Coherent MU-MIMO in Block Fading Channels: A Finite Blocklength Analysis

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Coherent MU-MIMO in Block Fading Channels: A Finite Blocklength Analysis

Junjuan Feng, Hien Quoc Ngo, and Michail Matthaiou

Institute of Electronics, Communications and Information Technology (ECIT)

Queen's University Belfast, BT3 9DT, Belfast, U.K.

Email:{junjuan.feng, hien.ngo, m.matthaiou}@qub.ac.uk

Abstract—This paper investigates the maximum coding rate of the multiple-user multiple-input-multiple-output (MU-MIMO) uplink in coherent block-fading channels and with finite blocklength. The backoff of the maximum coding rate from the capacity caused by finite blocklength is precisely characterized by a parameter, called the channel dispersion. In particular, we derive exact analytical and approximation results for a large number of the base station (BS) antennas. By analyzing these results, we observe that even when considering the channel dispersion, the maximum coding rate still increases with respect to the number of BS antennas, whilst the SNR for each user can also improve the performance to a ceiling limited by the inter-user interference. Fast channel dynamics (shorter coherence time) and high diversity gain (large number of blocks) are beneficial for the maximum coding rate under finite blocklength and coherent setting. Moreover, to obtain a certain fraction of the capacity with fixed error probability, the minimum required blocklength (delay) can be reduced by increasing the number of BS antennas. We also show that the channel dispersion will converge to a constant while the minimum required blocklength can approach to zero with massive number of the BS antennas. Hence, from a theoretical viewpoint, deploying a large number of the BS antennas is beneficial for low latency communications.

Index Terms—Channel dispersion, coherent block-fading channel, finite blocklength, maximum coding rate, multiple-user MIMO.

I. INTRODUCTION

The future Internet-of-Things (IoT), ranging from machine-type communications to mission-critical communications, entails new requirements, such as higher spectral efficiency, ultra-higher reliabilities and low latency. The future 5G networks and beyond are expected to support ultra reliable low latency communication (URLLC) with latency going from less than 1 ms to few milliseconds and reliability higher than 99.999% [1] [2]. Thus, providing URLLC is important, challenging and has attracted lot of research attention over the past decade. If it is not necessary to consider latency constraints, the conventional Shannon capacity is sufficient enough for the performance analysis. However, when low latency communication is required, the assumption of long codewords becomes irrelevant and we need to resort to short codewords in order to fulfill the stringent latency constraints [3]. Having said this, finite blocklength systems have come at the forefront of wireless

communications research, particularly after the pioneering work of [4]. Most importantly, Shannon capacity yields an overestimation of the performance, and another refined performance metric, *the maximum coding rate*, which is as a function of finite blocklength and coding error probability, becomes mostly relevant.

A. Related Works

Polyanskiy, Poor and Verdú [4] showed that with an error probability no larger than $0 < \epsilon < 1$ and under finite blocklength n, the normal approximation of the maximum coding rate under real additive white Gaussian noise (AWGN) channel (no fading) with unit variance can be expressed as

$$R^*(n, \epsilon) = C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) + \mathcal{O}\left(\frac{\log_2 n}{n}\right), \quad (1)$$

where $Q^{-1}(\epsilon)$ is the inverse Q-function, $C = \frac{1}{2}\log_2(1+\rho)$ is the capacity, and $V = \frac{\rho(\rho+2)}{2(1+\rho)^2}\log_2^2 e$ denotes the channel dispersion with the definition as [4, Definition 1], in which ρ is the SNR. The term $\mathcal{O}\left(\frac{\log_2 n}{n}\right)$ represents the remaining terms that vary with the same speed as $\frac{\log_2 n}{n}$.

Following [4], a series of works have extended the result (1) to other kinds of point-to-point channels. For instance, [5] studied the channel dispersion under additive non-Gaussian noise channels by using random Gaussian codebooks and nearest neighbor decoding. For the generalization to fading channels, [6] derived the dispersion of a single-input single-output (SISO), stationary coherent fading channel with real valued AWGN. The result showed the significant effect of the channel fading dynamics on the dispersion. Quasi-static single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) fading channels under finite block-length with unit variance AWGN were thoroughly investigated in [7] and [8], respectively. The results showed that under mild conditions on the channel gain, the channel dispersion is zero regardless whether the channel state information (CSI) is available at transmitter and/or receiver. Regarding the block-fading channel, [9]–[12] studied the performance metrics from SISO to MIMO systems with unit variance AWGN. Specifically, in [9], the non-asymptotic bounds of the maximum coding rate were presented in the noncoherent setting for SISO systems. It was shown that the maximum coding rate is not monotonic with respect to the channel coherence time and there exists a coherence time that maximizes the coding rate and makes a trade-off between the diversity gain obtained by a large number of blocks and the

resources used for channel estimation. Following this, a high-SNR normal approximation of the maximum coding rate was presented in [10] under noncoherent SISO Rayleigh block-fading channels, which complements the non-asymptotic bounds in [9] and provides us with a tractable formula for performance analysis. Very recently, [11] and [12] have provided a channel dispersion formula for AWGN coherent MIMO block-fading channel and the analysis showed that as we increase the number of receive antennas, the normalized dispersion decreases and the exact dispersion will approach to a constant. Also, [12] pointed out that the transmit and the receive antennas are not symmetric in affecting the channel dispersion.

