

LMMSE Receivers in Uplink Massive MIMO Systems with Correlated Rician Fading

Ikram Boukhedimi, Abla Kammoun, and Mohamed-Slim Alouini
Computer, Electrical, and Mathematical Sciences and Engineering Division,
King Abdullah University of Science and Technology (KAUST),
Thuwal, Makkah Province, Kingdom of Saudi Arabia.
Email: {ikram.boukhedimi, abla.kammoun, slim.alouini}@kaust.edu.sa

Abstract—We carry out a theoretical analysis of the uplink (UL) of a massive MIMO system with per-user channel correlation and Rician fading, using two processing approaches. Firstly, we examine the linear-minimum-mean-square-error receiver under training-based imperfect channel estimates. Secondly, we propose a statistical combining technique that is more suitable in environments with strong Line-of-Sight (LoS) components. We derive closed-form asymptotic approximations of the UL spectral efficiency (SE) attained by each combining scheme in single and multi-cell settings, as a function of the system parameters. These expressions are insightful in how different factors such as LoS propagation conditions and pilot contamination impact the overall system performance. Furthermore, they are exploited to determine the optimal number of training symbols which is shown to be of significant interest at low Rician factors. The study and numerical results substantiate that stronger LoS signals lead to better performances, and under such conditions, the statistical combining entails higher SE gains than the conventional receiver.

Index Terms—Massive MIMO, correlated Rician fading, imperfect channel estimation, optimal training, LMMSE combining.

I. INTRODUCTION

Future 5G networks are expected to support substantial amounts of mobile data traffic and ensure a reliable quality of service to the end users [1]–[3]. Over the past couple of years, massive MIMO has been extensively investigated, (see [1]–[7] and references therein). In general, most of these works consider Rayleigh fading channels that model scattered signals, yet do not encompass the possibility of a Line-of-Sight (LoS) component which is commonly present in practical wireless propagation scenarios and modeled by Rician-fading. At the same time, in order to meet the 5G performance demands, massive MIMO is expected to be omnipresent and thus, all propagation conditions ought to be examined. This is the case for indoor applications or small areas operating over mmWave communications wherein the presence of LoS is conceivable [8]. In fact, active research is being conducted to study the pairing of massive MIMO systems with mmWave communications to jointly reap their benefits [9]–[11]. Accordingly, understanding the performance of massive MIMO systems operating under LoS conditions is of growing importance and is the focus of the present work.

The main literature related to this work is represented by [12]–[26]. The authors in [12] investigate the power scaling laws and the uplink (UL) rates using zero-forcing (ZF) and maximum-ratio (MR) combiners, however assuming uncorrelated channels.

In [14], an analytical study of the rates of the downlink (DL) of a MU-MIMO system assuming ZF beamforming and uncorrelated Rician fading channels is performed. Conversely, in [16], the authors use tools from random matrix theory (RMT) to conduct an asymptotic analysis of the DL of a spatially correlated MIMO Rician fading model, yet under the assumption of perfect channel state information at the base station (BS). A similar large system analysis relying on some recent RMT results is led in [19]. In this latter, the authors investigate regularized-ZF and MR-transmit schemes in the DL of a large-scale MIMO system under uncorrelated Rician fading system and assuming imperfect channel estimates. Moreover, in the same line, the authors in [21] use similar tools to analyze the ergodic UL rates of a massive MIMO system and determine the optimal fraction of the coherence time used for channel training in a Rayleigh fading setting. The work in [25] analyzes transmit and receive LoS-based beamforming designs which treat the scattered signals as interference in Rician fading massive MIMO systems. Finally, the work [26] recently appeared after our submission, and it focuses on correlated Rician massive MIMO systems using the less involved MR combining and precoding scheme.

