ON ACHIEVABLE RATES FOR MASSIVE MIMO SYSTEM WITH IMPERFECT CHANNEL COVARIANCE INFORMATION

Atchutaram K. Kocharlakota[⋆] *Karthik Upadhya*[†] *Sergiy A. Vorobyov*[⋆]

[⋆] Dept. of Signal Processing and Acoustics, Aalto University, Espoo, Finland
[†] Nokia Bell Labs, Espoo, Finland
Emails: kameswara.kocharlakota@aalto.fi, karthik.upadhya[at]nokia-bell-labs[dot]com, svor@ieee.org

ABSTRACT

An analytical lower bound on uplink channel capacity of a user in a massive multiple-input multiple-output system where the channel vector and the covariance matrices of the users in that cell are unknown is derived in this paper. This analytical bound enables us to choose appropriate sample size for covariance matrix estimation to meet the spectral efficiency requirements. The accurate agreement between the derived bound and the simulated bound based on random samples of channel vectors and covariance matrices is shown.

***Index Terms*—** Achievable rate, spectral efficiency, Massive multiple-input multiple-output (MIMO), covariance estimation, channel estimation, pilot contamination.

1. INTRODUCTION

Increased spectral efficiency through spatial multiplexing makes massive multiple-input multiple-output (MIMO) system, where each base station (BS) has large number of antennas to serve multiple users within the cell, one of the key technologies for the next generation mobile networks [1–3]. However, knowledge of channel state information (CSI) at the BS is essential for the communication between the users and the BS.

A limited number of pilots are used for channel estimation due to finite coherence time and finite coherence bandwidth, which results in the pilot contamination problem in multi-cell scenarios [1]. It has been shown recently that, under certain assumptions on the spatial covariance matrices, the sum rate for the massive MIMO system is unbounded despite the presence of pilot contamination [4]. However, this result assumes availability of the individual user covariance matrices at the BS, which, in practice, are also contaminated. Several methods have been proposed in recent literature for estimating the spatial covariance matrices [5–8], and with the estimated channel and covariance information, bounds on the channel capacity are numerically studied in these papers.

However closed form expressions for such bounds utilizing estimated covariance matrices are not available in the literature, to the best of our knowledge. Such bounds can provide useful insights into the number of observations needed for estimating the covariance matrices to achieve a target spectral efficiency (SE).

In this paper, we derive closed form expressions for the SE in a massive MIMO system with imperfect channel and covariance matrix estimates at the BS, thereby characterizing the SE in terms of the number of samples required to estimate the covariance matrices.

2. SYSTEM MODEL

A massive MIMO system with L cells having K users within each of them is considered. All the BSs are assumed to have M antennas each, and all the users have single antenna.

The uplink (UL) channel between the k^{th} user in the l^{th} cell, indexed as (l, k), and a BS j is denoted as $\mathbf{h}_{lk} \in \mathbb{C}^M$ (the subscript j is dropped for the sake of simplicity), and is assumed to be a zero mean circularly symmetric complex Gaussian random vector, denoted as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_{lk})$, where $\mathbf{R}_{lk}$ is the covariance matrix. The channel is assumed to be constant for τ_c symbols –length of the coherence block, while its second order statistics are assumed to be constant for τ_s coherence blocks. C_u symbols are used for UL communication within each coherence block.

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

$$\mathbf{Y}[n] = \sum_{l=1}^{L} \sum_{k=1}^{K} \sqrt{\mu} \mathbf{h}_{lk} \mathbf{x}_{lk}^{\mathsf{T}}[n] + \mathbf{N}[n] \tag{1}$$

where μ is the transmit power at each user, $\mathbf{N} \in \mathbb{C}^{M \times C_u}$ is the additive white Gaussian noise whose elements are distributed as $\mathcal{CN}(0, 1)$, $\mathbf{x}_{lk} \in \mathbb{C}^{C_u}$ is the signal transmitted by user (l, k) whose elements are distributed as $\mathcal{CN}(0, 1)$. It contains the data signal as well as the pilot signals for estimating the channel and the covariance matrices.

