Uplink Performance of Conventional and Massive MIMO Cellular Systems with Delayed CSIT

Anastasios K. Papazafeiropoulos*, Hien Quoc Ngo†, Michail Matthaiou§,‡, and Tharmalingam Ratnarajah♯

*Communications and Signal Processing Group, Imperial College London, London, U.K.
†Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden
§School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K.
‡Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
♯Institute for Digital Communications (IDCoM), University of Edinburgh, Edinburgh, U.K.
Email: a.papazafeiropoulos@imperial.ac.uk, nqhien@isy.liu.se, m.matthaiou@qub.ac.uk, t.ratnarajah@ed.ac.uk

***Abstract*—This work studies the uplink of a cellular network with zero-forcing (ZF) receivers under imperfect channel state information at the base station. More specifically, apart from the pilot contamination, we investigate the effect of time variation of the channel due to the relative users' movement with regard to the base station. Our contributions include analytical expressions for the sum-rate with finite number of BS antennas, and also the asymptotic limits with infinite power and number of BS antennas, respectively. The numerical results provide interesting insights on how the user mobility degrades the system performance which extends previous results in the literature.**

I. Introduction

Among the emerging technological breakthroughs of next generation (5G) networks, massive MIMO, the extension of the well-known multi-user multiple-input multiple-output (MU-MIMO), has attracted a lot of interest recently [1]. In general, MU-MIMO can achieve higher throughput by exploiting user multiplexing [2]. In massive MIMO, capitalizing on this principle, a base station (BS) equipped with several tens to hundreds of antennas, communicates with a few tens of users on the same time-frequency resource [3]. The key advantages of these configurations are: higher data rates, low-complexity precoding and decoding algorithms with near-optimal performance, as well as improved energy efficiency [3]–[5].

On a similar note, zero-forcing (ZF) processing is regarded as a low-complexity alternative of maximum-likelihood multiuser detector and "dirty paper coding" [6], especially, when the BSs are equipped with massive antenna arrays. A lot of research has been conducted on single-cell systems with ZF receivers [7], but the main current interest has shifted to practical multi-cell scenarios, where pilot contamination degrades the system performance [3], [8].

This research was supported by a Marie Curie Intra European Fellowship and HARP project within the 7th European Community Framework Programme for Research of the European Commission under grant agreements no. [330806], IAWICOM and no. [318489], HARP.

An important question, which has been largely overlooked in the massive MIMO literature, is how the performance of a massive MIMO topology is affected by the relative movement of users. This scenario is of high practical importance, in urban environments, where users move rapidly within a geographical area. The main challenge in time-varying environments is to perform robust channel estimation, when the propagation channel changes over time. The dynamic channel behavior was modeled in terms of a stationary ergodic Gauss-Markov block fading channel model [9]–[11], where an autoregressive model was combined with the Jakes' autocorrelation function that captures the time variation of the channel.

Motivated by the above observation, this paper makes the following contributions: First, we consider the uplink of a multi-cell topology containing multiple users per cell and obtain an analytical expression for the achievable sum rate with ZF receivers. Note that, our analysis holds for any *finite number of antennas*, which stands in contrast to the works in [9], [11], where only a deterministic equivalent analysis was pursued for a matched filter and minimum-mean-square-error (MMSE) detector, respectively. We recall that this type of analysis is asymptotic and requires several conditions to be fulfilled. Using the methodology of [12], which completely ignored the user mobility, we present a detailed performance analysis of massive MIMO suffering from both pilot contamination and time-varying channels due to user mobility.

Notation: Throughout this paper, we use boldface lowercase and uppercase letters to denote vectors and matrices, respectively. The notation $(\bullet)^{\mathsf{H}}$ stands for the conjugate transpose, and $\| \bullet \|$ denotes the Euclidean norm of a vector, while $(\bullet)^{\dagger}$ denotes the pseudo-inverse of a matrix. We use the notation $x \overset{\mathrm{d}}{\sim} y$ to imply that x and y have the same distribution.

978-1-4799-4912-0/14/$31.00

II. SYSTEM MODEL

We consider a cellular network with L non-cooperative BSs. Each BS, equipped with N antennas, serves K single-antenna users. Here, we assume that: i) $N \geq K$, and ii) all KL users share the same time-frequency resource. Moreover, the channels are frequency flat and obey the conditions of a quasi-static block fading model, i.e., they vary from symbol to symbol, while during the symbol period they are considered constant. We assume that the channel vector $\boldsymbol{g}_{lik}[n] \in \mathbb{C}^{N\times 1}$ between the kth user in the ith cell and the BS in cell l at the nth symbol undergoes independent small-scale fading and large-scale fading (shadow fading and path loss). In particular, we have

$$\boldsymbol{g}_{lik}[n] = \sqrt{\beta_{lik}}\boldsymbol{h}_{lik}[n], \tag{1}$$

where β_{lik} represents shadow fading and path loss, while $\boldsymbol{h}_{lik} \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_N)$ represents the uncorrelated small-scale fading vector between the lth BS and the kth user in the ith cell. The $N \times 1$ received uplink signal vector at the lth BS is given by