In this paper, we investigate the UL performance of a massive MIMO system wherein every user is allotted a distinct channel correlation matrix and Rician factor, and above all, assuming imperfect CSI, and ultimately, pilot contamination. Furthermore, we consider two different combining schemes. The first method consists in the involved Linear-Minimum-Mean-Square-Error (LMMSE) receiver which we will refer to as the ‘conventional’ receiver in the sequel. Note that LMMSE’s performance underlying such an intricate system model render the analysis comprehensive and unprecedented as, to the best of our knowledge, it has not been conducted thus far. As to the second technique, we propose a ‘statistical’ combiner that only utilizes the long-term statistics of the channels like the Rician factors, spatial correlation matrices, etc. In essence, this technique is purposely designed for LoS-prevailing environments to circumvent training and channel estimation and associated errors, and exploit the presence of LoS components in a more efficient manner.

We first consider a single-cell network with the aforementioned comprehensive channel model, and carry out a theoretical

A part of this paper has been presented at the IEEE ICC18, May 2018, Kansas City, MO, USA.

analysis of the achievable UL spectral efficiency (SE) for each receiver. Assuming the large antenna-limit with a fixed number of users, we harness some rudimentary asymptotic tools such as the law of large numbers (LLN) and convergence of quadratic forms [27], to derive closed-form approximations of the SEs. These approximations provide insights on the impact of the system parameters such as the training sequence's length, the Rician factor as well as the propagation conditions on the overall performances. Furthermore, we exploit them to determine the optimal number of training symbols that maximizes the UL SE whose value is shown to be particularly crucial at low Rician factors. A relevant outcome of the study reveals that high Rician factors generate far better performances, and that in such environments, longer training sequences are rather counterproductive since they degrade the achievable SEs. This result, led us to propose the statistical receiver for systems with strong specular signals. As will be elaborated later, this scheme is obtained through the maximization of the corresponding UL SE and proven to outperform the conventional receiver in LoS-prevailing systems. Note that in the conference paper [28], we present the preliminary results of this single-cell setting.

For a more thorough analysis, we extend our study to a multi-cell system in order to examine the performances of both combining schemes when they are subject to inter-cell interference, and especially to pilot contamination. Similarly to the single-cell, we derive closed-form approximations of the achievable SEs under the asymptotic antenna regime. Accordingly, the impacts of LoS propagation conditions and pilot-contamination-induced interference are meticulously analyzed for both receivers. Ultimately, the discussion highlights interesting aspects on the interplay between the Rician factors and pilot contamination. Additionally, it unveils that in a multi-cell setup, the proposed combiner outperforms the conventional one to an even higher extent than it is the case for a single-cell system. Evidently, numerical results are provided to validate the accuracy of our analytical findings and better illustrate the efficiency of both schemes in all settings for finite system dimensions, albeit computed in the asymptotic regime.

The remainder of the paper is organized as follows. Section II encompasses the single-cell system starting by the corresponding UL system model, followed by the detection schemes, and finally the theoretical analysis of the achievable spectral efficiencies that shed light on some interesting aspects. Then, pursuing a similar rationale as in Section II, Section III focuses on the multi-cell systems. After that, Section IV consists of a selection of numerical results that confirm the theoretical derivations given in Sections II and III. Finally, conclusions are drawn in Section V.

II. PERFORMANCE ANALYSIS IN A SINGLE-CELL SETTING

We consider uplink transmissions of a TDD single-cell system with K mono-antenna users (UEs) and a BS equipped with N antennas. Assuming Gaussian codebooks, the vector of the transmitted data symbols sent by all the UEs is denoted $\sqrt{\frac{p_u}{N}} \mathbf{x} \sim \mathcal{CN}(0, \frac{p_u}{N} \mathbf{I}_K)$ and therefore, the received signal at

the BS writes:

$$\mathbf{y} = \sqrt{\frac{p_u}{N}} \mathbf{H} \mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]$ is the $N \times K$ aggregated MIMO channel matrix from all UEs to the BS and \mathbf{n} represents a zero-mean additive Gaussian noise with variance σ^2 . Correlated Rician fading channels are considered such that the channel between the k -th UE and the BS is modeled as:

$$\mathbf{h}_k = \sqrt{\beta_k} \left(\sqrt{\frac{1}{1 + \kappa_k}} \Theta_k^{\frac{1}{2}} \mathbf{z}_k + \sqrt{\frac{\kappa_k}{1 + \kappa_k}} \bar{\mathbf{z}}_k \right), \quad (2)$$