B. Contributions

In this paper, we will focus on block-fading channels, in which the channel coefficient remains constant during one coherence block and varies independently to a new realization in the next block [9]. We also assume coherent communication, which means that the BS has perfect CSI. Different from [4]–[12], which focused on point-to-point networks, we take a step further by exploiting the multiple-user uplink networks, in which multiple users transmit signals to a multiple-antenna BS. Thus, the interference from other users must be considered, which is actually not Gaussian. A more refined performance metric, the maximum coding rate under finite blocklength and non-Gaussian noise is presented and analyzed. The results enable us to obtain some important insights, which are listed as follows:

1) The new maximum coding rate under finite blocklength is an increasing function with respect to the number of BS antennas and the signal-to-noise-ratio (SNR) of each user boosts the performance to a ceiling determined by the inter-user interference;
2) Fast fading channel dynamics (shorter coherence time) and high diversity gain (large number of blocks) can strongly improve the performance even under the inter-user interference, which has been proved in [12] for point-to-point MIMO systems;
3) The minimum blocklength required for obtaining a fixed fraction of capacity with given error probability reduces when a massive number of antennas is deployed at the BS. In addition, longer blocklength is needed if the number of users increases;
4) Lastly, the channel dispersion will converge to a constant when the number of antennas at the BS goes to infinity.

Notation: Boldface lower and upper case letters are used to denote vectors and matrices, respectively. The notation $\|\cdot\|_2$ and $(\cdot)^T$ denote the 2-norm and the transpose of a vector or matrix, respectively. The operators $\mathbb{E}\{x\}$, $\text{Var}\{x\}$, $\Re(x)$ and $\Im(x)$ denote the mean, variance, real and imaginary part of a random variable x, respectively. The notation $\mathcal{CN}(0,\sigma^2)$ represents the complex Gaussian distribution with zero mean and variance σ^2 and $\mathcal{C}^{M \times N}$ denotes complex arrays with dimension $M \times N$. Finally, $f(x) = \mathcal{O}(g(x))$ means that $\lim_{x \to \infty} \frac{f(x)}{g(x)} = c < \infty$.

II. SYSTEM MODEL

In the uplink, K single-antenna users transmit signals to the common BS with N antennas. A coherent block fading channel is considered, i.e., the BS has perfect CSI. Assume that each block has T_c symbols and there are in total L blocks. Thus, the finite coding blocklength is $n = LT_c$. The transmission process is as follows.

At a certain block j, for $j = 1, 2, \ldots, L$, for each symbol $i \in [T_c] = \{1, 2, \ldots, T_c\}$, the channel input-output relation or the received signal at the BS is

$$\mathbf{y}_i(j) = \sum_{k=1}^{K} \mathbf{h}_k(j) x_{k,i}(j) + \mathbf{n}_i(j), \qquad (2)$$

where $\mathbf{y}_i \in \mathcal{C}^{N \times 1}$; $\mathbf{h}_k(j) = [h_{k,1}(j), h_{k,2}(j), \ldots, h_{k,N}(j)]^T \in \mathcal{C}^{N \times 1}$ denotes the channel vector from the k-th user to the BS at block j, which has independent and identical distributed (i.i.d.) elements and each element is distributed as $\mathcal{CN}(0, \gamma_k^2)$; $x_{k,i}(j)$ denotes the i-th transmit symbol for user k at block j; and $\mathbf{n}_i(j) = [n_{i,1}(j), \ldots, n_{i,N}(j)]^T \in \mathcal{C}^{N \times 1}$ is the AWGN vector, which has i.i.d. elements with zero mean and variance σ^2.

The received signal matrix at BS, $\mathbf{Y}(j) \in \mathcal{C}^{N \times T_c}$, for all symbols $[T_c]$ at block j, is

$$\mathbf{Y}(j) = \sum_{k=1}^{K} \mathbf{h}_k(j) \mathbf{x}_k^T(j) + \mathbf{N}(j) = \mathbf{H}(j)\mathbf{X}(j) + \mathbf{N}(j), \quad (3)$$

where $\mathbf{Y}(j) = [\mathbf{y}_1(j), \ldots, \mathbf{y}_{T_c}(j)] \in \mathcal{C}^{N \times T_c}$; $\mathbf{x}_k(j) = [x_{k,1}(j), \ldots, x_{k,T_c}(j)]^T$ contains the transmitted symbols of length T_c for user k; $\mathbf{N}(j) = [\mathbf{n}_1(j), \ldots, \mathbf{n}_{T_c}(j)] \in \mathcal{C}^{N \times T_c}$ is the AWGN matrix with i.i.d. $\mathcal{CN}(0, \sigma^2)$ elements; $\mathbf{H}(j) = [\mathbf{h}_1(j), \ldots, \mathbf{h}_K(j)] \in \mathcal{C}^{N \times K}$ denotes the uplink channel matrix from all users to the BS; $\mathbf{X}(j) = [\mathbf{x}_1(j), \ldots, \mathbf{x}_K(j)]^T \in \mathcal{C}^{K \times T_c}$ contains the transmitted signals from all users.