$$\boldsymbol{y}_l[n] = \sqrt{p_{\mathrm{u}}}\sum_{i=1}^{L}\boldsymbol{G}_{li}[n]\boldsymbol{x}_i[n] + \boldsymbol{z}_l[n], \quad l = 1, 2, ..., L, \tag{2}$$

where $\sqrt{p_{\mathrm{u}}}\boldsymbol{x}_i[n] \in \mathbb{C}^{K\times 1}$ is the zero-mean stochastic data signal vector of K users allocated in the ith cell with $p_{\mathrm{u}} > 0$ being the average transmitted power per user, $\boldsymbol{G}_{li} \triangleq [\boldsymbol{g}_{li1}, \ldots, \boldsymbol{g}_{liK}] \in \mathbb{C}^{N\times K}$ denotes the combined matrix between all users in cell i and BS l, and $\boldsymbol{z}_l$ is the additive white Gaussian noise (AWGN) vector, modeled as $\mathcal{CN}(\mathbf{0}, \boldsymbol{I}_N)$.

To detect the signals transmitted from the K users in the lth cell, the lth BS has to acquire CSI. Conventionally, the lth BS estimates the CSI via uplink training. We assume that in each cell, K users are assigned K orthogonal pilots sequences of length τ symbols. Due to the limitations of the coherence interval, we further assume that all L cells use the same set of orthogonal pilot sequences [3], [4]. Denote by $\boldsymbol{\Psi} \in \mathbb{C}^{K\times\tau}$ ($\tau \geq K$) the pilot matrix transmitted from the K users in each cell with $\boldsymbol{\Psi}$ satisfying $\boldsymbol{\Psi}\boldsymbol{\Psi}^{\mathsf{H}} = \mathbf{I}_K$. Then, the $N \times \tau$ received pilot signal at BS l is[1]

$$\boldsymbol{Y}_l^{\mathrm{tr}}[n] = \sqrt{p_{\mathrm{p}}}\sum_{i=1}^{L}\boldsymbol{G}_{li}[n]\boldsymbol{\Psi} + \boldsymbol{Z}_l^{\mathrm{tr}}[n], \quad l = 1, 2, ..., L \tag{3}$$

where $p_{\mathrm{p}} \triangleq \tau p_{\mathrm{u}}$, the superscript tr denotes the uplink training stage, and $\boldsymbol{Z}_l^{\mathrm{tr}}1[n] \in \mathbb{C}^{N\times\tau}$ is spatially AWGN at BS l during the training phase. The MMSE channel estimate of $\boldsymbol{g}_{lik}[n]$ is given by [9]

$$\hat{\boldsymbol{g}}_{lik}[n] = \beta_{lik}\boldsymbol{Q}_{lk}\left(\sum_{j=1}^{L}\boldsymbol{g}_{ljk}[n] + \frac{1}{\sqrt{p_{\mathrm{p}}}}\tilde{\boldsymbol{z}}_{lk}^{\mathrm{tr}}[n]\right), \tag{4}$$

where $\boldsymbol{Q}_{lk} \triangleq \left(\frac{1}{p_{\mathrm{p}}} + \sum_{i=1}^{L}\beta_{lik}\right)^{-1}\boldsymbol{I}_N$, and $\tilde{\boldsymbol{z}}_{lk}^{\mathrm{tr}}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represents the additive noise.

The orthogonality property of the MMSE channel estimation allows us to decompose $\boldsymbol{g}_{lik}[n]$ as:

$$\boldsymbol{g}_{lik}[n] = \hat{\boldsymbol{g}}_{lik}[n] + \tilde{\boldsymbol{g}}_{lik}[n], \tag{5}$$

where $\hat{\boldsymbol{g}}_{lik}[n] \sim \mathcal{CN}\left(\mathbf{0}, \hat{\beta}_{lik}\right)$ and $\tilde{\boldsymbol{g}}_{lik}[n] \sim \mathcal{CN}\left(\mathbf{0}, \left(\beta_{lik} - \hat{\beta}_{lik}\right)\mathbf{I}_N\right)$ with $\hat{\beta}_{lik} \triangleq \frac{\beta_{lik}^2}{\sum_{j=1}^{L}\beta_{ljk} + 1/p_{\mathrm{p}}}$ are the independent channel estimate and channel estimation error, respectively. Note that β_{lik}, $\hat{\beta}_{lik}$, and $\boldsymbol{Q}_{lk}$ are independent of n, $\forall l$, i, and k, since by assuming that the coherence time is large relative to any delay constraint of the channel, the shadowing can be kept constant.