where β_k accounts for the large-scale channel fading of UE $_k$ and the second term represents the small-scale fading channel. This latter consists of the Rayleigh component $\mathbf{z}_k \sim \mathcal{CN}(0, \mathbf{I}_N)$ to depict scattered or Non-LoS signals and the deterministic component $\bar{\mathbf{z}}_k \in \mathbb{C}^N$ to represent the specular (LoS) signals. For each UE, the ratio between these components is depicted by the Rician factor κ_k . Plus, for each UE k , we consider a different channel correlation matrix Θ_k . Throughout the paper, $\forall k$, Θ_k is assumed to be slowly varying compared to the channel coherence time and thus is supposed to be perfectly known to the BS. Finally, for notational convenience, we let: $\bar{\mathbf{h}}_k = \sqrt{\frac{\beta_k \kappa_k}{1 + \kappa_k}} \bar{\mathbf{z}}_k$ and $\mathbf{R}_k = \frac{\beta_k}{1 + \kappa_k} \Theta_k$. Therefore, $\mathbf{h}_k \sim \mathcal{CN}(\bar{\mathbf{h}}_k, \mathbf{R}_k)$.

A. Channel Estimation

In practice, prior to processing the received signal, the BS estimates the channel matrix \mathbf{H} . Let $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \dots, \hat{\mathbf{h}}_K]$ denote the aggregate matrix of these estimates. In TDD systems, each UL channel coherence block of length T is split into two phases starting by training and followed with data transmission. In the pilot training interval of $\tau \geq K$ symbols, all K UEs broadcast orthogonal sequences of known pilot symbols with average power τp_p . It is important to note that in the considered Rician fading, since the specular components are hardly changing, it is reasonable to assume that both the LoS component and Rician factors of all UEs are known to both the transmitter and receiver. Accordingly, using single-cell LMMSE estimation, the estimate $\hat{\mathbf{h}}_k$ of the channel \mathbf{h}_k is given by [4] :

$$\hat{\mathbf{h}}_k = \mathbf{R}_k \Phi_k (\mathbf{y}_k^{tr} - \bar{\mathbf{h}}_k) + \bar{\mathbf{h}}_k, \quad (3)$$

where, $\Phi_k = \left(\mathbf{R}_k + \frac{1}{\tau \rho_{tr}} \mathbf{I}_N \right)^{-1}$, $\mathbf{y}_k^{tr} = \mathbf{h}_k + \frac{1}{\sqrt{\tau \rho_{tr}}} \mathbf{n}_k^{tr}$, and $\rho_{tr} = \frac{p_p}{\sigma^2}$ is the SNR corresponding to the training phase. The higher value $\tau \rho_{tr}$ takes, the better quality of channel estimation becomes. In fact, as $\tau \rho_{tr} \rightarrow \infty$, $\hat{\mathbf{H}} \rightarrow \mathbf{H}$ which corresponds to the perfect CSI scenario. From (3), it can be shown that $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\bar{\mathbf{h}}_k, \tilde{\mathbf{R}}_k)$, with $\tilde{\mathbf{R}}_k = \mathbf{R}_k \Phi_k \mathbf{R}_k$. Plus,

Notations: In the sequel, bold upper and lowercase characters refer to matrices and vectors, respectively. We also use $(\cdot)^H$, $\text{tr}(\cdot)$ and $(\cdot)^{-1}$ to denote the conjugate transpose, the trace of a matrix and the inverse operations, respectively. $\log(\cdot)$ is the natural logarithm, the $N \times N$ identity matrix is denoted \mathbf{I}_N , and $\delta_{j\ell}$ is the Kronecker delta. Finally, $[\mathbf{A}]_{ij}$ is the element on the i -th row and j -th column of matrix \mathbf{A} and $\text{diag}\{a_i\}_{i=1}^N$ is the $N \times N$ diagonal matrix with a_i being its i -th diagonal element.

considering the orthogonality property of LMMSE estimation, the estimation error $\xi_k = \mathbf{h}_k - \hat{\mathbf{h}}_k$ follows the distribution $\mathcal{CN}(0, \mathbf{R}_k - \tilde{\mathbf{R}}_k)$.