In the following subsections, the pilot structures and estimation techniques for the channel vector and the covariance matrices are explained.

2.1. LMMSE Channel Estimation

The BSs and users are assumed to be perfectly synchronized, and K symbols in each coherence block are dedicated for UL channel estimation where each user in a cell is allocated a pilot from a set of K orthogonal sequences. Let $\mathbf{p}_k \in \mathbb{C}^K$, such that $\mathbf{p}_k^H \mathbf{p}_m = K\delta_{km}$, be the pilot transmitted by the k^{th} user in every cell, and $\mathbf{Y}^{(p)}[n] \in \mathbb{C}^{M\times K}$ be the received signal corresponding to pilot transmissions in the n^{th} coherence block.

The linear minimum mean squared error (LMMSE) estimate of the channel from the target user (j,u) in the n^{th} coherent block is given by:

$$\hat{\mathbf{h}}_{ju}[n] = \hat{\mathbf{R}}_{ju}\hat{\mathbf{Q}}_u^{-1}\hat{\mathbf{h}}_{ju}^{LS}[n], \quad n = 1,\ldots,\tau_s \tag{2}$$

where $\hat{\mathbf{R}}_{ju}$ is the estimated covariance matrix, $\hat{\mathbf{Q}}_u$ is an estimate of $\mathbf{Q}_u$, and

$$\hat{\mathbf{h}}_{ju}^{LS}[n] = \frac{1}{K\sqrt{\mu}}\mathbf{Y}^{(p)}[n]\mathbf{p}_u^*, \tag{3}$$

$$\mathbf{Q}_u = \mathbb{E}\{\hat{\mathbf{h}}_{ju}^{LS}(\hat{\mathbf{h}}_{ju}^{LS})^H\} = \sum_{l=0}^{L-1}\mathbf{R}_{lu} + \frac{1}{K\mu}\mathbf{I}. \tag{4}$$

In the following subsection, we describe pilot structure and estimation techniques for these covariance matrices.

2.2. Covariance Matrix Estimation

Here, a covariance matrix estimation technique using the pilot structure introduced in [7] for estimating both $\hat{\mathbf{R}}_{ju}$ and $\hat{\mathbf{Q}}_u$ is described.

An additional set of pilot sequences $\{\boldsymbol{\phi}_{lk}[n]\}_{n=1}^{N_R}$ for estimating $\mathbf{R}_{ju}$ is transmitted by user (l,k). Then $\boldsymbol{\phi}_{lk}[n]$ is the pilot sequence transmitted in n^{th} coherent block, and it is given by $\boldsymbol{\phi}_{lk}[n] = [\bar{\boldsymbol{\phi}}_k^{\intercal}, e^{j\theta_{ln}}\bar{\boldsymbol{\phi}}_k^{\intercal}]^{\intercal} \in \mathbb{C}^{C_r}$ $(2K \le C_r < C_u)$, and $\bar{\boldsymbol{\phi}}_k$ is the sub-sequence used by k^{th} user in all the cells. It is infact a column chosen from $\boldsymbol{\Phi}$ where $\boldsymbol{\Phi}^H\boldsymbol{\Phi} = K\mathbf{I}_{C_r/2}$. Here also $\{\theta_{ln}\}_{n=1}^{N_R}$ is the random phase sequence generated for all the users in l^{th} cell such that it is independent of the channel vectors [7]. Furthermore, it is chosen such that $\mathbb{E}(e^{j\theta_{ln}}) = 0$. These sequences are independently and identically distributed (i.i.d) over different cells, and are assumed to be known at corresponding BS and all the users in the cell.

