{jlk} \in \mathbb{C}^M$ and is assumed to be distributed as $\mathcal{C}\mathcal{N}(0, R_{jlk})$. The channel is assumed to be constant for the length of the coherence block, i.e., $\tau_c$ symbols, while its second order statistics are assumed to be constant for $\tau_s$ coherence blocks. TDD mode is considered such that the reciprocity property of the channel can be exploited for DL channel estimation. That is, the DL channel between user $(l, k)$ and BS $j$ is given by $h_{jlk}^*$.

The UL received signal, $Y_j \in \mathbb{C}^{M \times C_u}$, in the $n^{th}$ coherence block at BS $j$ is given by:

$$Y_j[n] = \sum_{l=1}^{L} \sum_{k=1}^{K} \sqrt{\mu} h_{jlk} x_{lk}[n] + N_j[n]$$

where $x_{lk} \in \mathbb{C}^{C_u}$ is the transmitted signal by the user $(l, k)$, and it is assumed to be distributed as $\mathcal{C}\mathcal{N}(0, I)$; $N_j \in \mathbb{C}^{M \times C_u}$ is the additive white Gaussian noise whose elements are identically and independently distributed (i.i.d) as $\mathcal{C}\mathcal{N}(0, 1)$; $\mu$ is the UL transmit power; and $C_u$ is the number of symbols that are dedicated for UL transmission within each coherence block.

The DL received signal, $z \in \mathbb{C}^{C_d}$, in the $n^{th}$ coherence block at user $(j, u)$ is given by:

$$z[n] = \sum_{l=1}^{L} \sum_{k=1}^{K} \sqrt{\lambda} (h_{jlu}^H b_{lk}) d_{lk}[n] + e[n]$$

where $d_{lk} \in \mathbb{C}^{C_d}$ is the DL data from BS $l$ to the user $(l, k)$ and $b_{lk} \in \mathbb{C}^M$ is the corresponding beamforming vector normalized so that $\|b_{lk}\|^2 = 1$; $e \in \mathbb{C}^{C_d}$ is the additive white Gaussian noise distributed as $\mathcal{C}\mathcal{N}(0, I)$; $\lambda$ is the DL transmit power; and $C_d$ is the number of symbols that are dedicated for DL communication within each coherence block.
In the following subsections, pilot structures and estimation techniques for the UL channel and the same for the UL covariance matrices are explained. Since reciprocity property of the channel is assumed for TDD systems, DL channel estimation and DL covariance matrices estimation are avoided.

A. Channel Estimation

A dedicated set of \( P (\geq K) \) symbols is allocated for UL pilots for the channel estimation in each coherence block. Each user in a cell is assumed to choose pilot from a set of \( K \) orthogonal pilots that could be reused across cells. Then \( p_k \in \mathbb{C}^P \) be the pilot used by the \( k^{th} \) user in all the cells such that \( p_k^H p_m = K \delta_{km} \), and \( Y^{(p)}[n] \) is the received signal that corresponds to pilot transmissions in the \( n^{th} \) coherence block and it is given by:

\[
Y^{(p)}[n] = \sum_{l=1}^{L} \sum_{k=1}^{K} \sqrt{\mu} h_{jlk} p_k^* + N_{j}^{(p)}[n]
\]

(1)

where \( N_{j}^{(p)}[n] \) is the noise signal corresponding to the pilot signal.

In what follows, we discuss the channel estimation techniques considered in the rest of the paper.

1) LMMSE channel estimation: The LMMSE estimate of the channel from a target user \((j, u)\) at BS \( j \) in the \( n^{th} \) coherent block is given by:

\[
\hat{h}_{jju}^{LMMSE}[n] = R_{jju} Q_{j}^{-1} \hat{h}_{jju}^{LS}[n]
\]

(2)

where \( \hat{h}_{jju}^{LS}[n] \) is the least squares (LS) channel estimate given by:

\[
\hat{h}_{jju}^{LS}[n] = \frac{1}{K \sqrt{\mu}} Y_j^{(p)}[n] p_u^*
\]

and

\[
Q_{j} = \mathbb{E}\{\hat{h}_{jju}^{LS}[n]\hat{h}_{jju}^{LS}[n]^H\} = \sum_{l=0}^{L-1} R_{jlu} + \frac{1}{K \mu} I.
\]

2) Element-wise LMMSE channel estimation: As computing LMMSE channel estimates is highly complex task, an alternative approach is to use element-wise LMMSE estimates, where only the diagonal elements of the covariance matrices are considered for channel estimation. This technique has the additional advantage that it requires fewer number of samples/pilots for the channel estimation that does not grow with \( M \) [15]. The element-wise LMMSE is given by:

\[
[\hat{h}_{jju}^{el}[n]]_p = \frac{[S_{jju}]_{pp}[\hat{h}_{jju}^{LS}[n]]_p}{[P_{j}]_{pp} [\hat{h}_{jju}^{LS}[n]]_p}, \quad p \in \{1, \ldots, M\}
\]

(3)

where \( S_{jju} \triangleq \text{diag}(R_{jju}) \) and \( P_{j} \triangleq \text{diag}(Q_{j}) \).
3) LMMSE and Element-wise LMMSE channel estimation with imperfect channel covariance matrices: Although the above channel estimates assume that the covariance information is known, it is not available at the receiver. Therefore, it is reasonable to replace these matrices with estimated covariance matrices. Then the LMMSE and element-wise LMMSE channel estimates given in (2) and (3) can be re-written as:

\[
\hat{h}_{jju}[n] = \hat{R}_{jju}^{-1}\hat{h}_{jju,LS}[n] \tag{4}
\]

\[
[\hat{h}_{jju,\text{el}}[n]]_p = \frac{[\hat{S}_{jju}]_{pp}^{-1} \hat{h}_{jju,LS}[n]]_{p}, \quad p \in \{1, \ldots, M\} \tag{5}
\]

where \(\hat{R}_{jju}\) is the estimated covariance matrix, \(\hat{Q}_{jju}\) is the estimate of \(Q_{jju}\), and \(\hat{S}_{jju}\) and \(\hat{P}_{jju}\) are estimates of \(S_{jju}\) and \(P_{jju}\), respectively.

B. Covariance Matrix Estimation

In this subsection, we briefly describe a covariance matrix estimation technique using the pilot structure introduced in [17] for both LMMSE and element-wise LMMSE channel estimation. In this pilot structure, each user transmits two pilot sequences, where the second pilot sequence is generated by introducing random phase shifts to the first one. Furthermore, it is assumed that the BSs know these random phase shifts introduced by the users within the corresponding cells. The covariance matrix of a particular user is estimated using sample cross-correlation of the LS channel estimates from the two pilot sequences.

In addition to the pilot sequence \(p_k\) that is transmitted for channel estimation, another set of pilot sequences \(\{\phi_{lk}[n]\}_{n=1}^{N_l}\) is transmitted by the user \((l, k)\) for estimating \(R_{jlk}\) (or only its diagonal elements if using the element-wise estimator), where the pilot sequence in \(n^{th}\) coherence block is given by \(\phi_{lk}[n] = e^{j\theta_n}p_k\). The random phase sequence \(\{\theta_n\}_{n=0}^{N_l-1}\) is generated by the users in \(l^{th}\) cell, such that it is independent of the channel vectors [17], and \(\mathbb{E}(e^{j\theta_n}) = 0\). Furthermore, the random phase sequences are assumed to be i.i.d across users and cells.