B. Detection and Achievable Uplink Spectral Efficiency

To process the signal \mathbf{y} (1), the BS uses a linear receiver. In this work, we are interested in the conventional LMMSE receiver that relies on acquired channel estimates, and propose a statistical receiver that is mainly based on the long-term parameters of the system.

1) *Conventional LMMSE Receiver*: Let $\mathbf{g}_k \in \mathbb{C}^{N \times 1}$ denote the conventional combining vector used to process the signal sent by UE k . Under imperfect channel estimation conditions, \mathbf{g}_k is defined as [4], [29]:

$$\mathbf{g}_k = \left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \sum_{i=1}^K (\mathbf{R}_i - \tilde{\mathbf{R}}_i) + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_k. \quad (4)$$

In order to retrieve useful data, the BS generates the signal $\mathbf{r} = \mathbf{G}^H \mathbf{y}$. As shown in (4), the design of \mathbf{g}_k leverages the channel estimates $\hat{\mathbf{H}}$, thus making the performances sensitive to channel estimation errors. Therefore, the k -th element of \mathbf{r} can be decomposed as:

$$r_k = \sqrt{\frac{p_u}{N}} \sum_{i=1}^K \mathbf{g}_k^H \hat{\mathbf{h}}_i x_i + \sqrt{\frac{p_u}{N}} \sum_{i=1}^K \mathbf{g}_k^H \xi_i x_i + \mathbf{g}_k^H n_k, \quad (5)$$

This expression respectively, separates the signal and intra-cell interference, channel estimation errors, and noise terms. Additionally, when a pre-training phase of τ symbols is performed, only a fraction of the total coherence block is used for useful data transmission. Therefore, denoting $\rho_d = \frac{p_u}{\sigma^2}$, the UL achievable SE for UE k , in case of channel-estimate based conventional processing is defined as [30]:³

$$\text{SE}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T} \right) \mathbb{E} \left[\log \left(1 + \frac{|\mathbf{g}_k^H \hat{\mathbf{h}}_k|^2}{\mathbb{E} \left[\sum_{\substack{i=1 \\ i \neq k}}^K |\mathbf{g}_k^H \hat{\mathbf{h}}_i|^2 + \sum_{i=1}^K |\mathbf{g}_k^H \xi_i|^2 + \frac{N}{\rho_d} \|\mathbf{g}_k\|^2 \right]} \right) \right]. \quad (6)$$

2) *Statistical LMMSE Receiver*: Due to its slow varying pace, the LoS component $\bar{\mathbf{H}} = [\bar{\mathbf{h}}_1, \dots, \bar{\mathbf{h}}_K]$ can be easily estimated. For example, the BS may estimate the specular signals in a previous transmission from the UEs, in contrast to the Rayleigh signals which must be estimated at every T . In addition, choosing the right number of training symbols, τ , is paramount to ensure the overall UL performances since a small τ entails significant estimation errors and a larger τ suggests less transmitted data. Motivated by these factors, we propose in this work a statistical receiver denoted $\bar{\mathbf{g}}_k$, that exclusively exploits the presence of the quasi-deterministic

³In the sequel, we add the superscripts $(\cdot)^{\text{conv},S}$, $(\cdot)^{\text{stat},S}$, $(\cdot)^{\text{conv},M}$ and $(\cdot)^{\text{stat},M}$ to, respectively, distinguish the relevant quantities corresponding to conventional combining and statistical combining, in single-cell and multi-cell schemes.