Regularized covariance matrix $\mathbf{R}_{ju}$ estimate and its expectation are given as follows:

$$\hat{\mathbf{R}}_{ju} \triangleq \alpha\ddot{\mathbf{R}}_{ju} + (1-\alpha)\mathbf{R}_b, \tag{5}$$

where $\ddot{\mathbf{R}}_{ju} = \frac{1}{N_R}\sum_{n=1}^{N_R}\mathrm{sym}\big(\hat{\mathbf{h}}_{ju}^{(1)}[n](\hat{\mathbf{h}}_{ju}^{(2)}[n])^H\big)$, $\mathrm{sym}(\cdot)$ is defined as $\mathrm{sym}(\mathbf{A}) \triangleq \frac{1}{2}(\mathbf{A}+\mathbf{A}^H)$ and

$$\begin{aligned}\hat{\mathbf{h}}_{ju}^{(1)}[n] &= \mathbf{Y}^{(1)}[n]\sqrt{\mu}\bar{\boldsymbol{\phi}}_u^*(\mu\bar{\boldsymbol{\phi}}_u^{\intercal}\bar{\boldsymbol{\phi}}_u^*)^{-1}\\ &= \mathbf{h}_{ju} + \sum_{l\neq j}\mathbf{h}_{lu} + \frac{1}{K\sqrt{\mu}}\mathbf{N}^{(1)}[n]\bar{\boldsymbol{\phi}}_u^*,\end{aligned} \tag{6}$$

$$\begin{aligned}\hat{\mathbf{h}}_{ju}^{(2)}[n] &= \mathbf{Y}^{(2)}[n]\sqrt{\mu}e^{-j\theta_{jn}}\bar{\boldsymbol{\phi}}_u^*(\mu\bar{\boldsymbol{\phi}}_u^{\intercal}\bar{\boldsymbol{\phi}}_u^*)^{-1}\\ &= \mathbf{h}_{ju} + \sum_{l\neq j}\mathbf{h}_{lu}e^{-j\theta_{jn}} + \frac{1}{K\sqrt{\mu}}\mathbf{N}^{(2)}[n]\bar{\boldsymbol{\phi}}_u^*e^{-j\theta_{jn}},\end{aligned} \tag{7}$$

$\mathbf{Y}^{(1)}[n]$ and $\mathbf{Y}^{(2)}[n]$ are the received signals that correspond to the first and second sub-sequences of ϕ, respectively, $\mathbf{N}^{(1)}[n]$ and $\mathbf{N}^{(2)}[n]$ are the noise signals that are additive to first and second pilot sub-sequence, respectively, $\mathbf{R}_b$ is an arbitrary symmetric positive definite bias-matrix, and α is a design parameter. For later use, it is useful to define $\bar{\mathbf{R}}_{ju} \triangleq \mathbb{E}\{\hat{\mathbf{R}}_{ju}\} = \alpha\mathbf{R}_{ju} + (1-\alpha)\mathbf{R}_b$.

For estimating $\mathbf{Q}_u$, an additional set of pilots is not required, but the received signal that corresponds to pilot sequence $\mathbf{p}_u$ can be utilized to compute an unbiased estimator of $\mathbf{Q}_u$, which is given as follows:

$$\hat{\mathbf{Q}}_u = \frac{1}{N_Q}\sum_{n=1}^{N_Q}\hat{\mathbf{h}}_{ju}^{LS}[n](\hat{\mathbf{h}}_{ju}^{LS}[n])^H. \tag{8}$$

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

3. UL SPECTRAL EFFICIENCY

To obtain a lower bound on the channel capacity, we assume that the codeword is spread over multiple realizations of the covariance estimates. Then, a lower bound on capacity of the UL channel from user (j,u) to BS j is given by [6]:

$$R_{ju} = \left(1 - \frac{K}{C_u} - \frac{N_R C_r}{C_u\tau_s}\right)\log_2(1+\gamma_u) \quad [bits/s/Hz] \tag{9}$$

where γ_u is given in (10) at the top of the next page and $\mathbf{R}_s \triangleq \sum_{l=0}^{L-1}\sum_{k=0}^{K-1}\mathbf{R}_{lk} + \frac{1}{\mu}\mathbf{I}$. The expectation taken in all the terms is over the random matrix $\hat{\mathbf{W}}_{ju}$.