Let \(Y^{(r)}[n]\) be the received signal that corresponds to \(\phi_{ju}[n]\), and it is given by:

\[
Y^{(r)}[n] = \sum_{l=1}^{L} \sum_{k=1}^{K} \sqrt{\mu} h_{jlk}^* \phi_{lk}^T[n] + N^{(r)}_j[n] \tag{6}
\]
where $N_j^{(r)}[n]$ is the noise signal corresponding to $\phi_{jju}[n]$. Additionally, we denote LS channel estimates that correspond to the first ($p_u$) and second ($\phi_{jju}$) pilot sequences as $\hat{h}_{jju}[n]^{(1)}$ and $\hat{h}_{jju}[n]^{(2)}$, respectively. These are given by:

$$\hat{h}_{jju}[n]^{(1)} \triangleq \hat{h}_{jju}[n]^{LS} = h_{jju} + \sum_{l \neq j} h_{jlu} + \frac{\sigma}{K} \sqrt{\mu} N[n] p_u^*$$

$$\hat{h}_{jju}[n]^{(2)} \triangleq \frac{1}{K} Y^{(r)}[n] \sqrt{\mu} e^{-j\theta_j n} p_u^* = h_{jju} + \sum_{l \neq j} h_{jlu} e^{-j\theta_j n} + \frac{\sigma}{K} \sqrt{\mu} N^{(r)}[n] p_u^* e^{-j\theta_j n}.$$  

Above defined LS channel estimates are used for computing sample cross-correlation in order to obtain covariance matrix estimation. In the following subsections, we describe both cases of complete and diagonal matrix estimation.

1) Full matrix estimation for LMMSE channel estimation: The regularized estimate for the covariance matrix is given by:

$$\hat{R}_{jju} = \alpha_R \hat{R}_{jju} + (1 - \alpha_R) R_b$$

where the unbiased estimator $\hat{R}_{jju}$ is given by [17]:

$$\hat{R}_{jju} = \frac{1}{2} \sum_{n=1}^{N_R} \left( \hat{h}_{jju}[n]^{(1)} \left( \hat{h}_{jju}[n]^{(2)} \right)^H + \hat{h}_{jju}[n]^{(2)} \left( \hat{h}_{jju}[n]^{(1)} \right)^H \right).$$

$R_b$ is an arbitrary symmetric positive definite bias-matrix, and $\alpha_R$ is a design parameter. Additionally, it is useful to define $\hat{R}_{jju}$ to denote the expected value of $\hat{R}_{jju}$ given by $\hat{R}_{jju} \triangleq \alpha_R \hat{R}_{jju} + (1 - \alpha_R) R_b$.

The pilot sequence $p_u$ can be reused for estimating $Q_{jju}$. The estimate is given as $\hat{Q}_{jju} = \frac{1}{N_Q} \sum_{n=1}^{N_Q} \hat{h}_{jju}^{LS}[n] (\hat{h}_{jju}^{LS}[n])^H$.

2) Diagonal matrix estimation for element-wise LMMSE channel estimation: For element-wise LMMSE, it is sufficient to estimate the diagonal matrices $S_{jju}$ and $P_{jju}$. The regularized estimate for $S_{jju}$ is given by:

$$\hat{S}_{jju} = \alpha_R \hat{S}_{jju} + (1 - \alpha_R) \text{diag}(R_b)$$

where the unbiased estimator $\hat{S}_{jju}$ is given by:

$$[\hat{S}_{jju}]_{pp} = \frac{1}{2} \sum_{n=1}^{N_R} \left[ \hat{h}_{jju}[n]^{(1)} \right]_p \left[ \hat{h}_{jju}[n]^{(2)} \right]_p^* + \frac{1}{2} \sum_{n=1}^{N_R} \left[ \hat{h}_{jju}[n]^{(2)} \right]_p \left[ \hat{h}_{jju}[n]^{(1)} \right]_p^*, \quad \forall p \in 1 \ldots M.$$
Let us define $\bar{S}_{jju}$ be the expected value of $S_{jju}$ given by $\bar{S}_{jju} \triangleq \alpha_R S_{jju} + (1 - \alpha_R) \text{diag}(R_b)$ for future use. Finally, the diagonal matrix $P_{ju}$ can be estimated as $[\hat{P}_{ju}]_{pp} = \frac{1}{N_Q} \sum_{n=1}^{N_Q} |\hat{h}^{LS}_{jju}[n]|_p^2$, $\forall p \in 1 \ldots M$.

In the following section, the SE for the UL channel of a target user $(j, u)$ is derived. For the derivation, we consider a matched filter receiver combiner, $v_{ju}[n] = \hat{h}_{ju}[n] = \hat{W}_{ju} \hat{h}^{LS}_{ju}[n]$, where $\hat{W}_{ju} = \hat{R}_{jju} \hat{Q}_{ju}^{-1}$ for the LMMSE case and $\hat{W}_{ju} = \hat{S}_{jju} \hat{P}_{ju}^{-1}$ for the element-wise LMMSE case. It is assumed that the covariance matrices and $\hat{h}^{LS}_{ju}[n]$ are uncorrelated within a coherence block $n$, i.e., $\hat{R}_{jju}$ ($\hat{S}_{jju}$) and $\hat{Q}_{ju}$ ($\hat{P}_{ju}$) are computed each from a different set of coherence blocks that does not include $n^{th}$ block. Furthermore, it is assumed that $N_Q > M (>2$).

### III. Main Results: UL and DL Spectral Efficiency

#### A. Uplink Spectral Efficiency

To obtain a lower bound on channel SE, we assume that the codeword is spread over multiple realizations of the covariance estimates. Then, a lower bound on SE of the UL channel from user $(j, u)$ to BS $j$ is given by [20]:

$$r_{ju}^{(ul)} = \left(1 - \frac{K}{C_u} - \frac{N_R K}{C_u \tau_s}\right) \log_2 \left(1 + \gamma_{ju}^{(ul)}\right) \text{[bits/s/Hz]}$$

where $\gamma_{ju}$ is given by

$$\gamma_{ju}^{(ul)} = \frac{|\mathbb{E}\{\text{tr}(\hat{W}_{ju}^H R_{jju})\}|^2}{\mathbb{E}\{\text{tr}(\hat{W}_{ju} Q_{ju} \hat{W}_{ju}^H R_s)\} + \sum_{l=1}^{L} \mathbb{E}\{|\text{tr}(\hat{W}_{ju}^H R_{jlu})|^2\} - |\mathbb{E}\{\text{tr}(\hat{W}_{ju}^H R_{jju})\}|^2}$$

and $R_s \triangleq \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} R_{jlk} + \frac{1}{\mu} \mathbf{I}$. The expectation taken in all the terms of (13) is over random matrix $\hat{W}_{ju}$.

#### B. Uplink Spectral Efficiency for the LMMSE Case

In this subsection, expressions for all the terms given in (13) are derived for the LMMSE channel estimation. The matched filter receiver combiner for the LMMSE channel estimate is given by:

$$\hat{W}_{ju} = \hat{R}_{jju} \hat{Q}_{ju}^{-1}. \quad (14)$$

Before analytically deriving the expectations for the terms in (13), we present some useful lemmas. In what follows, $\mathbb{E}_R$ represents the expectation over $\hat{R}_{jju}$, $\mathbb{E}_Q$ represents the expectation...
over $\hat{Q}_{ju}$, and $\mathbb{E}$ represents the expectation over both. It should be noted that, as already mentioned, we have assumed that $\hat{R}_{jju}$ and $\hat{Q}_{ju}$ are estimated from different pilot resources such that the estimates are independent to each other. Therefore, $\mathbb{E}_R$ and $\mathbb{E}_Q$ can be evaluated independently.

**Lemma 1.** Given an arbitrary matrix $A \in \mathbb{C}^{M \times M}$, and for any mutually independent $M$-dimensional random vector $h$ distributed as $CN(0, R)$, we have

$$
\mathbb{E}\{hh^H Ahh^H\} = RAR + R\text{tr}(AR)
$$

(15)

$$
\mathbb{E}\{|h^H Ah|^2\} = |\text{tr}(A^H R)|^2 + \text{tr}(AR A^H R).
$$

(16)

*Proof.* Proof is available in Appendix A.