LoS component $\bar{\mathbf{H}}$ and the long-term parameters of the system, such as the spatial correlation matrices \mathbf{R}_k , the large-scale fading factors β_k , etc. Naturally, using such a receiver enables to avoid training and channel estimation altogether, thereby yielding the single-cell UL SE $\text{SE}_k^{\text{stat},S}$:

$$\text{SE}_k^{\text{stat},S} = \mathbb{E} \left[\log \left(1 + \frac{|\bar{\mathbf{g}}_k^H \bar{\mathbf{h}}_k|^2}{\mathbb{E} \left[\bar{\mathbf{g}}_k^H \left(\sum_{i=1}^K \mathbf{h}_i \mathbf{h}_i^H - \bar{\mathbf{h}}_k \bar{\mathbf{h}}_k^H + \frac{N}{\rho_d} \mathbf{I}_N \right) \bar{\mathbf{g}}_k \right]} \right) \right]. \quad (7)$$

We propose to design $\bar{\mathbf{g}}_k$ through the maximization of a deterministic equivalent of $\text{SE}_k^{\text{stat},S}$ in the infinite antenna limit which we denote $\overline{\text{SE}}_k^{\text{stat},S}$. Specifically, for $k = 1, \dots, K$, $\bar{\mathbf{g}}_k$ is defined as:

$$\begin{aligned} \bar{\mathbf{g}}_k &= \underset{\bar{\mathbf{g}}_k}{\text{argmax}} \overline{\text{SE}}_k^{\text{stat},S}, \\ \text{s.t. } \overline{\text{SE}}_k^{\text{stat},S} - \overline{\text{SE}}_k^{\text{stat},S} \xrightarrow[N \rightarrow \infty]{a.s.} 0. \end{aligned} \quad (\text{P1})$$

As shall be seen in the next section, on account that $\bar{\mathbf{g}}_k$ is deterministic, one should note that $\overline{\text{SE}}_k^{\text{stat},S}$ is obtained by means of quite rudimentary asymptotic tools.

C. Asymptotic Analysis of the Single-Cell Performances

In this section, we carry out a comparative theoretical analysis between the UL performances achieved by the conventional receiver, \mathbf{g}_k (4), and the proposed statistical combiner, $\bar{\mathbf{g}}_k$ (P1). The study is conducted under the assumption of imperfect channel state information, and a distinct Rician factor as well as channel correlation per user. Ultimately, the objective is to determine conditions in which the statistical receiver outperforms the conventional one. Nonetheless, as can be seen from $\text{SE}_k^{\text{conv},S}$ (6) and $\text{SE}_k^{\text{stat},S}$ (7), these expressions involve random quantities that are rather compact and do not lend themselves to simple interpretations nor manipulations. Accordingly, we first derive closed-form asymptotic approximations of both $\text{SE}_k^{\text{conv},S}$ and $\text{SE}_k^{\text{stat},S}$ which we exploit thereafter for the comparison. To obtain these approximations, we consider the large-antenna limit with a fixed number of UEs. This can be formulated as:

Assumption 1. We assume that K is fixed while N grows large without bound. We also consider that as $N \rightarrow \infty$, $\forall k$, the channel correlation matrix has a bounded spectral norm $\|\mathbf{R}_k\|_2$. For simplicity, this asymptotic regime will be denoted by \mathcal{H}^{∞} . *Conventional Combining in Single-Cell Systems*: Define the matrices:

$$\begin{aligned} \mathbf{Q} &= \left(\frac{1}{N} \bar{\mathbf{H}}^H \bar{\mathbf{H}} + \frac{1}{N} \text{diag} \left\{ \text{tr} \tilde{\mathbf{R}}_\ell \right\}_{\ell=1}^K + \frac{1}{\rho_d} \mathbf{I}_K \right)^{-1}, \\ \mathbf{T}_i &= \bar{\mathbf{H}}^H \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \bar{\mathbf{H}} + \text{diag} \left\{ \text{tr} \left(\tilde{\mathbf{R}}_\ell \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \right) \right\}_{\ell=1}^K, \end{aligned} \quad (8)$$

and let \mathbf{q}_k be the k -th column of the matrix \mathbf{Q} .

Theorem 1 (Conventional combining in single-cell systems). Under Assumption 1, we have : $\text{SE}_k^{\text{conv},S} - \overline{\text{SE}}_k^{\text{conv},S} \xrightarrow[N \rightarrow \infty]{a.s.} 0$,

such that $\overline{\text{SE}}_k^{\text{conv},S}$ is given by (10) : (in the top of the next page).