Nevertheless, the performance degradation due to the time variation of the channel, induced by the relative movement between the antennas and the scatterers, is inevitable in practical systems, making its study of crucial interest. For this reason, in our analysis, we incorporate the Gauss-Markov block fading model, which correlates the current channel state with its past samples by considering two-dimensional isotropic scattering. More precisely, the channel $\boldsymbol{g}_{lik}[n]$ can be approximately modeled via an autoregressive model of order 1 as [9]:

$$\boldsymbol{g}_{lik}[n] = \alpha\boldsymbol{g}_{lik}[n-1] + \boldsymbol{e}_{lik}[n], \tag{6}$$

where $\boldsymbol{g}_{lik}[n-1]$ and $\boldsymbol{e}_{lik}[n] \sim \mathcal{CN}\left(\mathbf{0}, \left(1-\alpha^2\right)\beta_{lik}\mathbf{I}_N\right)$ are uncorrelated, denoting the channel at the previous symbol duration and the stationary Gaussian channel error vector due to the time variation of the channel, respectively. Note that $\alpha \triangleq \mathrm{J}_0(2\pi f_D T_s)$, where $\mathrm{J}_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_D and T_s are the maximum Doppler shift and the channel sampling period. Basically, α corresponds to the temporal correlation parameter that describes the isotropic scattering according to the Jakes' model. In particular, the maximum Doppler shift f_D equals $f_D = \frac{v f_c}{c}$, where v (in m/s) is the relative velocity of the user, $c = 3\times 10^8$m/s is the speed of light, and f_c is the carrier frequency.

Substituting (5) into (6), we obtain a model which combines both the effects of channel estimation error and channel aging as follows:

$$\begin{aligned}\boldsymbol{g}_{lik}[n] &= \alpha\boldsymbol{g}_{lik}[n-1] + \boldsymbol{e}_{lik}[n]\\ &= \alpha\hat{\boldsymbol{g}}_{lik}[n-1] + \tilde{\boldsymbol{e}}_{lik}[n],\end{aligned} \tag{7}$$

where $\hat{\boldsymbol{g}}_{lik}[n-1]$ and $\tilde{\boldsymbol{e}}_{lik}[n] \triangleq \alpha\tilde{\boldsymbol{g}}_{lik}[n-1] + \boldsymbol{e}_{lik}[n] \sim \mathcal{CN}\left(\mathbf{0}, \left(\beta_{lik} - \alpha^2\hat{\beta}_{lik}\right)\mathbf{I}_N\right)$ are mutually independent. We now define $\hat{\boldsymbol{G}}_{li}[n] \triangleq [\hat{\boldsymbol{g}}_{li1}[n], \ldots, \hat{\boldsymbol{g}}_{liK}[n]] \in \mathbb{C}^{N\times K}$ and

[1] As in previous works, we assume that the channel is estimated within one block. So, it remains constant during the training phase [9].

$$\gamma_k = \frac{\alpha^2 p_{\mathrm{u}}}{\alpha^2 p_{\mathrm{u}} \sum_{i\neq l}^{L} \left\| \left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k \hat{\boldsymbol{G}}_{li}[n-1] \right\|^2 + p_{\mathrm{u}} \sum_{i=l}^{L} \left\| \left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k \tilde{\boldsymbol{E}}_{li}[n] \right\|^2 + \left\| \left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k \right\|^2}. \tag{12}$$

$\tilde{\boldsymbol{E}}_{li} \triangleq [\tilde{\boldsymbol{e}}_{li1}[n], \ldots, \tilde{\boldsymbol{e}}_{liK}[n]] \in \mathbb{C}^{N\times K}$ as the combined channel matrices from all users in cell i to BS l. In particular, $\hat{\boldsymbol{G}}_{li}[n]$ can be expressed as [14]:

$$\hat{\boldsymbol{G}}_{li}[n] = \hat{\boldsymbol{G}}_{ll}[n]\boldsymbol{D}_i, \tag{8}$$

where $\boldsymbol{D}_i = \mathrm{diag}\{\frac{\beta_{li1}}{\beta_{ll1}}, \frac{\beta_{li2}}{\beta_{ll2}}, \ldots, \frac{\beta_{liK}}{\beta_{llK}}\}$.

Making use of (7), we can rewrite the received signal $\boldsymbol{y}_l[n]$ at the lth BS ($l \in [1, L]$) as

$$\boldsymbol{y}_l[n] = \alpha\sqrt{p_{\mathrm{u}}}\sum_{i=1}^{L}\hat{\boldsymbol{G}}_{li}[n-1]\boldsymbol{x}_i[n] + \sqrt{p_{\mathrm{u}}}\sum_{i=1}^{L}\tilde{\boldsymbol{E}}_{li}[n]\boldsymbol{x}_i[n] + \boldsymbol{z}_l[n]. \tag{9}$$

Moreover, we assume that the lth BS uses the ZF technique to detect the signals transmitted from K users in its cells. With ZF, the received signal vector $\boldsymbol{y}_l[n]$ is pre-multiplied with $\alpha^{-1}\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]$, where $\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]$ is the $K\times N$ pseudo-inverse matrix of $\hat{\boldsymbol{G}}_{ll}[n-1]$:

$$\begin{aligned} \boldsymbol{r}_l[n] &= \sqrt{p_{\mathrm{u}}}\boldsymbol{x}_l[n] + \sqrt{p_{\mathrm{u}}}\sum_{i\neq l}^{L}\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\hat{\boldsymbol{G}}_{li}[n-1]\boldsymbol{x}_i[n] \\ &+ \alpha^{-1}\sqrt{p_{\mathrm{u}}}\sum_{i=1}^{L}\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\tilde{\boldsymbol{E}}_{li}[n]\boldsymbol{x}_i[n] + \alpha^{-1}\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\boldsymbol{z}_l[n]. \end{aligned} \tag{10}$$