**Lemma 2.** Given a Hermitian matrix $C \in \mathbb{C}^{M \times M}$, an arbitrary matrix $A \in \mathbb{C}^{M \times M}$, and a complex Wishart matrix, $X \in \mathbb{C}^{M \times M}$, distributed as $\mathcal{W}(N, I)$, we have

$$
\mathbb{E}\{|X^{-1}|_{ij}\} = \frac{|I|}{N - M}
$$

(17)

$$
\mathbb{E}\{|X^{-1}|_{ij}|X^{-1}|_{lk}\} = \frac{|I|_{ij}|I|_{lk} + \frac{1}{N-M}|I|_{ij}|I|_{lk}}{(N - M)^2 - 1}
$$

(18)

$$
\mathbb{E}\{|\text{tr}(X^{-2} C)|\} = \frac{N}{(N - M)^3 - (N - M)\text{tr}(C)}
$$

(19)

$$
\mathbb{E}\{|\text{tr}(X^{-1} A)|^2\} = \frac{|\text{tr}(A)|^2 + \frac{1}{N-M}\text{tr}(AA^H)}{(N - M)^2 - 1}.
$$

(20)

*Proof.* Proof is available in Appendix B.

**Lemma 3.** Given an arbitrary matrix $A \in \mathbb{C}^{M \times M}$, we have

$$
\mathbb{E}\{\hat{R}_{jju} A \hat{R}_{jju}\} = R_{jju} AR_{jju} + \frac{1}{2N_R} Q_{ju} \text{tr}(AQ_{ju}) + \frac{1}{2N_R} R_{jju} \text{tr}(AR_{jju})
$$

(21)

and

$$
\mathbb{E}\{|\text{tr}(\hat{R}_{jju} A)|^2\} = |\text{tr}(R_{jju} A)|^2 + \frac{1}{2N_R} \text{tr}(AQ_{ju}A^H Q_{ju}) + \frac{1}{2N_R} \text{tr}(AR_{jju}A^H R_{jju}).
$$

(22)

*Proof.* Proof of this lemma uses Lemma 1 and is presented in Appendix C.

Now we are ready to formulate the key theorem of this subsection.

**Theorem 1.** The numerator term of (13), for the case of LMMSE estimation, is given by

$$
\mathbb{E}\{\text{tr}(\hat{W}_{ju}^H R_{jju})\} = \frac{N_Q}{N_Q - M} \text{tr}(\hat{W}_{ju}^H R_{jju}).
$$

(23)
The first and second terms of the denominator in (13) are given by:

\[
\mathbb{E}\{ \text{tr}(\hat{W}_{ju} Q_{ju} \hat{W}_{ju}^H R_s) \} = \kappa_1 \text{tr}(W_{ju} Q_{ju} W_{ju}^H R_s) + \frac{\alpha_R^2 \kappa_1}{2N_R} M \text{tr}(R_s Q_{ju}) \\
+ \frac{\alpha_R^2 \kappa_1}{2N_R} \text{tr}(W_{ju}) \text{tr}(R_s R_{ju})
\] (24)

and

\[
\mathbb{E}\{ |\text{tr}(\hat{W}_{ju}^H R_{ju})|^2 \} = \kappa_2 |\text{tr}(W_{ju}^H R_{ju})|^2 + \frac{\alpha_R^2 \kappa_2}{2N_R} \text{tr}(W_{lu} Q_{ju} W_{lu}^H Q_{ju}) \\
+ \frac{\alpha_R^2 \kappa_1}{2N_R} \text{tr}(W_{lu} R_{ju} W_{lu}^H R_{ju}) + \frac{\kappa_1}{N_Q} \text{tr}(\hat{W}_{ju}^2 Q_{ju}) \\
+ \frac{\alpha_R^2 \kappa_1}{2N_Q N_R} M \text{tr}(W_{ju}^2 Q_{ju}^2) + \frac{\alpha_R^2 \kappa_1}{2N_Q N_R} \text{tr}(W_{ju}) \text{tr}(W_{ju} Q_{ju} R_{ju})
\] (25)

where \( \kappa_1 \triangleq N_Q \kappa_2/(N_Q - M), \kappa_2 \triangleq N^2_Q/((N_Q - M)^2 - 1), W_{ju} \triangleq \hat{R}_{ju} Q_{ju}^{-1} \) and \( W_{lu} \triangleq R_{ju} Q_{ju}^{-1} \).

**Proof.** We define a matrix \( \hat{Q}_{ju} \) as follows:

\[
\hat{Q}_{ju} \triangleq N_Q (Q_{ju}^{-\frac{1}{2}} \delta_{ju} Q_{ju}^{-\frac{1}{2}}).
\] (26)

It can be seen that \( \hat{Q}_{ju} \) is a Wishart matrix distributed as \( \mathcal{W}(N_Q, I) \).

Using (14) and (26), the numerator term of (13) can be written as:

\[
\mathbb{E}\{ \text{tr}(\hat{W}_{ju}^H R_{ju}) \} = N_Q \mathbb{E}\{ \text{tr}(Q_{ju}^{-\frac{1}{2}} \hat{Q}_{ju}^{-\frac{1}{2}} \hat{R}_{ju} R_{ju}) \}.
\] (27)

Taking direct expectation over \( \hat{R}_{ju} \) in (27) and also using Lemma 2, (23) can be obtained.

Proof of (24) and (25) is as follows. Substituting (14) into the first and second terms in the denominator of (13) and using Lemma 2, we get the following equations:

\[
\mathbb{E}\{ \text{tr}(\hat{W}_{ju} Q_{ju} \hat{W}_{ju}^H R_s) \} = \kappa_1 \mathbb{E}_R \{ \text{tr}(Q_{ju}^{-1} \hat{R}_{ju} R_s \hat{R}_{ju}) \} \\
\mathbb{E}\{ |\text{tr}(\hat{W}_{ju}^H R_{ju})|^2 \} = \kappa_2 \mathbb{E}_R \{ |\text{tr}(Q_{ju}^{-1} \hat{R}_{ju} R_{ju})|^2 \} \\
+ \frac{\kappa_1}{N_Q} \mathbb{E}_R \{ \text{tr}(Q_{ju}^{-1} \hat{R}_{ju} R_{ju}^2 \hat{R}_{ju} Q_{ju}^{-1}) \}.
\] (28) (29)

Then using Lemma 3 and substituting (9) into (28) and (29), we get (24) and (25), respectively. \( \square \)
C. Uplink Spectral Efficiency for the Element-Wise LMMSE Case

In this subsection, derivations for all the terms given in (13) for the element-wise LMMSE case are presented. The matched filter receiver combiner for the element-wise LMMSE channel estimate is given by

\[ \hat{W}_{ju} = \hat{S}_{jju} \hat{P}_{ju}^{-1}. \]  

(30)

Before analytically deriving the expectations for the terms in (13), we present some useful lemmas. In what follows, \( E_S \) represents the expectation over \( \hat{S}_{jju} \), \( E_P \) represents the expectation over \( \hat{P}_{ju} \), and \( E \) represents the expectation over both.

**Lemma 4.** Given a zero mean complex Gaussian \( 2 \times 1 \) random vector

\[ h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \]

with covariance matrix

\[ R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \]

then \( E\{|h_1|^2|h_2|^2\} = r_{11}r_{22} + r_{12}r_{21} \).

**Proof.** Proof is available in Appendix D.

**Lemma 5.** Given arbitrary matrices \( A_1 \in \mathbb{C}^{M \times M}, \ A_2 \in \mathbb{C}^{M \times M}, \ A \in \mathbb{C}^{M \times M} \), and a matrix \( Y = Z/2 \), where \( Z \) is a diagonal matrix whose elements are i.i.d \( \chi^2 \) random variables with \( 2N \)-degrees of freedom \( (N > 2) \), we have

\[ E\{\text{tr}(Y^{-1}A_1Y^{-1}A_2)\} = \tau_1 \text{tr}(A_1A_2) + \tau_2 \text{tr}(A_{1d}A_{2d}) \]  

(31)

and

\[ E\{|\text{tr}(Y^{-1}A)|^2\} = \tau_1 |\text{tr}(A)|^2 + \tau_2 \text{tr}(A_d^H A_d) \]  

(32)

where \( \tau_1 \triangleq 1/(N-1)^2 \), \( \tau_2 \triangleq \tau_1/(N-2) \), \( A_{1d} \triangleq \text{diag}(A_1) \), \( A_{2d} \triangleq \text{diag}(A_2) \), and \( A_d \triangleq \text{diag}(A) \).