Before deriving the expectation terms of (10), we give lemmas that will be useful in the derivation. In what follows, $\mathbb{E}_R$ represents the expectation over $\hat{\mathbf{R}}_{ju}$, $\mathbb{E}_Q$ represents expectation over $\hat{\mathbf{Q}}_u$, and $\mathbb{E}$ represents expectation over both.

Lemma 1. *Given an arbitrary matrix $\mathbf{A} \in \mathbb{C}^{M\times M}$, and for any mutually independent M-dimensional random vectors $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}$ distributed as $\mathcal{CN}(\mathbf{0},\mathbf{R}_1)$, $\mathcal{CN}(\mathbf{0},\mathbf{R}_2)$, and $\mathcal{CN}(\mathbf{0},\mathbf{R})$, respectively, we have*

$$\mathbb{E}\{\mathbf{h}_1\mathbf{h}_2^H\mathbf{A}\mathbf{h}_2\mathbf{h}_1^H\} = \mathbf{R}_1\mathrm{tr}(\mathbf{A}\mathbf{R}_2), \tag{13}$$

$$\mathbb{E}\{\mathbf{h}\mathbf{h}^H\mathbf{A}\mathbf{h}\mathbf{h}^H\} = \mathbf{R}\mathbf{A}\mathbf{R} + \mathbf{R}\mathrm{tr}(\mathbf{A}\mathbf{R}). \tag{14}$$

$$\gamma_u = \frac{|\mathbb{E}\{\mathrm{tr}(\hat{\mathbf{W}}_{ju}^H \mathbf{R}_{ju})\}|^2}{\mathbb{E}\{\mathrm{tr}(\hat{\mathbf{W}}_{ju}\mathbf{Q}_u\hat{\mathbf{W}}_{ju}^H\mathbf{R}_s)\} + \sum_{l=1}^{L}\mathbb{E}\{|\mathrm{tr}(\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{lu})|^2\} - |\mathbb{E}\{\mathrm{tr}(\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{ju})\}|^2} \tag{10}$$

$$\mathbb{E}\{\mathrm{tr}(\hat{\mathbf{W}}_{ju}\mathbf{Q}_u\hat{\mathbf{W}}_{ju}^H\mathbf{R}_s)\} = \kappa_1\mathrm{tr}(\bar{\mathbf{W}}_{ju}\mathbf{Q}_u\bar{\mathbf{W}}_{ju}^H\mathbf{R}_s) + \frac{\alpha^2\kappa_1}{2N_R}\left\{M\mathrm{tr}(\mathbf{R}_s\mathbf{Q}_u) + \mathrm{tr}(\mathbf{W}_{ju})\mathrm{tr}(\mathbf{R}_s\mathbf{R}_{ju})\right\} \tag{11}$$

$$\begin{aligned}\mathbb{E}\{|\mathrm{tr}(\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{lu})|^2\} &= \kappa_2|\mathrm{tr}(\bar{\mathbf{W}}_{ju}^H\mathbf{R}_{lu})|^2 + \frac{\alpha^2\kappa_2}{2N_R}\mathrm{tr}(\mathbf{W}_{lu}\mathbf{Q}_u\mathbf{W}_{lu}^H\mathbf{Q}_u) + \frac{\alpha^2\kappa_2}{2N_R}\mathrm{tr}(\mathbf{W}_{lu}\mathbf{R}_{ju}\mathbf{W}_{lu}^H\mathbf{R}_{ju}) \\ &+ \frac{\kappa_1}{N_Q}\mathrm{tr}(\bar{\mathbf{W}}_{ju}^2\mathbf{Q}_u\mathbf{W}_{lu}^2\mathbf{Q}_u) + \frac{\alpha^2\kappa_1}{2N_QN_R}M\mathrm{tr}(\mathbf{W}_{lu}^2\mathbf{Q}_u^2) + \frac{\alpha^2\kappa_1}{2N_QN_R}\mathrm{tr}(\mathbf{W}_{ju})\mathrm{tr}(\mathbf{W}_{lu}^2\mathbf{Q}_u\mathbf{R}_{ju})\end{aligned} \tag{12}$$