Proof: A proof is given in Appendix A \blacksquare

We provide in the closed-form expression (10) approximations of all the different terms constituting $\overline{\text{SE}}_k^{\text{conv},S}$. This allows to have some insights on the behavior of these signals and their impact on the achievable SE. Note nonetheless that further simplifications can be made in the infinite antenna limit. For instance, we can see that as N grows infinitely large and for a fixed $K : \frac{1}{N^2} \sum_{i=1}^K \mathbf{q}_k^H \mathbf{T}_i \mathbf{q}_k \xrightarrow[N \rightarrow \infty]{a.s.} 0$, therefore implying that channel estimation errors vanish in the UL massive MIMO setting. Accordingly, $\overline{\text{SE}}_k^{\text{conv},S}$ (10) amounts to:

$$\overline{\text{SE}}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log \left(\frac{\rho_d}{|\mathbf{Q}|_{kk}} \right) + \mathcal{O} \left(\frac{1}{N} \right). \quad (11)$$

Another key point in $\overline{\text{SE}}_k^{\text{conv},S}$ (10), is the term $\sum_{i \neq k}^K \left| \frac{1}{\rho_d} [\mathbf{Q}]_{ki} \right|^2$, which represents an approximation of intra-cell interference. If Rayleigh fading is considered (*i.e.* $\overline{\mathbf{H}} = \mathbf{0}_{N \times K}$), this term cancels out. Pursuant to [4], this result confirms that in the setting $N \rightarrow \infty$ while K is fixed, intra-cell interference due to Rayleigh fading dissipates. However, as we can see in this paper, in Rician fading, the specular signals generate intra-cell interference which is embodied by the inner products between $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{h}}_i$, ($i, k, 1, \dots, K$). In light of this outcome, one way to eliminate interference is to have LoS components that are mutually orthogonal between users. This circumstance can be accomplished under asymptotic favorable propagation conditions where, $\frac{1}{N} \hat{\mathbf{h}}_i^H \hat{\mathbf{h}}_j \xrightarrow[N \rightarrow \infty]{a.s.} 0$, $i \neq j$, therefore yielding $\sum_{i \neq k}^K \left| \frac{1}{\rho_d} [\mathbf{Q}]_{ki} \right|^2 \xrightarrow[N \rightarrow \infty]{a.s.} 0$. Hence, with the elimination of interference, we can conclude that for Rician fading, better performances are achieved in favorable propagation environments, specifically:

Corollary 1 (Favorable Propagation). *if $\frac{1}{N} \hat{\mathbf{h}}_i^H \hat{\mathbf{h}}_j \xrightarrow[N \rightarrow \infty]{a.s.} 0$, for $i \neq j$, we have:*

$$\overline{\text{SE}}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log \left[1 + \frac{\rho_d}{N} \left(\text{tr} \tilde{\mathbf{R}}_k + \|\hat{\mathbf{h}}_k\|^2 \right) \right] + \mathcal{O} \left(\frac{1}{N} \right). \quad (12)$$

Furthermore, under the same settings of corollary 1, we demonstrate in [28], where we compare LMMSE and Matched Filters (MF), that the UL SE (12) is in fact identical to when MF is used. Accordingly, we find that in massive MIMO with Rician fading channels, LMMSE and MF receivers attain comparable performances only under favorable propagation conditions. This is, however, different from Rayleigh fading, wherein similar performances are obtained by the receivers (*see* [1, Eq.(13)] and [4, Remark 3.4]). On another note, from the expression of channel estimates $\hat{\mathbf{h}}_k$ (3), it can be shown by a simple eigenvalue decomposition that for low CSI, (*i.e.* $\tau \rho_{tr} \rightarrow 0$), $\text{tr} \tilde{\mathbf{R}}_i \rightarrow 0$, $\forall i$. In such a case, we can see from the SE expressions $\overline{\text{SE}}_k^{\text{conv},S}$ (11) and (12), that the UL performances degrade with the deterioration of the CSI quality, and $\overline{\text{SE}}_k^{\text{conv},S}$ will be mainly determined by the specular signals. Accordingly, we can state that the strength of the LoS component is peculiarly

beneficial when the channels are poorly estimated. By the same token, having reliable CSI becomes of greater importance as the LoS component weakens. Consequently, good channel estimates highly impact the performances; however, in Rician fading channels, it is of utmost relevance to have a receiver that exploits the presence and strength of the specular signals in an efficient manner.