Then, the kth element of $\boldsymbol{r}_l[n]$ is used to detect the signal transmitted from the kth user. The post-processed received signal corresponding to the kth user is

$$\begin{aligned} r_{lk}[n] &= \sqrt{p_{\mathrm{u}}}x_{lk}[n] + \sqrt{p_{\mathrm{u}}}\sum_{i\neq l}^{L}\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\hat{\boldsymbol{G}}_{li}[n-1]\boldsymbol{x}_i[n] \\ &+ \alpha^{-1}\sqrt{p_{\mathrm{u}}}\sum_{i=1}^{L}\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\tilde{\boldsymbol{E}}_{li}[n]\boldsymbol{x}_i[n] + \alpha^{-1}\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\boldsymbol{z}_l[n], \end{aligned} \tag{11}$$

where the notation $[\boldsymbol{A}]_k$ refers to the kth row of matrix $\boldsymbol{A}$, and $x_{lk}[n]$ is the kth element of $\boldsymbol{x}_l[n]$, i.e., it is the transmit signal from the kth user in the lth cell at the nth time slot. Treating (11) as a single-input single-output (SISO) system, we obtain the SINR of the transmission from the kth user in the lth cell to its BS as in (12) shown at the top of the page.

III. Achievable Uplink Sum Rate

In this section, we pursue a sum-rate analysis for finite and infinite number of BS antennas taking into account the effects of pilot contamination and user mobility.

A. Finite-N Analysis

Proposition 1: The uplink SINR of transmission between the kth user in the lth cell to its BS is distributed as

$$\gamma_k \overset{\mathrm{d}}{\sim} \frac{\alpha^2 p_{\mathrm{u}} X_k[n-1]}{\alpha^2 p_{\mathrm{u}} C X_k[n-1] + p_{\mathrm{u}} Y_k[n] + 1}, \tag{13}$$

where X_k and Y_k are independent random variables (RVs) whose probability density functions (PDFs) are respectively given by

$$p_{X_k}(x) = \frac{e^{-x/\hat{\beta}_{llk}}}{(N-K)!\hat{\beta}_{llk}}\left(\frac{x}{\hat{\beta}_{llk}}\right)^{N-K}, \; x \geq 0 \tag{14}$$

$$p_{Y_k}(y) = \sum_{p=1}^{\varrho(\mathcal{A}_k)}\sum_{q=1}^{\tau_p(\mathcal{A}_k)}\mathcal{X}_{p,q}(\mathcal{A}_k)\frac{\mu_{k,p}^{-q}}{(q-1)!}y^{q-1}e^{\frac{-y}{\mu_{k,p}}}, \quad y \geq 0 \tag{15}$$

where $\mathcal{A}_k \triangleq \mathrm{diag}\left(\tilde{\boldsymbol{D}}_{l1}, \ldots, \tilde{\boldsymbol{D}}_{lL}\right) \in \mathbb{C}^{KL\times KL}$ with $\tilde{\boldsymbol{D}}_{li}$ a $K\times K$ diagonal matrix having elements $\left[\tilde{\boldsymbol{D}}_{li}\right]_{kk} = \left(\beta_{lik} - \alpha^2\hat{\beta}_{lik}\right)$, as well as $\varrho(\mathcal{A}_k)$ denotes the numbers of distinct diagonal elements of $\mathcal{A}_k$. Similarly, $\mu_{k,1}, \mu_{k,2}, \ldots, \mu_{k,\varrho(\mathcal{A}_k)}$ are the associated distinct diagonal elements in decreasing order and $\tau_p(\mathcal{A}_k)$ are the multiplicities of $\mu_{k,p}$, while $\mathcal{X}_{p,q}(\mathcal{A}_k)$ is the (p,q)th characteristic coefficient of $\mathcal{A}_k$, as defined in [15, Definition 4]. Regarding C, it is a constant value: $C \triangleq \sum_{i\neq l}^{L}\left(\frac{\beta_{lik}}{\beta_{llk}}\right)^2$.

Proof: Division of each term of (12) by $\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|^2$ leads to

$$\gamma_k = \frac{\alpha^2 p_{\mathrm{u}}\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|^{-2}}{\alpha^2 p_{\mathrm{u}} C\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|^{-2} + p_{\mathrm{u}}\sum_{i=1}^{L}\left\|\hat{\boldsymbol{Y}}_i[n]\right\|^2 + 1}, \tag{16}$$

where $C \triangleq \sum_{i\neq l}^{L}\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\hat{\boldsymbol{G}}_{li}[n-1]\right\| \overset{(8)}{=} \sum_{i\neq l}^{L}\left(\frac{\beta_{lik}}{\beta_{llk}}\right)^2$ and $\hat{\boldsymbol{Y}}_i[n] \triangleq \frac{\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\tilde{\boldsymbol{E}}_{li}[n]}{\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|}$.