Based on [6] and [13], when the distribution of the channel realization $\mathbf{H}(j)$ is isotropic and the input signal $\mathbf{X}(j)$ satisfies $\mathbf{X}(j) \sim \mathcal{CN}(\mathbf{0}, T_c\mathbf{P})$, in which $\mathbf{P} \in \mathcal{C}^{K \times K}$ is a diagonal matrix with the k-th diagonal element equal to the transmit power constraint P_k of user k, for $k = 1, \ldots, K$, the Shannon capacity can be achieved and this capacity achieving input distribution is unique when the number of BS antennas is $N \geq 2$ [11, Proposition 3]. Thus, in the following, we will always assume that $\mathbf{X}(j) \sim \mathcal{CN}(\mathbf{0}, T_c\mathbf{P})$. With this assumption, the transmitted signal for each user at each block satisfies the power constraint

$$\mathbb{E}\left\{\|\mathbf{x}_k(j)\|_2^2\right\} = T_c P_k, \quad k = 1, 2, \ldots, K. \qquad (4)$$

Meanwhile, we also assume that $\mathbf{x}_k(j) \in \mathcal{CN}(\mathbf{0}, P_k\mathbf{I}_{T_c})$, for $j = 1, \ldots, L$ and $k = 1, 2, \ldots, K$.

Without any combination scheme at the BS, (3) can be rewritten as

$$\mathbf{Y}(j) = \mathbf{h}_k(j)\mathbf{x}_k^T(j) + \sum_{\substack{m=1 \\ m \neq k}}^{K} \mathbf{h}_m(j)\mathbf{x}_m^T(j) + \mathbf{N}(j), \qquad (5)$$

where the first term denotes the desired signal, the second term represents interference from other users and the last term is the AWGN.

III. PERFORMANCE ANALYSIS

Denoting the interference matrix in (5) as

$$\mathbf{N}^{\mathrm{I}}(j) \triangleq \sum_{\substack{m=1 \\ m \neq k}}^{K} \mathbf{h}_m(j)\mathbf{x}_m^T(j), \qquad (6)$$

and treating the interference plus noise term in (5) as the effective noise denoted as

$$\mathbf{N}^{\mathrm{total}}(j) \triangleq \mathbf{N}^{\mathrm{I}}(j) + \mathbf{N}(j). \qquad (7)$$

Then, (5) can be rewritten as

$$\mathbf{Y}(j) = \mathbf{h}_k(j)\mathbf{x}_k^T(j) + \mathbf{N}^{\mathrm{total}}(j), \qquad (8)$$

which now can be treated as the single-input-multiple-output (SIMO) fading channel with the input signal $\mathbf{x}_k^T(j)$, the fading channel $\mathbf{h}_k(j)$ and the noise $\mathbf{N}^{\mathrm{total}}(j)$. Here, we should note that the noise $\mathbf{N}^{\mathrm{total}}(j)$ is not Gaussian distributed due to the effect of the interference term $\mathbf{N}^{\mathrm{I}}(j)$.

In (8), the signal-to-interference-plus-noise (SINR) ratio for the user k is

$$\mathrm{SINR}_k = \frac{P_k\|\mathbf{h}_k(j)\|_2^2}{\sigma^2 + \sigma_{K/k}^2}, \qquad (9)$$

where $\sigma_{K/k}^2 = \sum_{\substack{m=1 \\ m \neq k}}^{K} \gamma_m^2 P_m$ is the second-order moment of the interference term $\mathbf{N}^{\mathrm{I}}(j)$ in (6).

Then, by treating the noise $\mathbf{N}^{\mathrm{total}}(j)$ in (7) as Gaussian noise and with the mismatched decoder at the BS as in [14], the maximum coding rate in nats per channel use of the k-th user $R_k(n, \epsilon_k)$ with blocklength n and block error probability not greater than ϵ_k can be asymptotically expressed as[1]

$$R_k(n, \epsilon_k) = C_k - \sqrt{\frac{V_k}{n}}Q^{-1}(\epsilon_k) + O\left(\frac{1}{L}\right), \qquad (10)$$

where $0 < \epsilon_k < 1/2$ is the maximum error probability[2] that is allowed for transmission for user k; C_k denotes the Shannon capacity which admits the form

$$C_k = \mathbb{E}\left\{\ln\left[1 + \mathrm{SINR}_k\right]\right\}. \qquad (11)$$