Proof. (13) can be derived by splitting the expectation over $\mathbf{h}_1$ and $\mathbf{h}_2$. Proof of (14) is straightforward, but it involves the values of second and fourth order moments of Gaussian random variables. Thus, it is left to the reader because of space limitation. □

Lemma 2. *Given a Hermitian matrix $\mathbf{C} \in \mathbb{C}^{M\times M}$, an arbitrary matrix $\mathbf{A} \in \mathbb{C}^{M\times M}$, and a complex Wishart matrix, $\mathbf{X} \in \mathbb{C}^{M\times M}$, with N degrees of freedom (represented as $\mathcal{CW}(N,\mathbf{I})$), we have*

$$\mathbb{E}\{[\mathbf{X}^{-1}]_{ij}\} = \frac{[\mathbf{I}]_{ij}}{N-M}, \tag{15}$$

$$\mathbb{E}\{[\mathbf{X}^{-1}]_{ij}[\mathbf{X}^{-1}]_{lk}\} = \frac{[\mathbf{I}]_{ij}[\mathbf{I}]_{lk} + \frac{1}{N-M}[\mathbf{I}]_{lj}[\mathbf{I}]_{ik}}{(N-M)^2-1}, \tag{16}$$

$$\mathbb{E}\{\mathrm{tr}(\mathbf{X}^{-2}\mathbf{C})\} = \frac{N}{(N-M)^3-(N-M)}\mathrm{tr}(\mathbf{C}), \tag{17}$$

$$\mathbb{E}\{|\mathrm{tr}(\mathbf{X}^{-1}\mathbf{A})|^2\} = \frac{|\mathrm{tr}(\mathbf{A})|^2 + \frac{1}{N-M}\mathrm{tr}(\mathbf{A}\mathbf{A}^H)}{(N-M)^2-1}. \tag{18}$$

Proof. Proof is available in Appendix. □

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

$$\begin{aligned}\mathbb{E}\{\ddot{\mathbf{R}}_{ju}\mathbf{A}\ddot{\mathbf{R}}_{ju}\} &= \mathbf{R}_{ju}\mathbf{A}\mathbf{R}_{ju} + \frac{1}{2N_R}\mathbf{Q}_u\mathrm{tr}(\mathbf{A}\mathbf{Q}_u) \\ &+ \frac{1}{2N_R}\mathbf{R}_{ju}\mathrm{tr}(\mathbf{A}\mathbf{R}_{ju})\end{aligned} \tag{19}$$

and

$$\begin{aligned}\mathbb{E}\{|\mathrm{tr}(\ddot{\mathbf{R}}_{ju}\mathbf{A})|^2\} &= |\mathrm{tr}(\mathbf{R}_{ju}\mathbf{A})|^2 + \frac{1}{2N_R}\mathrm{tr}(\mathbf{A}\mathbf{Q}_u\mathbf{A}^H\mathbf{Q}_u) \\ &+ \frac{1}{2N_R}\mathrm{tr}(\mathbf{A}\mathbf{R}_{ju}\mathbf{A}^H\mathbf{R}_{ju})\end{aligned} \tag{20}$$

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

Now we are ready to formulate the key theorem.