**Proof.** Proof is available in Appendix E.

**Lemma 6.** Given an arbitrary matrix \( A \in \mathbb{C}^{M \times M} \) and an arbitrary diagonal matrix \( D \in \mathbb{R}^{M \times M} \), then

\[ E\{\hat{S}_{jju}A\hat{S}_{jju}\} = S_{jju}AS_{jju} + \frac{1}{2N}\text{tr}(A \circ Q_{ju} \circ Q_{ju}) + \frac{1}{2N}\text{tr}(A \circ R_{jju} \circ R_{jju}) \]  

(33)
and

\[ E\{|\text{tr}(\mathbf{S}_{ju}\mathbf{D})|^{2}\} = |\text{tr}(\mathbf{S}_{ju}\mathbf{D})|^{2} + \frac{1}{2NR} \sum_{p=1}^{M} \sum_{q=1}^{M} [\text{D}(\mathbf{Q}_{ju} \circ \mathbf{Q}_{ju})\mathbf{D}]_{pq} \]

\[ + \frac{1}{2NR} \sum_{p=1}^{M} \sum_{q=1}^{M} [\text{D}(\mathbf{R}_{ju} \circ \mathbf{R}_{ju})\mathbf{D}]_{pq} \]  (34)

**Proof.** Proof is available in Appendix F.

Now we are ready to formulate the key theorem of this subsection.

**Theorem 2.** The numerator term of (13), for the case of element-wise LMMSE estimation, is given by:

\[ E\{\text{tr}(\mathbf{W}_{ju}^{H}\mathbf{R}_{ju})\} = \frac{N_{Q}}{\left(N_{Q} - 1\right)} \text{tr}(\mathbf{W}_{ju}^{H}\mathbf{R}_{ju}). \]  (35)

The first and second terms of the denominator in (13) are given by:

\[ E\{\text{tr}(\mathbf{W}_{ju}^{H}\mathbf{Q}_{ju}\hat{}^{H}\mathbf{R}_{ju})\} = \kappa_{3}\text{tr}(\mathbf{W}_{ju}^{H}\hat{}^{H}\mathbf{R}_{ju}) + \frac{\alpha_{3}^{2}\kappa_{3}}{2NR} \text{tr}(\mathbf{P}^{-1}_{ju}\mathbf{Q}_{ju}\mathbf{P}^{-1}_{ju}\{\mathbf{R}_{s} \circ \mathbf{Q}_{ju} \circ \mathbf{Q}_{ju}\}) \]

\[ + \frac{\alpha_{3}^{2}\kappa_{3}}{2NR} \text{tr}(\mathbf{P}^{-1}_{ju}\mathbf{Q}_{ju}\mathbf{P}^{-1}_{ju}\{\mathbf{R}_{s} \circ \mathbf{R}_{ju} \circ \mathbf{R}_{ju}\}) + \kappa_{4}\text{tr}(\mathbf{W}_{ju}^{H}\hat{}^{H}\mathbf{W}_{ju}^{H}\mathbf{S}_{s}) \]

\[ + \frac{\alpha_{3}^{2}\kappa_{4}}{2NR} \text{tr}(\mathbf{S}_{s}\mathbf{P}_{ju}) + \frac{\alpha_{3}^{2}\kappa_{4}}{2NR} \text{tr}(\mathbf{W}_{ju}^{H}\mathbf{S}_{s}\mathbf{W}_{ju}) \]  (36)

and

\[ E\{|\text{tr}(\mathbf{W}_{ju}^{H}\mathbf{R}_{ju})|^{2}\} = \kappa_{3}|\text{tr}(\mathbf{W}_{ju}^{H}\hat{}^{H}\mathbf{R}_{ju})|^{2} + \frac{\alpha_{3}^{2}\kappa_{3}}{2NR} \sum_{p,q=1}^{M} [\mathbf{W}_{lu}(\mathbf{Q}_{ju} \circ \mathbf{Q}_{ju})\mathbf{W}_{lu}]_{pq} \]

\[ + \frac{\alpha_{3}^{2}\kappa_{3}}{2NR} \sum_{p,q=1}^{M} [\mathbf{W}_{lu}(\mathbf{R}_{ju} \circ \mathbf{R}_{ju})\mathbf{W}_{lu}]_{pq} + \kappa_{4}\text{tr}(\mathbf{W}_{ju}^{H}\mathbf{S}_{s}^{2}) \]

\[ + \frac{\alpha_{3}^{2}\kappa_{4}}{2NR} \text{tr}(\mathbf{W}_{lu}^{2}\mathbf{W}_{ju}) + \frac{\alpha_{3}^{2}\kappa_{4}}{2NR} \text{tr}(\mathbf{W}_{lu}^{2}\mathbf{S}_{s}^{2}) \]  (37)

where \( \kappa_{3} = N_{Q}^{2}/(N_{Q} - 1)^{2} \), \( \kappa_{4} = \kappa_{3}/(N_{Q} - 2) \), \( \hat{} \mathbf{W}_{ju} \triangleq \hat{}^{S}_{ju} \mathbf{P}_{ju}^{-1} \) and \( \hat{} \mathbf{W}_{lu} \triangleq \hat{}^{S}_{ju} \mathbf{P}_{ju}^{-1} \).

**Proof.** We define a diagonal matrix \( \hat{} \mathbf{P}_{ju} \) as follows:

\[ \hat{} \mathbf{P}_{ju} \triangleq N_{Q}(\mathbf{P}_{ju}^{-1} \hat{} \mathbf{P}_{ju}). \]  (38)

It can be seen that the elements of \( 2\hat{} \mathbf{P}_{ju} \) are i.i.d \( \chi^{2} \) random variables with \( 2N \)-degrees of freedom.
Using (30) and (38), the numerator term of (13) can be written as:

$$E\{\text{tr}(\hat{W}_{ju}^H R_{jju})\} = N_Q E\{\text{tr}(\tilde{P}_{ju}^{-1}\hat{S}_{jju} R_{jju})\}$$

$$= N_Q \sum_{p=1}^{M} E_P\{[\hat{P}_{ju}]_{pp}\} E_S\{[\hat{P}_{ju}]_{pp}^{-1}[\hat{S}_{jju}]_{pp} R_{jju} R_{jju}\}.$$  (39)

Taking direct expectation over $\hat{S}_{jju}$ in (39) and using the properties of inverse $\chi^2$ distribution, (35) can be obtained.

Proof of (36) and (37) is as follows. Substituting (30) and (38) into the first and second denominator terms of (13) and using Lemma 5, we get the following equations

$$E\{\text{tr}(\hat{W}_{ju}^H R_{jju})\} = \kappa_3 E_S\{\text{tr}(\tilde{P}_{ju}^{-1}\hat{S}_{jju} R_{jju})\}$$

$$+ \kappa_4 E_S\{\text{tr}(\tilde{P}_{ju}^{-1}\hat{S}_{jju} S_{jju} R_{jju})\}.$$  (40)

and

$$E\{|\text{tr}(\hat{W}_{ju}^H R_{jju})|^2\} = E\{|\text{tr}(\hat{W}_{ju}^H S_{jju})|^2\}$$

$$= \kappa_3 E_S\{|\text{tr}(\tilde{P}_{ju}^{-1}\hat{S}_{jju} S_{jju})|^2\} + \kappa_4 E_S\{|\text{tr}(\tilde{P}_{ju}^{-2}\hat{S}_{jju}^2 S_{jju})|^2\}.$$  (41)

Then using Lemma 6 and substituting (11) into (40) and (41), we get (36) and (37), respectively.