2) *Optimal Training:* Define $\gamma_k(\tau)$, such that (11) writes: $\overline{\text{SE}}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log(1 + \gamma_k(\tau))$. We determine in the next Theorem the optimal value τ^* that maximizes the achievable average SE. The objective is to determine the optimal number of symbols out of the total coherence symbols to be dedicated for training, for a fixed power allocation. Therefore, $\tau^* \geq K$ to preserve orthogonality of the pilot sequences, and evidently, $\tau^* < T$. Accordingly, τ^* is solution to the optimization problem:

$$\tau^* = \underset{\tau}{\text{argmax}} \frac{1}{K} \sum_{k=1}^K \overline{\text{SE}}_k^{\text{conv},S}, \quad (P2)$$

s.t. $K \leq \tau < T$.

Theorem 2 (Optimal training). *Under imperfect channel estimates, the optimal training length is given by :*

• *If:*

$$\frac{1}{K} \sum_{k=1}^K \left[(T - K) \frac{\gamma'_k(K)}{1 + \gamma_k(K)} - \log(1 + \gamma_k(K)) \right] \leq 0, \quad (13)$$

then $\tau^* = K$.

• *Otherwise: τ^* is the solution to the fixed point equation:*

$$\tau^* = T - \frac{\frac{1}{K} \sum_{k=1}^K \log(1 + \gamma_k(\tau))}{\frac{1}{K} \sum_{k=1}^K \frac{\gamma'_k(\tau)}{1 + \gamma_k(\tau)}}, \quad (14)$$

where $\gamma'_k(\tau)$ is the derivative of $\gamma_k(\tau)$ with respect to τ .

Proof: A proof of Theorem 2 is given in Appendix B. \blacksquare

Note that the derivation of τ^* relies on $\overline{\text{SE}}_k^{\text{conv},S}$ considering its explicit form relatively to the intractable alternative, $\text{SE}_k^{\text{conv},S}$. The results of Theorem 2 will be validated by simulations. Nevertheless, they can be exploited to get some insights on the behavioural tendencies of the choice of τ and the overall uplink performances. For instance, an interesting direction is the impact of the Rician factor κ on the choice of τ . In order to investigate this point, we consider the following case study.

Case study: Let us examine the case where if K is kept fixed and N grows without bound: $\frac{1}{N} \hat{\mathbf{h}}_i^H \hat{\mathbf{h}}_j \xrightarrow[N \rightarrow \infty]{a.s.} \frac{\beta_k \kappa_k}{1 + \kappa_k} \delta_{ij}$, where δ_{ij} is the Kronecker delta. Additionally, let $\forall k : \mathbf{R}_k = \frac{\beta_k}{1 + \kappa_k} \mathbf{I}_N$, $\beta_k = \beta$ and $\kappa_k = \kappa$. Thus, according to corollary 1:

$$\gamma(\tau) = \frac{\beta \rho_d}{1 + \kappa} \left(\frac{1}{1 + \kappa + \frac{1}{\tau \rho_{tr}}} + \kappa \right). \quad (15)$$

a) *Low Rician factor:* consider small values of κ

• At a low SNR level (ρ_d approaches 0), the solution (14) can be rewritten as :