Theorem 1. *The signal component of* (10) *is given by*

$$\mathbb{E}\mathrm{tr}\{\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{ju}\} = \frac{N_Q}{N_Q-M}\mathrm{tr}(\bar{\mathbf{W}}_{ju}^H\mathbf{R}_{ju}). \tag{21}$$

The first and second terms of the denominator in (10) *are given in* (11) *and* (12) *at the top of this page, where $\kappa_1 = N_Q\kappa_2/(N_Q-M)$, $\kappa_2 = N_Q^2/((N_Q-M)^2-1)$, $\bar{\mathbf{W}}_{ju} \triangleq \bar{\mathbf{R}}_{ju}\mathbf{Q}_u^{-1}$, and $\mathbf{W}_{lu} = \mathbf{R}_{lu}\mathbf{Q}_u^{-1}$.*

Proof. We define a matrix $\tilde{\mathbf{Q}}_{ju}$ as follows:

$$\tilde{\mathbf{Q}}_{ju} \triangleq N_Q(\mathbf{Q}_{ju}^{-\frac{1}{2}}\hat{\mathbf{Q}}_u\mathbf{Q}_{ju}^{-\frac{1}{2}}). \tag{22}$$

It can be seen that $\tilde{\mathbf{Q}}_{ju}$ is a Wishart matrix distributed as $\mathcal{W}(N_Q,\mathbf{I})$. Using $\hat{\mathbf{W}}_{ju} = \hat{\mathbf{R}}_{ju}\hat{\mathbf{Q}}_u^{-1}$ and (22), the numerator term of (10) can be written as:

$$\mathbb{E}\mathrm{tr}\{\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{ju}\} = N_Q\mathbb{E}\mathrm{tr}\{\mathbf{Q}_{ju}^{-\frac{1}{2}}\tilde{\mathbf{Q}}_{ju}^{-1}\mathbf{Q}_{ju}^{-\frac{1}{2}}\hat{\mathbf{R}}_{ju}\mathbf{R}_{ju}\}. \tag{23}$$

By taking direct expectation over $\hat{\mathbf{R}}_{ju}$ in (23) and also using Lemma 2, (21) can be obtained.

Proof of (11) and (12) is as follows. By substituting $\hat{\mathbf{W}}_{ju} = \hat{\mathbf{R}}_{ju}\hat{\mathbf{Q}}_u^{-1}$ and (22) into the first and second denominator terms of (10) and by using Lemma 2, we get the following equations

$$\mathbb{E}\mathrm{tr}\{\hat{\mathbf{W}}_{ju}\mathbf{Q}_u\hat{\mathbf{W}}_{ju}^H\mathbf{R}_s\} = \kappa_1\mathbb{E}_R\mathrm{tr}\{\mathbf{Q}_u^{-1}\hat{\mathbf{R}}_{ju}\mathbf{R}_s\hat{\mathbf{R}}_{ju}\}, \tag{24}$$

$$\begin{aligned}\mathbb{E}\{|\mathrm{tr}(\hat{\mathbf{W}}_{ju}^H\mathbf{R}_{lu})|^2\} &= \kappa_2\mathbb{E}_R\{|\mathrm{tr}(\mathbf{Q}_u^{-1}\hat{\mathbf{R}}_{ju}\mathbf{R}_{lu})|^2\} \\ &+ \frac{\kappa_1}{N_Q}\mathbb{E}_R\mathrm{tr}\{\mathbf{Q}_u^{-1}\hat{\mathbf{R}}_{ju}\mathbf{R}_{lu}\mathbf{Q}_u^{-1}\mathbf{R}_{lu}\hat{\mathbf{R}}_{ju}\}\end{aligned} \tag{25}$$

By using Lemma 3, and by substituting (5) into (24) and (25), we get (11) and (12), respectively. □