D. Uplink Spectral Efficiency for the Element-Wise LMMSE Case With Regularized $\tilde{P}_{ju}$

In this section, derivations for all the terms given in (13) for element-wise LMMSE case, with regularized $\tilde{P}_{ju}$, are presented. The regularized estimate of $P_{ju}$ is given by:

$$\hat{P}_{ju} = \alpha_Q \tilde{P}_{ju} + (1 - \alpha_Q) P_b$$  (42)

where $[\hat{P}_{ju}]_{pp} = \frac{1}{N_Q} \sum_{n=1}^{N_Q} |\hat{h}_{ju}^L[n]|^2$, $\forall p \in 1 \ldots M$ is the unbiased estimate of $P_{ju}$; $P_b$ is an arbitrary diagonal bias-matrix with positive elements; and $\alpha_Q$ is a design parameter. Furthermore, let us define the matrix $\hat{P}_{ju} \triangleq N_Q (\tilde{P}_{ju}^{-1} \hat{P}_{ju})$ such that the elements of $2\hat{P}_{ju}$ are $\chi^2$ distributed with $2N_Q$ degrees of freedom. Now, we define two diagonal matrices, $E$ and $G$, whose elements are given by:

$$[E]_{pp} \triangleq E\{[\hat{P}^{-1}]_{pp}\} = E\left\{\left(\frac{1}{N_Q} \alpha_Q [P_{ju}]_{pp} [\hat{P}_{ju}]_{pp} + (1 - \alpha_Q) [P_b]_{pp}\right)^{-1}\right\}$$  (43)

$$[G]_{pp} \triangleq E\{[\hat{P}^{-1}]_{pp}^2\} = E\left\{\left(\frac{1}{N_Q} \alpha_Q [P_{ju}]_{pp} [\hat{P}_{ju}]_{pp} + (1 - \alpha_Q) [P_b]_{pp}\right)^{-2}\right\}.$$  (44)
It should be noted that expectation terms in the above equations can be evaluated numerically using the probability distribution function of $\chi^2$ distribution. The matched filter receiver combiner is given by

$$\hat{W}_{ju} = \hat{S}^{-1}_{ju},$$ \hfill (45) 

Before deriving the expectations for the terms in (13), we present a useful lemma.

**Lemma 7.** Given arbitrary matrices $A_1 \in \mathbb{C}^{M \times M}$, $A_2 \in \mathbb{C}^{M \times M}$, $A \in \mathbb{C}^{M \times M}$, we have

$$\mathbb{E}\{\text{tr}(\hat{P}^{-1}_{ju}A_1\hat{P}^{-1}_{ju}A_2)\} = \text{tr}(EA_1EA_2) + \text{tr}((G - E^2)A_{1d}A_{2d})$$ \hfill (46)

and

$$\mathbb{E}\{|\text{tr}(\hat{P}^{-1}A)|^2\} = |\text{tr}(EA)|^2 + \text{tr}((G - E^2)A_d^HA_d)$$ \hfill (47)

where $A_{1d} \triangleq \text{diag}(A_1)$, $A_{2d} \triangleq \text{diag}(A_2)$, and $A_d \triangleq \text{diag}(A)$.

**Proof.** Proof is available in Appendix G. \hfill \Box

Now we are ready to formulate the key theorem of this subsection.

**Theorem 3.** The numerator term of (13), for the case of element-wise LMMSE estimation, is given by

$$\mathbb{E}\{\text{tr}(\hat{W}_{ju}^HR_{ju})\} = \text{tr}(ES_{ju}R_{ju}).$$ \hfill (48)

The first and second terms of the denominator in (13) are given by

$$\mathbb{E}\{\text{tr}(\hat{W}_{ju}Q_{ju}\hat{W}_{ju}^HR_{s})\} = \text{tr}(EQ_{ju}ES_{ju}R_{s}S_{ju}) + \frac{\alpha_R^2}{2N_R}\text{tr}(EQ_{ju}E\{R_{s} \circ Q_{ju} \circ Q_{ju}\}) + \frac{\alpha_R^2}{2N_R}\text{tr}(G - E^2)P_{ju}^2S_{ju}S_{s})$$ \hfill (49)

and

$$\mathbb{E}\{|\text{tr}(\hat{W}_{ju}^HR_{ju})|^2\} = |\text{tr}(ES_{ju}S_{ju}|^2 + \frac{\alpha_R^2}{2N_R}\sum_{p,q=1}^M[ES_{ju}(Q_{ju} \circ Q_{ju})]_{pq} + \frac{\alpha_R^2}{2N_R}\sum_{p,q=1}^M[ES_{ju}(R_{ju} \circ R_{ju})]_{pq} + (1 + \frac{\alpha_R^2}{2N_R})\text{tr}((G - E^2)S_{ju}^2S_{ju}^2) + \frac{\alpha_R^2}{2N_R}\text{tr}((G - E^2)P_{ju}^2S_{ju}^2).$$ \hfill (50)
Proof. Using (45), the numerator term of (13) can be written as:

$$E\{ \text{tr}(\hat{W}_j^H R_{jju}) \} = E\{ \text{tr}(\hat{P}_{jju}^{-1}\hat{S}_{jju} R_{jju}) \}. $$

(51)

Taking direct expectation over $\hat{S}_{jju}$ in (51) and also using (43), (48) can be obtained.

Proof of (49) and (50) is as follows. Substituting (45) into the first and second denominator terms of (13) and using Lemma 7, we get the following equations

$$E\{ \text{tr}(\hat{W}_j^H \hat{Q}_{jju} \hat{W}_j^H R_{jju}) \} = E_S\{ \text{tr}(EQ_{jju}E\hat{S}_{jju} R_{jju}\hat{S}_{jju}) \}$$

$$+ E_S\{ \text{tr}((G - E^2)P_{jju}\hat{S}_{jju} S_s\hat{S}_{jju}) \} $$

(52)

and

$$E\{ |\text{tr}(\hat{W}_j^H R_{jju})|^2 \} = E\{ |\text{tr}(\hat{W}_j^H S_{jju})|^2 \}$$

$$= E_S\{ |\text{tr}(ES_{jju} S_{jlu})|^2 \} + E_S\{ \text{tr}((G - E^2)S^2_{jju} S^2_{jlu}) \}. $$

(53)

Then using Lemma 6 and substituting (11) into (52) and (53), we get (49) and (50), respectively.

\[ \square \]

E. Downlink Spectral Efficiency

The DL spectral efficiency for user $(j, u)$ is given in this section using the matched filter transmit precoding, that is, $b_{jju} = \hat{W}_{jju}\hat{h}_{jju}^L$.

A lower bound on DL channel SE of user $(j, u)$ is given by:

$$r_{jlu}^{(dl)} = \log_2 \left( 1 + \gamma_{jlu}^{(dl)} \right) \quad [\text{bits/s/Hz}]$$

where $\gamma_{jlu}^{(dl)}$ is given by:

$$\gamma_{jlu}^{(dl)} = \frac{E\{ |\text{tr}(\hat{W}_j^H R_{jju})|^2 \}}{E\{ |\text{tr}(\hat{W}_j^H \hat{Q}_{jju} \hat{W}_j^H R_s^{(dl)})| \} + \sum_{l=1}^{L} E\{ |\text{tr}(\hat{W}_j^H R_{jlu})|^2 \} - |E\{ |\text{tr}(\hat{W}_j^H R_{jju})|^2 \}| \} $$

and $R_s^{(dl)} \triangleq \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} R_{jlk} + \frac{1}{\lambda} I$. The expectation taken in all the terms of (54) is over random matrix $\hat{W}_{jju}$. However, $\hat{W}_{jju} = \hat{R}_{jju}Q_{jju}^{-1}$ for the LMMSE case and $\hat{W}_{jju} = \hat{S}_{jju} \hat{P}_{jju}^{-1}$ for the element-wise LMMSE case. The expectation terms in (54) are already presented in Theorem 1, 2, and 3 for LMMSE, element-wise LMMSE, and element-wise LMMSE with regularized $\hat{P}_{jju}$, respectively. Furthermore, $R_s$ should be replaced by $R_s^{(dl)}$ in computing the expectation terms.
IV. Simulation Results