The term V_k is the channel dispersion, which is used to characterize the backoff of the maximum coding rate from the capacity under finite blocklength, and has the form

$$V_k = T_c \operatorname{Var}\left\{\ln\left[1 + \mathrm{SINR}_k\right]\right\}$$
$$+ 2\left\{1 - \mathbb{E}\left\{\frac{1}{(1 + \mathrm{SINR}_k)}\right\}\right\}$$
$$+ \left(\frac{\xi}{T_c(\sigma_{K/k}^2 + \sigma^2)^2} - T_c - 1\right)\mathbb{E}\left\{\frac{1}{\left(1 + \frac{1}{\mathrm{SINR}_k}\right)^2}\right\}, \qquad (12)$$

where ξ is as

$$\xi \triangleq T_c\left\{(T_c + 1)(\sigma_{K/k}^2 + \sigma^2)^2\right.$$
$$\left. + (T_c + 2)\sum_{\substack{m=1 \\ m \neq k}}^{K_u} \gamma_m^4 P_m^2 - \left(\sum_{\substack{m=1 \\ m \neq k}}^{K_u} \gamma_m^2 P_m\right)^2\right\}. \qquad (13)$$

In the following, we will theoretically characterize the

maximum coding rate in (10). The exact analytical expression under the general case and the approximation result under large number of the BS antennas will be provided.

A. Analytical Result of Maximum Coding Rate

Proposition 1: The analytical expression of the maximum coding rate (10) under finite blocklength for an intended user k can be expressed as follows,

$$R_k(n, \epsilon_k) = \bar{R}_k^c(n, \epsilon_k)$$
$$= \bar{C}_k^c - \sqrt{\frac{\bar{V}_k^c}{n}}Q^{-1}(\epsilon_k) + O\left(\frac{1}{L}\right), \qquad (14)$$

where $\bar{C}_k^c$ and $\bar{V}_k^c$ denote the analytical results of the capacity C_k in (11) and the channel dispersion V_k in (12) respectively, which have the form as

$$\bar{C}_k^c = \frac{1}{\Gamma(N)}G_{2,3}^{3,1}\left(\frac{1}{2v_k}\middle|\begin{array}{c} 0, 1 \\ 0, 0, N \end{array}\right), \qquad (15)$$

$$\bar{V}_k^c = \frac{2T_c}{\Gamma(N)}e^{\frac{1}{2v_k}}\sum_{m=0}^{N-1}(-1)^m C_{N-1}^m$$
$$\times G_{3,4}^{4,0}\left(\frac{1}{2v_k}\middle|\begin{array}{c} m+1, m+1, m+1 \\ N, m, m, m \end{array}\right)$$
$$- T_c\left\{\frac{1}{\Gamma(N)}G_{2,3}^{3,1}\left(\frac{1}{2v_k}\middle|\begin{array}{c} 0, 1 \\ 0, 0, N \end{array}\right)\right\}^2$$
$$+ 2\left\{1 - \frac{1}{\Gamma(N)}G_{1,2}^{2,1}\left(\frac{1}{2v_k}\middle|\begin{array}{c} 1 \\ N, 1 \end{array}\right)\right\}$$
$$+ \left(\frac{\xi}{T_c(\sigma_{K/k}^2 + \sigma^2)^2} - T_c - 1\right)\frac{1}{\Gamma(N)}G_{1,2}^{2,1}\left(\frac{1}{2v_k}\middle|\begin{array}{c} -1 \\ N, 0 \end{array}\right), \qquad (16)$$

where $C_m^k = \frac{m!}{k!(m-k)!}$ denotes the binomial coefficient. The symbol $G_{p,q}^{m,n}\left(x\middle|\begin{array}{c} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{array}\right)$ is Meijer's G-function [15, Eq. (9.301)] and the parameter $\frac{1}{2v_k} \triangleq \frac{\sigma^2 + \sigma_{K/k}^2}{\gamma_k^2 P_k}$.

Proof: See Appendix A. ∎

In Proposition 1, define the signal-to-noise-ratio (SNR) as $\mathrm{SNR} = \frac{P_1}{\sigma^2} = \ldots = \frac{P_K}{\sigma^2} = \frac{P}{\sigma^2}$. When $\mathrm{SNR} = \frac{P}{\sigma^2} \to \infty$, we have

$$\frac{1}{2v_k} = \frac{\sigma_{K/k}^2}{\gamma_k^2} + \frac{1}{\gamma_k^2}\frac{1}{\mathrm{SNR}} \to \frac{\sigma_{K/k}^2}{\gamma_k^2}, \qquad (17)$$

which is a constant limited by the inter-user interference. By replacing all the coefficient $\frac{1}{2v_k}$ in Proposition 1 with the term $\frac{\sigma_{K/k}^2}{\gamma_k^2}$, we can conjecture that the maximum coding rate will be limited to a constant under high SNR due to the existence of the inter-user interference.

The result in Proposition 1, though analytical, provides limited insights due to presence of the non-linear Meijer's G-function. For this reason, we now present an approximation result of the maximum coding rate.

B. Approximation Result of the Maximum Coding Rate

In this part, we will provide the approximation of the maximum coding rate for a large number of the BS antennas.

Firstly, recall the functions and variables defined in Appendix A, from the variance approximation in [16, Chapt.