Since $\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|^2 = \left[\left(\hat{\boldsymbol{G}}_{ll}^{\mathsf{H}}[n-1]\hat{\boldsymbol{G}}_{ll}[n-1]\right)^{-1}\right]_{kk}$, $\left\|\left[\boldsymbol{G}_{ll}^{\dagger}[n-1]\right]_k\right\|^{-2}$ has an Erlang distribution with shape parameter $N-K+1$ and scale parameter $\hat{\beta}_{llk}$ [16]. Then,

$$\left\|\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k\right\|^{-2} \overset{\mathrm{d}}{\sim} X_k[n-1]. \tag{17}$$

Furthermore, conditioned on $\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k$, $\hat{\boldsymbol{Y}}_i[n]$ is a zero-mean complex Gaussian vector with covariance matrix $\tilde{\boldsymbol{D}}_{li}$ which is independent of $\left[\hat{\boldsymbol{G}}_{ll}^{\dagger}[n-1]\right]_k$. Hence,

$\sum_{i=1}^{L}\left\|\hat{\boldsymbol{Y}}_i[n]\right\|^2$ is the sum of KL statistically independent but not necessarily identically distributed exponential RVs. According to [17, Theorem 2], we obtain

$$\sum_{i=l}^{L}\left\|\hat{\boldsymbol{Y}}_i[n]\right\|^2 \overset{\mathrm{d}}{\sim} Y_k[n]. \tag{18}$$

Combining (16)–(18), we deduce (13). ∎

Remark 1: The PDF of the uplink SINR (13) accounts for both the effects of pilot contamination and Doppler shift. More specifically, the time variation of the channel reduces both the desired and interference signal powers by a factor of α^2, thus, degrading the SINR.

Corollary 1: Consider the high uplink power regime. We have

$$\lim_{p_{\mathrm{u}}\to\infty}\gamma_k \overset{\mathrm{d}}{\sim} \frac{\alpha^2 X_k[n-1]}{\alpha^2 C X_k[n-1]+Y_k[n]}. \tag{19}$$

This corollary shows that when p_{u} increases asymptotically, there is a finite SINR ceiling due to the simultaneous increase of the desired signal power as well as of the interference and channel estimation error powers.

From Proposition 1, the uplink ergodic rate from the kth user in the lth cell to its BS (in bits/s/Hz) is given by

$$\begin{aligned}
\langle R_{l,k}\rangle &= \mathbb{E}_{X_k,Y_k}\left\{\log_2\!\left(1+\frac{p_{\mathrm{u}}\alpha^2 X_k[n-1]}{p_{\mathrm{u}}\alpha^2 C X_k[n-1]+p_{\mathrm{u}}Y_k[n]+1}\right)\right\}\\
&= \int_0^\infty\!\!\int_0^\infty \log_2\!\left(1+\frac{p_{\mathrm{u}}\alpha^2 x}{p_{\mathrm{u}}\alpha^2 Cx+p_{\mathrm{u}}y+1}\right) p_{X_k}(x)\, p_{Y_k}(y)\, dxdy\\
&= \sum_{p=1}^{\varrho(\mathcal{A}_k)}\sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k)\,\mu_{k,p}^{-q}\log_2 e}{(q-1)!\,(N-K)!\hat{\beta}_{llk}^{N-K+1}}\\
&\quad\times\int_0^\infty\!\!\int_0^\infty \ln\!\left(1+\frac{p_{\mathrm{u}}\alpha^2 x}{p_{\mathrm{u}}\alpha^2 Cx+p_{\mathrm{u}}y+1}\right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}}\, dxdy\\
&= \sum_{p=1}^{\varrho(\mathcal{A}_k)}\sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k)\,\mu_{k,p}^{-q}\log_2 e}{(q-1)!\,(N-K)!\hat{\beta}_{llk}^{N-K+1}}\\
&\quad\times\Bigg(\underbrace{\int_0^\infty\!\!\int_0^\infty \ln\!\left(1+\frac{p_{\mathrm{u}}\alpha^2(C+1)x}{p_{\mathrm{u}}y+1}\right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}}\, dxdy}_{\triangleq\mathcal{I}_1}\\
&\quad\underbrace{-\int_0^\infty\!\!\int_0^\infty \ln\!\left(1+\frac{p_{\mathrm{u}}\alpha^2 Cx}{p_{\mathrm{u}}y+1}\right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}}\, dxdy}_{\triangleq\mathcal{I}_2}\Bigg)\\
&= \sum_{p=1}^{\varrho(\mathcal{A}_k)}\sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k)\,\mu_{k,p}^{-q}\log_2 e}{(q-1)!\,(N-K)!\hat{\beta}_{llk}^{N-K+1}}\left(\mathcal{I}_1-\mathcal{I}_2\right),
\end{aligned} \tag{20}$$

where $\mathcal{I}_1$ and $\mathcal{I}_2$ are given by (21) and (22) shown at the top of next page, while $\mathcal{I}_{m,n}(a,b,\alpha)$ is given by (23), $\mathrm{Ei}(\cdot)$ is the exponential integral function [18, Eq. (8.211.1)], and $U(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function of the second kind [18, Eq. (9.210.2)].