We consider a massive MIMO system with \( L = 7 \) cells each having \( K = 10 \) users, and the number of antennas at the BS is \( M = 100 \). The BSs are at a distance of 300\( m \) apart from each other and the users are uniformly spread across a circle of radius 120\( m \) from the BS in that cell. Angular spread of the channel cluster is assumed to be 20° within which the received paths from a user are assumed to be uniformly distributed. We consider a single-slope close-in free space reference distance path loss model from [22], and the mean path loss is given by:

\[
P_{L}(f, d) = 20 \log_{10}\left(\frac{4\pi f}{c}\right) + 10n \log_{10}(d)
\]

where \( n \) is the path loss exponent, and \( f \) is the operating frequency being, and \( c \) is the speed of light in \( m/s \). Therefore, the mean received signal to noise ratio (SNR) is given by:

\[
SNR = P_{T} - P_{L} - 10 \log_{10}(kT_{0}B) - NF
\]

where \( P_{T} \) is the transmit power, \( k \) is the Boltzmann constant, \( T_{0} = 290K \) is the nominal temperature and \( B \) is the signal bandwidth. In this setup, we consider \( n = 3.76 \), \( f = 3.4GHz \), \( P_{T} = -3dBm \), \( B = 40MHz \), and \( NF = 10dB \). That is, the mean signal to noise ratio of the received signal from a user that is at the distance \( d \) is given by \( 71.89 - 37.6 \log_{10} d \). The SNR expression as in (55) is used for the DL simulations.

A. Uplink Spectral Efficiency

We compare UL SE of a user for the case of known covariance matrices, SE for simulated system for the case of estimated covariance matrices, and derived theoretical bounds for SE for the case of estimated covariance matrices that correspond to three channel estimation techniques: the LMMSE channel estimation, the element-wise LMMSE channel estimation, and the element-wise LMMSE channel estimation with regularized \( \hat{P}_{ju} \). Fig. 1 shows the SE plots corresponding to the LMMSE channel estimation, Fig. 2 shows the SE plots corresponding to the element-wise LMMSE channel estimation, and Fig. 3 shows the SE plots corresponding to the element-wise LMMSE channel estimation with regularized \( \hat{P}_{ju} \).

Number of symbols that are dedicated for UL transmission within each coherence block is chosen to be \( C_{u} = 100 \) symbols. Second order statistics of the channel are assumed to be constant for \( \tau_{x} = 25000 \) coherence blocks and the UL transmit power is \( \mu = 1 \). Additionally,
Fig. 1: UL SE for the LMMSE channel estimate.

we choose $N_Q = N_R(> M)$, $\alpha_R = 0.95$, and $R_b = I$. Sample averaging for all the expectation terms is computed for 500 trails for different values of $N_R = \{170, 850, 1700, 3400, 4250\}$.

It can be seen from Fig. 1, 2, and 3 that the theoretical UL SEs for the case of unknown covariance matrix asymptotically approach the SE for the known covariance case. Also, SEs for simulated systems match the theoretical values.

**B. Downlink Spectral Efficiency**

We compare DL SE of a user for the case of known covariance matrices, SE for simulated systems for the case of estimated covariance matrices, and derived theoretical bounds on SE for the case of estimated covariance matrices that correspond to three channel estimation techniques: the LMMSE channel estimation, the element-wise LMMSE channel estimation, and the element-wise LMMSE channel estimation with regularized $\hat{P}_{ju}$. Fig. 4 shows the SE plots corresponding to the LMMSE channel estimation, Fig. 5 shows the SE plots corresponding to the element-wise
LMMSE channel estimation, and Fig. 6 shows the SE plots corresponding to the element-wise LMMSE channel estimation with regularized \( \hat{\mathbf{P}}_{ju} \).

Number of symbols that are dedicated for DL transmission within each coherence block is chosen to be \( C_d = 100 \) symbols. Second order statistics of the channel are assumed to be constant for \( \tau_s = 25000 \) coherence blocks and the DL transmit power is \( \lambda = 10 \). Additionally, we choose \( N_Q = N_R(> M) \), \( \alpha_R = 0.95 \), and \( \mathbf{R}_h = \mathbf{I} \). Sample averaging for all the expectation terms is computed for 500 trails for different values of \( N_R = \{170, 850, 1700, 3400, 4250\} \). However, for generating the regularized element-wise LMMSE plot shown in Fig. 6 we choose \( \alpha_Q = 0.95 \) in addition to the aforementioned parameters.

It can be seen from the Fig. 4, 5, and 6 that the theoretical DL SEs for the case of unknown covariance matrix asymptotically approach the SE for the known covariance case. Also, SEs for the simulated system match the theoretical values.
Fig. 3: UL SE for the element-wise LMMSE channel estimate with regularized $\hat{P}_{ju}$.

V. CONCLUSION

We derived analytical expressions for UL and DL SEs of a user in a massive MIMO system with matched filter receiver combiner that uses channel estimates with imperfect covariance information. We first derived the SE bound for the case of LMMSE channel estimates with imperfect covariance information that does not include regularization. Furthermore, we derived the SE bound for the case of element-wise LMMSE channel estimates with imperfect covariance information with and without regularization. We compared the derived analytical bounds with the SEs for perfectly known covariance, and with SEs obtained for a simulated system. Finally, accurate agreement between the derived analytical bounds for the SE and the results from the simulations is demonstrated.

APPENDIX A: PROOF OF LEMMA 1

Let us start with a proof of (15). Using the eigenvalue decomposition of $R = U\Lambda U^H$, let us define $\hat{\Lambda} \in \mathbb{R}^{K \times K}$ to be the diagonal matrix that contains only the non zero diagonal elements of...
Λ (K is rank of R), and \( \tilde{U}^{M \times K} \) to be the matrix that contains K eigenvectors corresponding to the diagonal element of \( \tilde{A} \). Let us also define \( B \triangleq \tilde{U} \tilde{A}^{1/2} \in \mathbb{C}^{M \times N} \). Then, there exists a unique \( g \in \mathbb{C}^N \) such that \( h = Bg \) and \( \mathbb{E}\{gg^H\} = I \). Now, \( \mathbb{E}\{hh^H A hh^H\} = B \mathbb{E}\{gg^H \tilde{A} gg^H\}B^H \) where \( \tilde{A} \triangleq B^H AB \). Since \( g \) is distributed as \( CN(0, I) \), the term \( \mathbb{E}\{gg^H \tilde{A} gg^H\} \) can be evaluated as follows:

\[
\mathbb{E}\{gg^H \tilde{A} gg^H\}_{ij} = \sum_{p=1}^{N} \sum_{q=1}^{N} \mathbb{E}\{[g]_i^*[g]_p [g]_q^*[g]_j\} [\tilde{A}]_{pq}
\]

\[
= \begin{cases} 
[\tilde{A}]_{ij} & \text{if } i \neq j \\
2[\tilde{A}]_{ii} + \sum_{p \neq i} [\tilde{A}]_{pp} & \text{otherwise}
\end{cases}
\]

\[
= \begin{cases} 
[\tilde{A}]_{ij} & \text{if } i \neq j \\
[\tilde{A}]_{ii} + \text{tr}(\tilde{A}) & \text{otherwise}
\end{cases}
\]

and \( \mathbb{E}\{gg^H \tilde{A} gg^H\} = \tilde{A} + \text{tr}(\tilde{A}) \). Therefore, \( \mathbb{E}\{hh^H A hh^H\} = RAR + R\text{tr}(AR) \).
Proof of (16) is as follows:

\[ E\{|h^H Ah|^2\} = E\{h^H A h h^H A^H h\} = E\{\text{tr}(A h h^H A^H h h^H)\}. \]

Using (15), we have

\[ E\{|h^H Ah|^2\} = |\text{tr}(A^H R)|^2 + \text{tr}(A R A^H R). \]

**APPENDIX B: PROOF OF LEMMA 2**

Proof of (17) and (18) is given in [23].