---

[1]The maximum coding rate in our paper is obtained under the assumption that the input signals are i.i.d. Gaussian distributed, while the result obtained by Polyanskiy, Poor and Verdú in [4] assumed that the input signals are under the unitary space time modulation (USTM).

[2]Here, the error probability $0 < \epsilon_k < 1/2$ is used to guarantee $Q^{-1}(\epsilon_k)$ is positive.

4.3.2], the approximation of the variance of $g(x_k(j))$ in (28) of Appendix A can be obtained as in the following lemma.

Lemma 1: The variance of the Shannon capacity $g(x_k(j)) = \ln\left(1 + v_k x_k(j)\right)$ can be approximated as

$$\text{Var}\{g(x_k(j))\} = \text{Var}\left\{\ln\left[1 + \text{SINR}_k\right]\right\} \approx \frac{4N v_k^2}{(1 + 2N v_k)^2}. \quad (18)$$

Then, the defined random variable $x_k(j) \triangleq \|\frac{\sqrt{2}}{\gamma_k}\mathbf{h}_k(j)\|_2^2$ in Appendix A, is distributed as $x_k(j) \sim \chi^2(2N)$. With [17, Eq. (74)], when the number of the BS antennas N is large, the following approximation holds tight,

$$x_k(j) \approx \mathbb{E}\{x_k(j)\} = 2N. \quad (19)$$

By substituting the above approximation result in Lemma 1 and the approximation result in (19) into the maximum coding rate in (10) and the corresponding terms (11) and (12), we obtain the approximating maximum coding rate shown in the following proposition.

Proposition 2: For a large number of the BS antennas N, the maximum coding rate of (10) can be approximated as

$$\bar{R}_k^{\text{app}}(n, \epsilon_k) = \bar{C}_k^{\text{app}} - \sqrt{\frac{\bar{V}_k^{\text{app}}}{n}} Q^{-1}(\epsilon_k) + O\left(\frac{1}{L}\right), \quad (20)$$

where $\bar{C}_k^{\text{app}}$ and $\bar{V}_k^{\text{app}}$ are the approximations for the capacity and channel dispersion, respectively, given by

$$\bar{C}_k^{\text{app}} = \ln\left[1 + 2v_k N\right], \quad (21)$$

$$\bar{V}_k^{\text{app}} = T_c \frac{4N v_k^2}{(1 + 2N v_k)^2} + 2\left(1 - \frac{1}{1 + 2v_k N}\right)$$
$$+ \left(\frac{\xi}{T_c(\sigma_{K/k}^2 + \sigma^2)^2} - T_c - 1\right) \frac{1}{\left(1 + \frac{1}{2v_k N}\right)^2}. \quad (22)$$

In Proposition 2, by increasing the number of BS antennas N, the capacity in (21) will grow without bound, whilst the channel dispersion in (22) will approach to a constant as

$$\bar{V}_k^{\text{app}} \xrightarrow{N \to \infty} \bar{V}^{\infty} = 2 + \left(\frac{\xi}{T_c(\sigma_{K/k}^2 + \sigma^2)^2} - T_c - 1\right). \quad (23)$$

As given in [6, Eq. (3)], by ignoring the third term $O\left(\frac{1}{L}\right)$ of (10) as [9, Eq. (9)], the minimum blocklength required for achieving a certain fraction $0 < \eta_k < 1$ of the Shannon capacity can be approximated as

$$n \gtrsim \left[\frac{Q^{-1}(\epsilon_k)}{1 - \eta_k}\right]^2 \frac{V_k}{C_k^2}, \quad (24)$$

where $V_k^{\text{nor}} = \frac{V_k}{C_k^2}$ is defined as the normalized channel dispersion. From Proposition 2, we have

$$V_k^{\text{nor}} = \frac{V_k}{C_k^2} \approx \frac{\bar{V}_k^{\text{app}}}{(\bar{C}_k^{\text{app}})^2} \xrightarrow{N \to \infty} 0. \quad (25)$$

Thus, the minimum blocklength required in (24) will approach to zero with increasing number of the BS antennas. This implies that employing large number of antennas at the BS is beneficial in the URLLC systems.

IV. NUMERICAL RESULTS

In this section, a set of numerical results is provided to verify the correctness of the analytical results as well as to enable us to find some new insights. These numerical results

are given under some basic system parameters setting as in Table I unless otherwise stated.

Please note that in this numerical section, all the numerical results are given by ignoring the term $O\left(\frac{1}{L}\right)$ in (10). This is reasonable as this term is very small and can be neglected without affecting the performance but is convenient for the theoretical analysis, as done in the majority of relevant papers [4, Eq. (1)] [9, Eq. (9)].