The derivations of $\mathcal{I}_1$ is as follows: we first compute the integral over x by using [18, Eq. (4.337.5)]. Then, by using [22, Eq. (19)] together with [20, Eq. (39)], we obtain (21). Similarly, we obtain $\mathcal{I}_2$, given in (22).

B. Large Antenna Limit Analysis

In this section, we consider the large system limit by accounting for specific assumptions. More precisely, we assume that $N\to\infty$, while p_{u} and K are fixed. The purpose of this analysis is to exploit the reduction of the interference and thermal noise due to the property of orthogonal channels vectors between the BS and the users as $N\to\infty$, as well as to achieve increase of the sum-rate, since it depends on N.

Keeping in mind that an Erlang distribution with shape and scale parameters given by $N-K+1$ and $\hat{\beta}_{llk}$, respectively, can be related with the sum of the squares of independent normal random variables $W_1[n-1], W_2[n-1], ..., W_{2(N-K+1)}[n-1]$, $X_k[n-1]$ is given by

$$X_k[n-1] = \frac{\hat{\beta}_{llk}}{2}\sum_{i=1}^{2(N-K+1)} W_i^2[n-1]. \tag{24}$$

Hence, the substitution of (24) into (13) as well as the use of the law of large numbers give (25), shown at the top of next page. This can be explained by noting that the nominator and the first term of the denominator converge almost surely to $\alpha^2 p_{\mathrm{u}}\hat{\beta}_{llk}/2$ and $\alpha^2 p_{\mathrm{u}} C\hat{\beta}_{llk}/2$ as $N\to\infty$, while the remaining terms of the denominator go to 0. If fact, the bounded SINR is expected because it is well-known that as the number of BS antennas tends to infinity, both the intra-cell interference and noise are cancelled out, while the inter-cell interference due to pilot contamination persists [3], [5].

IV. Numerical Results

In this section, we present numerical results to verify our analysis by considering a cellular network with $L=7$ cells and $K=10$ users per cell. The coherence interval is $T=200$ symbols and the length of training duration is $\tau=K$ symbols. Regarding the large-scale coefficients β_{lik}, we assume a simple scenario: $\beta_{llk}=1$ and $\beta_{lik}=a$, for $k=1,\ldots,K$, and $i\neq l$. Note that a can be considered as an intercell interference factor, while we define $\mathsf{SNR}\triangleq p_{\mathrm{u}}$.

In the following, we will examine the sum spectral efficiency which is defined as:

$$\mathcal{S}_l \triangleq \left(1-\frac{\tau}{T}\right)\sum_{k=1}^{K}\langle R_{l,k}\rangle, \tag{26}$$

where $\langle R_{l,k}\rangle$ is given by (20).

Figure 1 shows the sum spectral efficiency versus the SNR for $N=20$, 50, and 100, at the intercell interference factor $a=0.1$, and the temporal correlation parameter $\alpha=0.9$. The "analysis" curves are computed by using (20), while the "simulation" curves are obtained via (8) using Monte-Carlo simulations. The exact match between the analytical and simulated results validates our analysis. Furthermore, we can see that, at high SNR, the sum spectral efficiency saturates.

$$\mathcal{I}_1 = \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}}(C+1)}} \mathcal{I}_{q-1,N-K-t} \left(\frac{1}{\hat{\beta}_{llk}\alpha^2 (C+1)}, \frac{1}{\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}} (C+1)}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{llk}\alpha^2 (C+1)} \right) \right.$$
$$\left. + \sum_{u=1}^{N-K-t} \frac{(u-1)!\,(-1)^u\, p_{\mathrm{u}}^{-q}}{\left(\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}} (C+1)\right)^{N-K-t-u}} \Gamma(q)\, U\left(q, q+1+N-K-t-u, \frac{1}{\mu_{k,p} p_{\mathrm{u}}}\right) \right]. \tag{21}$$

$$\mathcal{I}_2 = \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}} C}} \mathcal{I}_{q-1,N-K-t} \left(\frac{1}{\hat{\beta}_{llk}\alpha^2 C}, \frac{1}{\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}} C}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{llk}\alpha^2 C} \right) \right.$$
$$\left. + \sum_{u=1}^{N-K-t} \frac{(u-1)!\,(-1)^u\, p_{\mathrm{u}}^{-q}}{\left(\hat{\beta}_{llk}\alpha^2 p_{\mathrm{u}} C\right)^{N-K-t-u}} \Gamma(q)\, U\left(q, q+1+N-K-t-u, \frac{1}{\mu_{k,p} p_{\mathrm{u}}}\right) \right]. \tag{22}$$