Using the eigenvalue decomposition of \( C = U \Lambda U^H \), let us define \( \bar{X} = U^H X U \). It should be noted that \( \bar{X} \) is distributed as \( \mathcal{W}(N, I) \). Then, (19) can be proved as follows. First, we compute

\[ E\{\text{tr}(X^{-2} C)\} = E\{\text{tr}(\bar{X}^{-2} \Lambda)\} = \sum_{i=1}^{M} [E\{\bar{X}^{-2}\}]_{ii} [\Lambda]_{ii}. \]
But from (18), we can derive that
\[
[E\{\tilde{X}^{-2}\}]_{ii} = \sum_{j=1}^{M} E\{[\tilde{X}^{-1}]_{ij}[\tilde{X}^{-1}]_{ji}\}
\]
\[
= \frac{N}{(N-M)^3 - (N-M)}, \quad \forall i = 1, \ldots M.
\]
Therefore, we finally find that
\[
E\{\text{tr}(X^{-2}C)\} = \sum_{i=1}^{M} \frac{N}{(N-M)^3 - (N-M)} [\Lambda]_{ii}
\]
\[
= \frac{N}{(N-M)^3 - (N-M)} \text{tr}(\Lambda)
\]
\[
= \frac{N}{(N-M)^3 - (N-M)} \text{tr}(C).
\]
For (19), we expand $\mathbb{E}\{\text{tr}(X^{-1}A)^2\}$ using (18) as follows:

$$
\mathbb{E}\{\text{tr}(X^{-1}A)^2\} = \sum_{p=1}^{M} \sum_{q=1}^{M} \sum_{r=1}^{M} \sum_{s=1}^{M} \mathbb{E}\{[X^{-1}]_{pp}[X^{-1}]_{ss}\}[A]_{pp}[A^H]_{ss}
$$

$$
= \sum_{p=1}^{M} \mathbb{E}\{[X^{-1}]_{pp}[X^{-1}]_{pp}\}[A]_{pp}[A^H]_{pp}
$$

$$
+ \sum_{p=1}^{M} \sum_{s=1, s \neq p}^{M} \mathbb{E}\{[X^{-1}]_{pp}[X^{-1}]_{ss}\}[A]_{pp}[A^H]_{ss}
$$

$$
+ \sum_{p=1}^{M} \sum_{s=1, s \neq p}^{M} \mathbb{E}\{[X^{-1}]_{ps}[X^{-1}]_{sp}\}[A]_{sp}[A^H]_{ps}.
$$

Using (18), the above equation can be further simplified to

$$
\mathbb{E}\{\text{tr}(X^{-1}A)^2\} = \frac{1}{(N-M)^2-1} \sum_{p=1}^{M} [A]_{pp}[A^H]_{pp}
$$

$$
+ \frac{1}{(N-M)^2-1} \sum_{p=1}^{M} \sum_{s=1, s \neq p}^{M} [A]_{pp}[A^H]_{ss}
$$

$$
+ \frac{1}{(N-M)^2-1} \sum_{p=1}^{M} \sum_{s=1, s \neq p}^{M} [A]_{sp}[A^H]_{ps}.
$$

Finally, the above equation can be re-written as:

$$
\mathbb{E}\{\text{tr}(X^{-1}A)^2\} = \frac{1}{(N-M)^2-1} \sum_{p=1}^{M} \sum_{s=1}^{M} [A]_{pp}[A^H]_{ss}
$$

$$
+ \frac{1}{(N-M)^2-1} \sum_{p=1}^{M} \sum_{s=1}^{M} [A]_{sp}[A^H]_{ps}
$$

$$
= \frac{|\text{tr}(A)|^2 + \frac{1}{N-M} \text{tr}(AA^H)}{(N-M)^2-1}.
$$

**APPENDIX C: PROOF OF LEMMA 3**

Let us define a pair of mutually independent random vectors as follows:

$$
g_{jju}^{(1)}[n] \triangleq \hat{h}_{jju}^{(1)}[n] - h_{jju}
$$

$$
g_{jju}^{(2)}[n] \triangleq \hat{h}_{jju}^{(2)}[n] - h_{jju}.
$$
The covariance matrices for $g^{(1)}[n]$ and $g^{(2)}[n]$ are identically equal to $Q_{ju} - R_{jju}$. Additionally, we also define mutually independent set of matrices

$$\tilde{R}_{jju}[n] \triangleq \hat{h}^{(1)}_{jju}[n](\hat{h}^{(2)}_{jju}[n])^H + \hat{h}^{(2)}_{jju}[n](\hat{h}^{(1)}_{jju}[n])^H$$

for all $n \in \{1 \ldots N_R\}$ such that $\tilde{R}_{jju} = \frac{1}{N_R} \sum_{n=1}^{N} \tilde{R}_{jju}[n]$ by definition (i.e., (10)).

Using the definition of $g^{(1)}_{jju}[n]$ and $g^{(2)}_{jju}[n]$, and also using Lemma 1, it can be shown that, for all $n = \{1 \ldots N_R\}$, we have

$$\mathbb{E}\{\tilde{R}_{jju}[n]A\tilde{R}_{jju}[n]\} = R_{jju}AR_{jju}$$

and

$$\mathbb{E}\{|\text{tr}(\tilde{R}_{jju}[n]A)|^2\} = |\text{tr}(R_{jju}A)|^2 + \frac{1}{2}\text{tr}(AQ_{ju}A^H) + \frac{1}{2}\text{tr}(AR_{jju}A^HR_{jju}).$$

Finally, along with the equation $\tilde{R}_{jju} = \frac{1}{N_R} \sum_{n=1}^{N} \tilde{R}_{jju}[n]$, (56) and (57) will result in (21) and (22).

**APPENDIX D: PROOF OF LEMMA 4**

Since it is trivial to prove the lemma for rank-1 case, we consider only a full rank $R$. Using the eigenvalue decomposition of $R = U\Lambda U^H$, let us define

$$B \triangleq U\Lambda^{1/2} = \begin{bmatrix} b_1^H \\ b_2^H \end{bmatrix}$$

and a unique

$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

such that $h =Bg$. Therefore, we can find that

$$\mathbb{E}\{|h_1|^2|h_2|^2\} = \mathbb{E}\{|b_1^Hg|^2|b_2^Hg|^2\} = \mathbb{E}\{b_1^Hgg^Hb_1b_2^Hgg^Hb_2\}$$

From Lemma 1 and using the fact that $R = BB^H$ and $\mathbb{E}\{gg^H\} = I$, we have

$$\mathbb{E}\{|h_1|^2|h_2|^2\} = b_1^Hb_1b_2^Hb_2 + b_1^Hb_2b_2^Hb_1 = r_{11}r_{22} + r_{12}r_{21}.$$
Since elements of the diagonal matrix $Z$ are $\chi^2$ distributed with $2N$ degrees of freedom, we have

$$
\mathbb{E}\{[Y^{-1}]_{pp}\} = 2\mathbb{E}\{[Z^{-1}]_{pp}\} = \frac{1}{(N-1)}
$$

$$
\mathbb{E}\{[Y^{-1}]^2\} = 4\mathbb{E}\{[Z^{-1}]^2\} = \frac{1}{(N-1)(N-2)}
$$

Using the above results, (31) can be derived as follows:

$$
\mathbb{E}\{\text{tr}(Y^{-1}A_1Y^{-1}A_2)\} = \left(\frac{1}{N-1}\right)^2 \sum_{p=1}^M \sum_{q\neq p}[A_1]_{pq}[A_2]_{qp}
$$

$$
+ \frac{1}{(N-1)(N-2)} \sum_{p=1}^M [A_1]_{pp}[A_2]_{pp}
$$

$$
= \tau_1 \sum_{q=1}^M \sum_{p=1}^M [A_1]_{pq}[A_2]_{qp} + \tau_2 \sum_{p=1}^M [A_1]_{pp}[A_2]_{pp}
$$

$$
= \tau_1 \text{tr}(A_1A_2) + \tau_2 \text{tr}(A_{1d}A_{2d})
$$

where $\tau_1 \triangleq 1/(N-1)^2$, $\tau_2 \triangleq 1/((N-1)^2(N-2))$, $A_{1d} \triangleq \text{diag}(A_1)$, and $A_{2d} \triangleq \text{diag}(A_2)$.