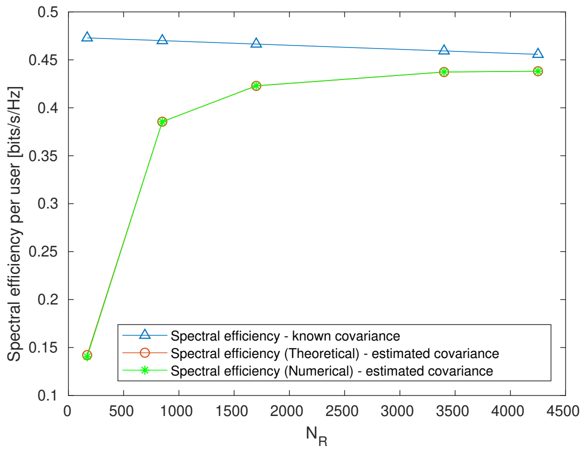

Fig. 1. UL SE of a user in massive MIMO system

4. SIMULATIONS

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 separated $300m$ and the users are uniformly located in a circle of radius $120m$ from the BS. The signal to noise ratio of the received signal from a user that is at a distance d is given by $78.6 - 37.6\log_{10} d$. In Fig 1, we compare SE of a user in the central cell for the case of known covariance matrices, theoretical lower bound on channel capacity for the estimated covariance case and simulated SE for the case of estimated covariance matrices.

Number of UL resources used in a coherence block is chosen to be $C_u = 100$ symbols, and second order statistics are assumed to be constant for $\tau_s = 25000$ coherence blocks. The transmit power of the target is $\mu = 1$. Additionally, we choose $C_r = 2K$ symbols, $N_Q = N_R\ (> M)$, $\alpha = 0.95$, and $\mathbf{R}_b = \mathbf{I}$. Sample averaging for all the expectation terms is computed for 500 iterations for different values of $N_R = (170, 850, 1700, 3400, 4250)$.

It can be seen from Fig 1 that the theoretical achievable rate for the case of unknown covariance matrix asymptotically approaches the SE for the known covariance case. Also, the simulated SE matches the theoretical values.

5. CONCLUSION

An analytical expression for the SE of a user in a massive MIMO system is derived for the case when the matched filter receiver combiner uses estimated channel covariance matrices and estimated channel vector. The simulation results matched that obtained by the theoretical expression derived in this paper. The accurate agreement between the derived analytical bound for the SE and the results of simulations is demonstrated.

6. APPENDIX

Proof of Lemma 2. Proofs of (15) and (16) are given in [9].

Using eigen value decomposition $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ and denoting $\tilde{\mathbf{X}} = \mathbf{U}^H\mathbf{X}\mathbf{U} \sim \mathcal{W}(N, \mathbf{I})$, (17) can be proved as follows:

$$\mathbb{E}\mathrm{tr}\{\mathbf{X}^{-2}\mathbf{C}\} = \mathbb{E}\mathrm{tr}\{\tilde{\mathbf{X}}^{-2}\mathbf{\Lambda}\} = \sum_{i=1}^{M}[\mathbb{E}\{\tilde{\mathbf{X}}^{-2}\}]_{ii}[\mathbf{\Lambda}]_{ii}$$
$$= [\mathbb{E}\{\tilde{\mathbf{X}}^{-2}\}]_{11}\mathrm{tr}(\mathbf{\Lambda}) = \sum_{j=1}^{M}\mathbb{E}\{[\tilde{\mathbf{X}}^{-1}]_{1j}[\tilde{\mathbf{X}}^{-1}]_{j1}\}\mathrm{tr}(\mathbf{C})$$
$$= \frac{N}{(N-M)^3 - (N-M)}\mathrm{tr}(\mathbf{C}).$$

The above derivation uses the fact that $[\mathbb{E}\{\tilde{\mathbf{X}}^{-2}\}]_{ii}$ is same for all $i \in \{1 \dots M\}$.

For (18), $\mathbb{E}\{|\mathrm{tr}(\mathbf{X}^{-1}\mathbf{A})|^2\}$ is expanded as follows:

$$\mathbb{E}\{|\mathrm{tr}(\mathbf{X}^{-1}\mathbf{A})|^2\} = \sum_{p,q,r,s=1}^{M}\mathbb{E}\{[\mathbf{X}^{-1}]_{pq}[\mathbf{X}^{-1}]_{sr}\}[\mathbf{A}]_{qp}[\mathbf{A}^H]_{rs}$$
$$= \sum_{p=1}^{M}\sum_{s=1}^{M}\mathbb{E}\{[\mathbf{X}^{-1}]_{pp}[\mathbf{X}^{-1}]_{ss}\}[\mathbf{A}]_{pp}[\mathbf{A}^H]_{ss}$$
$$+ \sum_{p=1}^{M}\sum_{s=1}^{M}\mathbb{E}\{[\mathbf{X}^{-1}]_{ps}[\mathbf{X}^{-1}]_{sp}\}[\mathbf{A}]_{sp}[\mathbf{A}^H]_{ps}.$$