TABLE I: Numerical Setting

Coherence interval T_c	$T_c = 30$
Number of blocks L	$L = 20$
Blocklength n	$n = LT_c = 600$
Number of the BS antennas	$N = 10$
Number of users	$K = 6$
Transmit power	$P_1 = \ldots = P_K = P = 20\text{dBm}$
Noise variances	$\sigma^2 = 1$
Error probability	$\epsilon_k = 10^{-5}$, for $k = 1, \ldots, K$
Channel gain	$\gamma_k^2 = 1$, for $k = 1, \ldots, K$

A. The Analytical Result in Proposition 1

Fig. 1 and Fig. 2 verify the correctness of the analytical result of the maximum coding rate presented in Proposition 1. Specifically, from Fig. 1, the maximum coding rate under finite blocklength will converge to the Shannon capacity with large number of blocks L and better performance can be obtained with longer coherence interval T_c. This is reasonable because the channel dispersion will be averaged under these two cases. Also, from Fig. 2, increasing the SNR of each user can boost the performance to a ceiling due to the presence of inter-user interference. Furthermore, for a fixed total blocklength n, with smaller coherence interval T_c (fast channel dynamics) and larger number of blocks L (high diversity gain), we can obtain higher maximum coding rate. This is because under this scenario, the information can be coded via a large number of independent channel realizations, which provides higher diversity gain. On the other hand, large number of blocks needs more symbols for channel estimation, which will lead to a rate loss. But in this paper, as we assume the coherent case, there is no rate loss. Thus, the performance can be improved and this result has also been proved in [9].

B. The Approximation Result in Proposition 2

The accuracy of the approximation result under a large number of the BS antennas in Proposition 2 is illustrated in Fig. 3. From this figure, the maximum coding rate, the capacity and the channel dispersion match very well with the corresponding simulation results. Futhermore, we can infer that even when the channel dispersion caused by finite blocklength is considered, the maximum coding rate is still an increasing function with respect to the number of the BS antennas. More interestingly, it shows that the channel dispersion converges to a constant with large number of the BS antennas.

From Fig. 4, though our approximation result in Proposition 2 is given under a large number of the BS antennas, it also matches well with respect to the number of users. Also, from this figure, increasing the number of users reduces the maximum coding rate of an individual user and no solid
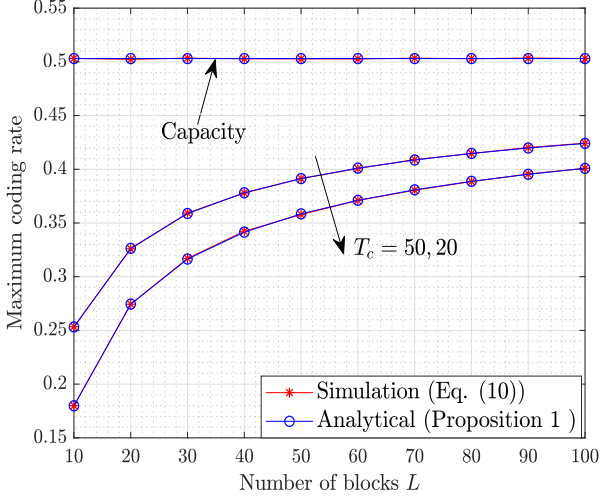
Fig. 1: The maximum coding rate versus the number of blocks with $K = 6$, $N = 10$ and $P = 20$dBm.

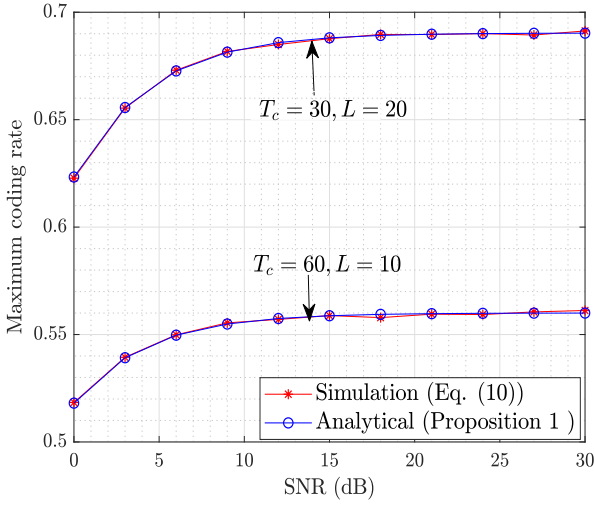Fig. 2: The maximum coding rate versus the SNR $= \frac{P}{\sigma^2}$ with $K = 6$, $N = 10$, and $n = LT_c = 600$.