$$\mathcal{I}_{m,n}(a,b,\alpha) \triangleq \sum_{r=0}^{m} \binom{m}{r} (-b)^{m-r} \left[\sum_{s=0}^{n+r} \frac{(n+r)^s\, b^{n+r-s}}{\alpha^{s+1} a^{m-s}} \mathrm{Ei}(-b) - \frac{(n+r)^{n+r}\, e^{\alpha b/a}}{\alpha^{n+r+1} a^{m-n-r}} \mathrm{Ei}\left(-\frac{\alpha b}{a} - b\right) \right.$$
$$\left. + \frac{e^{-b}}{\alpha} \sum_{s=0}^{n+r-1} \sum_{u=0}^{n+r-s-1} \frac{u!\,(n+r)^s \binom{n+r-s-1}{u} b^{n+r-s-u-1}}{\alpha^s a^{m-s} (\alpha/a+1)^{s+1}} \right]. \tag{23}$$

$$\lim_{N\to\infty} \gamma_k = \frac{\alpha^2 p_{\mathrm{u}} \frac{\hat{\beta}_{llk}}{2} \sum_{i=1}^{2(N-K+1)} W_i^2[n-1] / \left(2(N-K+1)\right)}{\left(\alpha^2 p_{\mathrm{u}} C \frac{\hat{\beta}_{llk}}{2} \sum_{i=1}^{2(N-K+1)} W_i^2[n-1] + p_{\mathrm{u}} \sum_{i=1}^{L} \left\| \hat{\boldsymbol{Y}}_i[n] \right\|^2 + 1\right) / \left(2(N-K+1)\right)} \xrightarrow{\text{a.s.}} \frac{1}{C}, \tag{25}$$

This is due to the fact that when SNR increases, both the desired signal power and intercell interference power are increased. To improve the system performance, we can use more antennas at the BS. At $\mathsf{SNR} = 5$dB, the sum spectral efficiencies can be increased by the factors of 2.5 or 5.5 if we increase N from 20 to 50 or from 20 to 100, respectively.

Next, we examine the effect of the temporal correlation parameter on the system performance. Figure 2 presents the sum spectral efficiency as a function of the temporal correlation parameter α, at $a = 0.1$ and $\mathsf{SNR} = 0$dB, for $N = 20, 50$, and 100. We can see that the system performance degrades significantly when the temporal correlation parameter decreases (or the time variation in the channel increases). Half of the spectral efficiency is reduced when α reduces from 1 to 0.6. Furthermore, at low α, using more antennas at the BS does not help improve the system performance much.

Finally, Fig. 3 shows the transmit power, p_{u}, that is needed to reach 1 bit/s/Hz per user. Here, we choose $a = 0.1$ and $\alpha = 0.7$ or 0.9. As expected, the required transmit power decreases significantly when we increase the number of BS antennas. By doubling the number of BS antennas, we can cut back the transmit power by at least 1.5dB. This observation is in line with the results of [5].

V. Conclusions

In this paper, we characterized the uplink performance of a cellular network taking into account both the well-known pilot contamination and the unavoidable, but less studied, time variation. The latter effect stems from user mobility which occurs in the vast majority of propagation scenarios. The main contribution of this work includes a new analytical expression for the achievable sum rate, that holds for any finite number of BS antennas. Furthermore, an asymptotic expression in the large antenna regime was also derived. Our numerical illustrations depicted how the time variation affects the performance, and verified our analysis.

References

[1] "5G: A Technology Vision," Huawei Technologies Co., Ltd., Shenzhen, China, Whitepaper Nov. 2013. [Online]. www.huawei.com/ilink/en/download/HW_314849

[2] D. Gesbert, M. Kountouris, R. W. Heath Jr., C. B. Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Sig. Proc. Mag.,* vol. 24, no. 5, pp. 36–46, 2007.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

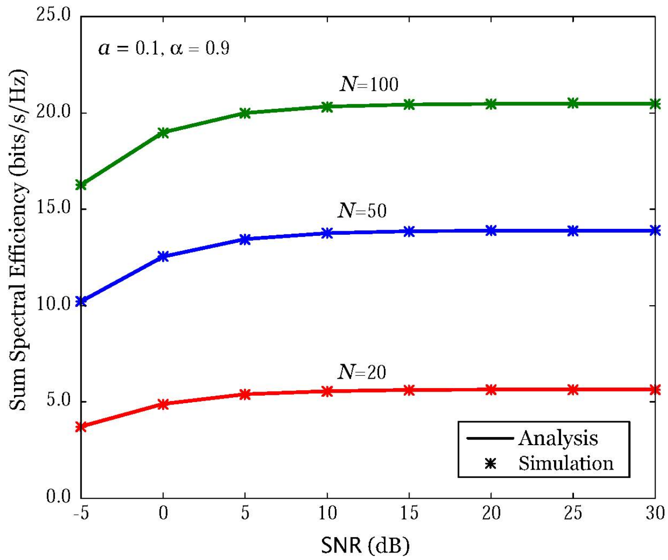

Fig. 1. Sum spectral efficiency versus SNR for different $N(a = 0.1$ and $\alpha = 0.9)$.