In what follows, proof of (32) is presented:

$$
\mathbb{E}\{\|\text{tr}(Y^{-1}A)\|^2\} = \frac{1}{(N-1)^2} \sum_{p=1}^M \sum_{q\neq p} [A]_{pp}[A]^*_{qq}
$$

$$
+ \frac{1}{(N-1)(N-2)} \sum_{p=1}^M |[A]_{pp}|^2
$$

$$
= \tau_1 \sum_{q=1}^M \sum_{p=1}^M [A]_{pp}[A]^*_{qq} + \tau_2 \sum_{p=1}^M |[A]_{pp}|^2
$$

$$
= \tau_1 |\text{tr}(A)|^2 + \tau_2 \text{tr}(A_d^H A_d)
$$

where $A_d \triangleq \text{diag}(A)$.

**APPENDIX F: PROOF OF LEMMA 6**

Let us define a pair of mutually independent random vectors as follows:

$$
g_{jju}^{(1)}[n] \triangleq \hat{h}_{jju}^{(1)}[n] - h_{jju}
$$

$$
g_{jju}^{(2)}[n] \triangleq \hat{h}_{jju}^{(2)}[n] - h_{jju}.$$
The covariance matrices for \( g_{jju}^{(1)}[n] \) and \( g_{jju}^{(2)}[n] \) are identically equal to \( Q_{jju} - R_{jju} \). Additionally, we also define mutually independent set of matrices as

\[
\breve{S}_{jju}[n] \triangleq \text{diag}(\hat{h}_{jju}^{(1)}[n] (\hat{h}_{jju}^{(2)}[n])^H + \hat{h}_{jju}^{(2)}[n] (\hat{h}_{jju}^{(1)}[n])^H)
\]

for all \( n \in \{1 \ldots N_R\} \) such that \( \breve{S}_{jju} = \frac{1}{N} \sum_{n=1}^{N} \breve{S}_{jju}[n] \) by definition (i.e., (12)).

Using the definitions of \( g_{jju}^{(1)}[n] \) and \( g_{jju}^{(1)}[n] \) together with Lemma 1 (for scalar case), and Lemma 4, it can be shown that

\[
\mathbb{E}\{[\breve{S}_{jju}]_{pp}[\breve{S}_{jju}]_{qq}\} = \mathbb{E}\{[h_{jju}]_p[h_{jju}]_q\} + \frac{1}{2} [Q_{jju}]_{pq}[Q_{jju}]_{pq} \]

\[
+ \frac{1}{2} [Q_{jju} - R_{jju}]_{pq}[Q_{jju} - R_{jju}]_{pq} + \frac{1}{2} [Q_{jju}]_{pq}[Q_{jju}]_{pq} + \frac{1}{2} [R_{jju}]_{pq}[R_{jju}]_{pq},
\]

Therefore,

\[
\mathbb{E}\{[\breve{S}_{jju}]_{pp}[\breve{S}_{jju}]_{eq}\} = [A]_{pq}\{[S_{jju}]_{pp}[S_{jju}]_{pq} + \frac{1}{2} [Q_{jju}]_{pq}[Q_{jju}]_{pq} \}
\]

\[
+ \frac{1}{2} [R_{jju}]_{pq}[R_{jju}]_{pq}\} \quad (58)
\]

and

\[
\mathbb{E}\{|\text{tr}(S_{jju}D)|^2\} = \sum_{p=1}^{M} \sum_{q=1}^{M} \left\{[S_{jju}]_{pp}[S_{jju}]_{qp} + \frac{1}{2} [Q_{jju}]_{pq}[Q_{jju}]_{pq} \right\}
\]

\[
+ \frac{1}{2} [R_{jju}]_{pq}[R_{jju}]_{pq}\} [D]_{pp}[D]_{qq}
\]

\[
= |\text{tr}(S_{jju}D)|^2 + \frac{1}{2} \sum_{p,q=1}^{M} [D(Q_{jju} \circ Q_{jju})D]_{pq}
\]

\[
+ \frac{1}{2} \sum_{p,q=1}^{M} [D(R_{jju} \circ R_{jju})D]_{pq}, \quad (59)
\]

Finally, along with the equation \( \breve{S}_{jju} = \frac{1}{N} \sum_{n=1}^{N} \breve{S}_{jju}[n] \), (58) and (59) will result in (33) and (34), respectively.
APPENDIX G: PROOF OF LEMMA 7

Expression (46) is derived as follows:

$$
\mathbb{E}\{\text{tr}(\hat{P}^{-1}_{j_u} A_1 \hat{P}^{-1}_{j_u} A_2)\} = \sum_{p=1}^{M} \sum_{q \neq p} \mathbb{E}\{[\hat{P}^{-1}]_{pp}\} \mathbb{E}\{[\hat{P}^{-1}]_{qq}\} [A_1]_{pp}[A_2]_{pp}
$$

$$
+ \sum_{p=1}^{M} \mathbb{E}\{[\hat{P}^{-1}]_{pp}\} [A_1]_{pp}[A_2]_{pp}
$$

$$
= \sum_{p=1}^{M} \sum_{q \neq p} [E]_{pp}[A_1]_{pq}[E]_{qq}[A_2]_{qp} + \sum_{p=1}^{M} [G]_{pp}[A_1]_{pp}[A_2]_{pp}
$$

$$
= \sum_{p=1}^{M} \sum_{q=1}^{M} [E]_{pp}[A_1]_{pq}[E]_{qq}[A_2]_{qp} + \sum_{p=1}^{M} [G - E^2]_{pp}[A_1]_{pp}[A_2]_{pp}
$$

$$
= \text{tr} (EA_1 E A_2) + \text{tr} ((G - E^2) A_{1d} A_{2d})
$$

where $A_{1d} \triangleq \text{diag}(A_1)$ and $A_{2d} \triangleq \text{diag}(A_2)$.

In what follows, proof of (47) is presented:

$$
\mathbb{E}\{|\text{tr}(\hat{P}^{-1} A)|^2\} = \sum_{p=1}^{M} \sum_{q \neq p} \mathbb{E}\{[\hat{P}^{-1}]_{pp}\} \mathbb{E}\{[\hat{P}^{-1}]_{qq}\} [A]_{pp}[A]_{qq}^*
$$

$$
+ \sum_{p=1}^{M} \mathbb{E}\{[\hat{P}^{-1}]_{pp}\} [A]_{pp}^2
$$

$$
= \sum_{p=1}^{M} \sum_{q \neq p} [E]_{pp}[E]_{qq}[A]_{pp}[A]_{qq}^* + \sum_{p=1}^{M} [G]_{pp}[A]_{pp}^2
$$

$$
= \sum_{p=1}^{M} \sum_{q=1}^{M} [E]_{pp}[E]_{qq}[A]_{pp}[A]_{qq}^* + \sum_{p=1}^{M} [G - E^2]_{pp}[A]_{pp}^2
$$

$$
= \text{tr}(EA) + \text{tr} ((G - E^2) A_d^H A_d)
$$

where $A_d \triangleq \text{diag}(A)$.

REFERENCES

[1] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3590–3600, November 2010.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive mimo for next generation wireless systems,” IEEE Communications Magazine, vol. 52, no. 2, pp. 186–195, February 2014.

[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive mimo: Benefits and challenges,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 5, pp. 742–758, Oct 2014.
[23] D. Maiwald and D. Kraus, “On moments of complex wishart and complex inverse wishart distributed matrices,” in IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, April 1997, pp. 3817–3820 vol.5.