Fig. 3: The maximum coding rate and the channel dispersion versus the number of BS antennas with $K = 6$.
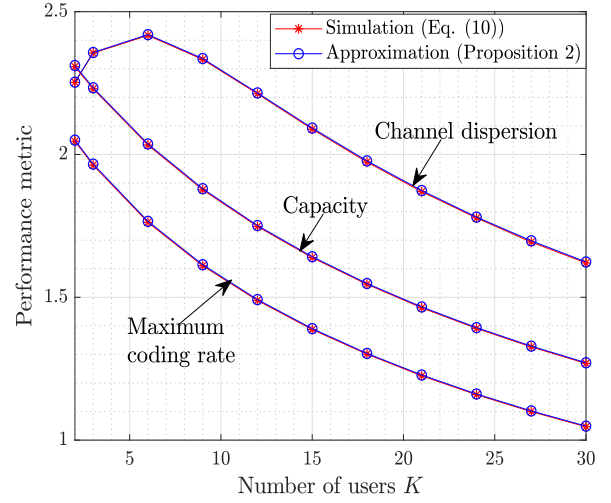
Fig. 4: Maximum coding rate, capacity and the channel dispersion versus the number of users with $N = 100$.
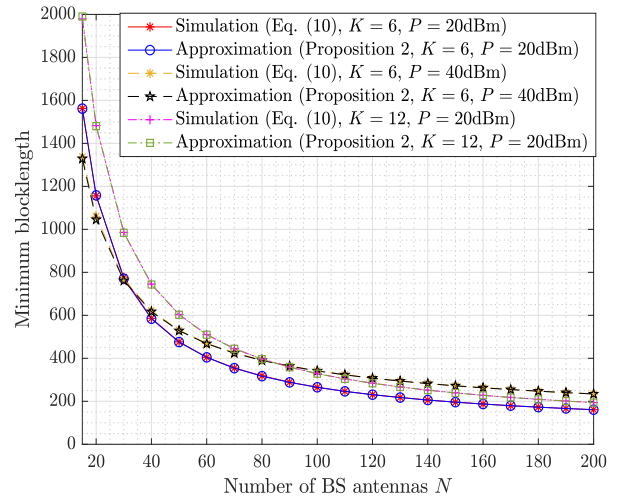
Fig. 5: The minimum blocklength versus the number of BS antennas with $\eta_k = 0.8$.

conclusion has been drawn about the channel dispersion with respect to the number of users.

Most importantly, in Fig. 5, a larger number of the BS antennas is beneficial in reducing the minimum blocklength (delay) required for obtaining a certain level of the capacity and the minimum blocklength approaches to zero with increasing the number of BS antennas. Moreover, increaing the number of users would need longer blocklength since the inter-user interference is increased. However, increasing the transmit power of each user may not be beneficial to the minimum blocklength, especially in the large number of the BS antennas region.

V. CONCLUSION

An uplink multiple-user MIMO system with a multiple-antenna BS under coherent block-fading channel and finite blocklength has been investigated. A new and more refined performance metric, the maximum coding rate under finite blocklength was studied. We derived a new, analytical ex-
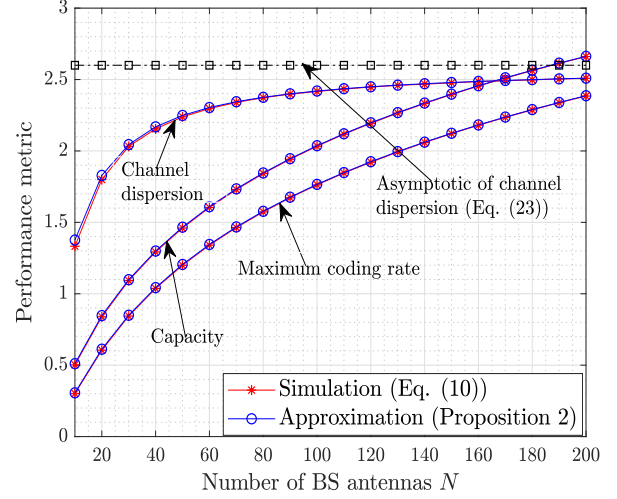
pression based on the Meijer's G-function, and approximation results under large number of the BS antennas. From these theoretical results, we obtained the following insights: i) faster channel fading dynamics (shorter coherence time) and higher diversity gain (large number of blocks) help to improve the maximum coding rate; ii) increasing the number of BS antennas boosts the performance and the SNR for each user improves the performance to a ceiling limited by the inter-user interference; iii) the minimum blocklength required for obtaining a fixed fraction of rate reduces with large number of the BS antennas; and iv) the channel dispersion will converge to a constant while the minimum blocklength required will approach to zero when a massive number of antennas is deployed at the BS. Thus, we can conclude that increasing the number of BS antennas will be beneficial in URLLC systems when considering finite blocklength channel coding.

APPENDIX A
PROOF OF PROPOSITION 1

Recall the maximum coding rate in (10) and the corresponding terms (11) and (12). We can now define the coefficient $v_k \triangleq \frac{\gamma_k^2}{2} \frac{P_k}{\sigma^2 + \sigma_{K/k}^2}$ and the random variable $x_k(j) \triangleq \|\frac{\sqrt{2}}{\gamma_k} \mathbf{h}_k(j)\|_2^2$, then, $x_k(j)$ is written as

$$x_k(j) = \frac{2}{\gamma_k^2} \sum_{m=1}^{N} \left\{ [\Re(h_{k,m}(j))]^2 + [\Im(h_{k,m}(j))]^2 \right\}, \quad (26)$$

which is distributed as $\chi^2(2N)$ and the corresponding PDF is

$$f(x_k(j)) = \begin{cases} \frac{(x_k(j))^{N-1} e^{-\frac{x_k(j)}{2}}}{2^N \Gamma(N)}, & x_k(j) > 0 \\ 0, & \text{otherwise} \end{cases}. \quad (27)$$