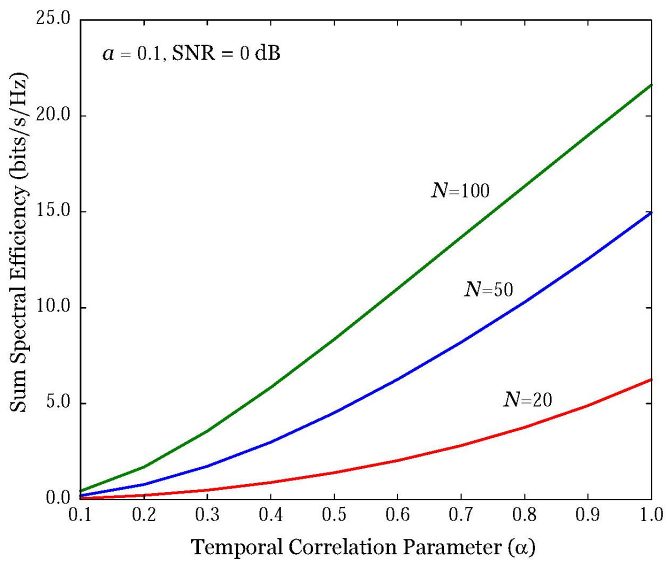

Fig. 2. Sum spectral efficiency versus α for different $N(a = 0.1$ and SNR $=$ 0 dB).

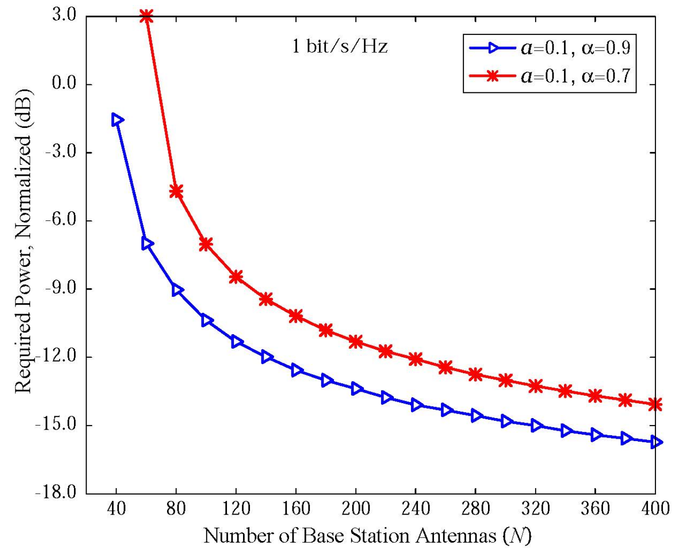

Fig. 3. Transmit power required to achieve 1 bit/s/Hz per user versus $M(a = 0.1,\ \alpha = 0.7$ and $0.9)$.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[6] C. J. Chen and L. C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.,* vol. 25, no. 7, pp. 1435–1445, Sep. 2007.

[7] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory,* vol. 49, no. 7, pp. 1691–1706, Jul. 2006.

[8] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.,* vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[9] K. T. Truong and R. W. Heath, Jr., "Effects of channel aging in massive MIMO systems," *IEEE/KICS J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[10] A. Papazafeiropoulos and T. Ratnarajah "Linear precoding for downlink massive MIMO with delayed CSIT and channel prediction" in *Proc. IEEE WCNC 2014*, Apr. 2014, pp. 821–826.

[11] A. Papazafeiropoulos and T. Ratnarajah "Uplink performance of massive MIMO subject to delayed CSIT and anticipated channel prediction," in *Proc. IEEE ICASSP 2014*, May. 2014, pp. 3186–3189.

[12] H. Q. Ngo, M. Matthaiou, T. Q. Duong, and E. G. Larsson, "Uplink performance analysis of multicell MU-SIMO systems with ZF receivers," *IEEE Trans. Veh. Tech.*, vol. 62, no. 9, pp. 4471–4483, Nov. 2013.

[13] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.

[14] H. Q. Ngo, M. Matthaiou, and E. G. Larsson. "Performance analysis of large scale MU-MIMO with optimal linear receivers," *in Proc. IEEE Swe-CTW*, Oct. 2012, pp. 59–64.

[15] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008.

[16] D. A. Gore, R. W. Heath Jr., and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.

[17] A. Bletsas, H. Shin, and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, Sep. 2007.

[18] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2007.

[19] Wolfram, "The Wolfram functions site." Available: http://functions.wolfram.com

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

[21] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series, Volume 3: More Special Functions*. New York: Gordon and Breach Science, 1990.

[22] H. Q. Ngo, T. Q. Duong, and E. G. Larsson, "Uplink performance analysis of multicell MU-MIMO with zero-forcing receivers and perfect CSI," in *Proc. IEEE Swe-CTW*, Sweden, Nov. 2011, pp. 40–45.