Using (16), the above equation can be re-written as in (18). □

Proof of Lemma 3. Let us define a pair of mutually independent random vectors $\mathbf{g}^{(1)}_{jju}[n] \triangleq \hat{\mathbf{h}}^{(1)}_{jju}[n] - \mathbf{h}_{ju}$ and $\mathbf{g}^{(2)}_{jju}[n] \triangleq \hat{\mathbf{h}}^{(2)}_{jju}[n] - \mathbf{h}_{ju}$. Their covariance matrices are identically equal to $\mathbf{Q}_u - \mathbf{R}_{ju}$. Additionally, we also define mutually independent set of matrices as $\breve{\mathbf{R}}_{ju}[n] \triangleq \mathrm{sym}(\hat{\mathbf{h}}^{(1)}_{ju}[n](\hat{\mathbf{h}}^{(2)}_{ju}[n])^H)$, $\forall n \in \{1, \dots N_R\}$ such that $\ddot{\mathbf{R}}_{ju} = \frac{1}{N_R}\sum_{n=1}^{N_R}\breve{\mathbf{R}}_{ju}[n]$.

Using the definitions of $\mathbf{g}^{(1)}_{jju}[n]$ and $\mathbf{g}^{(2)}_{jju}[n]$, and also Lemma 1, it can be shown that

$$\mathbb{E}\{\breve{\mathbf{R}}_{ju}[n]\mathbf{A}\breve{\mathbf{R}}_{ju}[n]\} = \mathbf{R}_{ju}\mathbf{A}\mathbf{R}_{ju} + \frac{1}{2}\mathbf{Q}_u\mathrm{tr}(\mathbf{A}\mathbf{Q}_u)$$
$$+ \frac{1}{2}\mathbf{R}_{ju}\mathrm{tr}(\mathbf{A}\mathbf{R}_{ju}), \quad \forall n = 1 \text{ to } N_R, \tag{26}$$

and

$$\mathbb{E}\{|\mathrm{tr}(\breve{\mathbf{R}}_{ju}[n]\mathbf{A})|^2\} = |\mathrm{tr}(\mathbf{R}_{ju}\mathbf{A})|^2 + \frac{1}{2}\mathrm{tr}(\mathbf{A}\mathbf{Q}_u\mathbf{A}^H\mathbf{Q}_u)$$
$$+ \frac{1}{2}\mathrm{tr}(\mathbf{A}\mathbf{R}_{ju}\mathbf{A}^H\mathbf{R}_{ju}), \quad \forall n = 1 \text{ to } N_R. \tag{27}$$

Finally, along with the equation $\ddot{\mathbf{R}}_{ju} = \frac{1}{N_R}\sum_{n=1}^{N_R}\breve{\mathbf{R}}_{ju}[n]$, (26) and (27) will result in (19) and (20), respectively. □

7. 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.

[4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive mimo has unlimited capacity," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 574–590, Jan 2018.

[5] S. Haghighatshoar and G. Caire, "Massive mimo pilot decontamination and channel interpolation via wideband sparse channel estimation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8316–8332, Dec 2017.

[6] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive mimo with imperfect channel covariance information," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 974–978.

[7] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive mimo," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 546–550, April 2018.

[8] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive mimo," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 863–867, June 2018.

[9] D. Maiwald and D. Kraus, "On moments of complex wishart and complex inverse wishart distributed matrices," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 1997, vol. 5, pp. 3817–3820 vol.5.