With this, we define the related terms in (10)-(12) as

$$g(x_k(j)) \triangleq \ln(1 + \text{SINR}_k) = \ln(1 + v_k x_k(j)), \quad (28)$$

$$h(x_k(j)) \triangleq \frac{1}{1 + \text{SINR}_k} = \frac{1}{1 + v_k x_k(j)}, \quad (29)$$

$$q(x_k(j)) \triangleq \frac{1}{\left(1 + \frac{1}{\text{SINR}_k}\right)^2} = \frac{1}{\left(1 + \frac{1}{v_k x_k(j)}\right)^2}. \quad (30)$$

Then, based on (27), the first-order moment of $g(x_k(j))$ is

$$\mathbb{E}\{g(x_k(j))\} = \int_0^{+\infty} g(x_k(j)) f(x_k(j)) dx_k(j)$$

$$= \frac{1}{\Gamma(N)} G_{2,3}^{3,1}\left(\frac{1}{2v_k} \middle| \begin{array}{c} 0,1 \\ 0,0,N \end{array} \right). \quad (31)$$

The second-order moment of $g(x_k(j))$ can be obtained from [18, Eq. (41)], which is as follows

$$\mathbb{E}\left\{ [g(x_k(j))]^2 \right\} = \int_0^{+\infty} [g(x_k(j))]^2 f(x_k(j)) dx_k(j)$$

$$= \frac{2}{\Gamma(N)} e^{\frac{1}{2v_k(j)}} \sum_{m=0}^{N-1} (-1)^m C_{N-1}^m$$

$$\times G_{3,4}^{4,0}\left(\frac{1}{2v_k} \middle| \begin{array}{c} m+1, m+1, m+1 \\ N, m, m, m \end{array} \right). \quad (32)$$

Combining the results in (31) and (32), the variance of the random variable $g(x_k(j))$ can be derived as

$$\text{Var}\{g(x_k(j))\} = \mathbb{E}\left\{ [g(x_k(j))]^2 \right\} - \left[\mathbb{E}\{g(x_k(j))\} \right]^2. \quad (33)$$

Now, the expectation of $h(x_k(j)) = \frac{1}{1+v_k(j)x_k(j)}$ is

$$\mathbb{E}\{h(x_k(j))\} = \int_0^{+\infty} h(x_k(j)) f(x_k(j)) dx_k(j)$$

$$= \frac{1}{\Gamma(N)} G_{1,2}^{2,1}\left(\frac{1}{2v_k} \middle| \begin{array}{c} 1 \\ N, 1 \end{array} \right). \quad (34)$$

The mean value of $q(x_k(j)) = \frac{1}{\left(1 + \frac{1}{v_k x_k(j)}\right)^2}$ is as

$$\mathbb{E}\{q(x_k(j))\} = \int_0^{+\infty} q(x_k(j)) f(x_k(j)) dx_k(j)$$

$$= \frac{1}{\Gamma(N)} G_{1,2}^{2,1}\left(\frac{1}{2v_k} \middle| \begin{array}{c} -1 \\ N, 0 \end{array} \right). \quad (35)$$

Substituting (31)-(35) into (10), Proposition 1 can be obtained, which completes the proof.

REFERENCES

[1] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[2] A. Avranas, M. Kountouris, and P. Ciblat, "Energy-latency tradeoff in ultra-reliable low-latency communication with retransmissions," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2475–2485, Nov. 2018.

[3] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Block-fading channels at finite blocklength," in *Proc. IEEE ISWCS*, Aug. 2013, pp. 1–4.

[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[5] J. Scarlett, V. Y. Tan, and G. Durisi, "The dispersion of nearest-neighbor decoding for additive non-Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 81–92, Jan. 2017.

[6] Y. Polyanskiy and S. Verdú, "Scalar coherent fading channel: Dispersion analysis," in *Proc. IEEE ISIT*, Aug. 2011, pp. 2959–2963.

[7] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," in *Proc. IEEE ISIT*, Jul. 2013, pp. 1531–1535.

[8] ——, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jul. 2014.

[9] ——, "Diversity versus channel knowledge at finite block-length," in *Proc. IEEE ITW*, Sep. 2012, pp. 572–576.

[10] A. Lancho, T. Koch, and G. Durisi, "On single-antenna Rayleigh block-fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 496–519, Jan. 2020.

[11] A. Collins and Y. Polyanskiy, "Dispersion of the coherent MIMO block-fading channel," in *Proc. IEEE ISIT*, Jul. 2016, pp. 1068–1072.

[12] ——, "Coherent multiple-antenna block-fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 380–405, Jan. 2019.

[13] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecom.*, vol. 10, no. 6, pp. 585–595, 1999.

[14] J. Östman, A. Lancho, and G. Durisi, "Short-packet transmission over a bidirectional massive MIMO link," *2019, [Online]. Available: https://arxiv.org/abs/1912.00718*.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic press, 2014.

[16] H. Benaroya, S. M. Han, and M. Nagurka, *Probability Models in Engineering and Science*. CRC press, 2005.

[17] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[18] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.