$$\lim_{\kappa \rightarrow 0} \tau^* =$$

$$\overline{\text{SE}}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log \left(1 + \frac{\left|1 - \frac{1}{\rho_d} [\mathbf{Q}]_{kk}\right|^2}{\sum_{\substack{i=1 \\ i \neq k}}^K \left|\frac{1}{\rho_d} [\mathbf{Q}]_{ki}\right|^2 + \frac{1}{N^2} \sum_{i=1}^K \mathbf{q}_k^H \mathbf{T}_i \mathbf{q}_k + \frac{1}{\rho_d} \left([\mathbf{Q}]_{kk} - \frac{1}{\rho_d} [\mathbf{Q}^2]_{kk}\right)}\right). \quad (10)$$

$$T - \frac{(1 + \beta \rho_{tr} \tau) (1 + \beta \rho_{tr} \tau + \beta^2 \rho_{tr} \tau \rho_d) \log \left(1 + \frac{\beta^2 \rho_{tr} \tau}{1 + \beta \rho_{tr} \tau}\right)}{\beta^2 \rho_t \rho_d}. \quad (16)$$

Using Taylor's expansion in the low SNR regime yields:

$$\lim_{\kappa \rightarrow 0} \tau^* = \max \left\{ K, \frac{-1 + \sqrt{1 + \beta \rho_{tr} T}}{\beta \rho_{tr}} \right\}. \quad (17)$$

b) *High Rician factor:* For high values of the Rician factor, κ , thus leading to $\gamma'(K) = 0$, and therefore, (13) is always verified, hence :

$$\lim_{\kappa \rightarrow \infty} \tau^* = K. \quad (18)$$

This case study sheds some light on how the optimal number of training symbols depends on the large-scale fading parameters, the number of users, the UL SNR and the coherence interval. The first example represents the case wherein the Rayleigh fading is governing at poor SNR levels. As can be seen from (17), τ^* depends on the system parameters and on the available SNR during training, ρ_{tr} . For instance, if this latter is also low, (17) yields $\lim_{\kappa \rightarrow 0} \tau^* = \max \left\{ K, \frac{T}{2} \right\}$. This result implies that in a network setting where $T > 2K$, to ensure the best performances, half of the total transmitted symbols should be dedicated to training and the other half to useful data. Conversely, if more users are considered such that $T \leq 2K$, then, the optimal number of training symbols should not go beyond the imposed minimum, K . In the second example, as the Rician factor takes higher values, we find that τ^* always approaches K (18). Consequently, in such circumstances, there is no need to perform any optimization since the optimal number of training symbols is limited to the minimum possible value to ensure pilot orthogonality, namely K . More importantly, we deduce that above a certain κ , investing in more training samples is not optimal in terms of spectral efficiency as it will have a minor impact on the achievable UL SE. In fact, it might even induce performance losses when $\tau \gg K$, as shall be illustrated in simulations. This result motivates us to analyse, in the next section, the potential outcomes of employing the statistical receiver $\bar{\mathbf{g}}_k$ (P1) in LoS-prevailing environments.

3) *Statistical Combining in Single-Cell Systems:* As previously mentioned, the objective of having a statistical receiver is to eliminate both training and channel estimation and to exploit the presence of the Rician component efficiently. To this end, the proposed combining vector $\bar{\mathbf{g}}_k$ is obtained through the maximization of a deterministic approximation of the UL spectral efficiency $\overline{\text{SE}}_k^{\text{stat},S}$ (7), as depicted in (P1). Taking into account that $\bar{\mathbf{g}}_k$ is deterministic itself, a direct application of the convergence of quadratic forms lemma [27] and continuous mapping Theorem [31], yields:

Theorem 3 (Statistical combining in single-cell systems). Under Assumption 1, $\overline{\text{SE}}_k^{\text{stat},S} - \overline{\text{SE}}_k^{\text{stat},S} \xrightarrow[N \rightarrow \infty]{a.s.} 0$, with:

$$\overline{\text{SE}}_k^{\text{stat},S} = \log \left[1 + \frac{\bar{\mathbf{g}}_k^H \left(\frac{\bar{\mathbf{h}}_k \bar{\mathbf{h}}_k^H}{N} \right) \bar{\mathbf{g}}_k}{\bar{\mathbf{g}}_k^H \left(\frac{1}{N} \sum_{i=1}^K \mathbf{R}_i + \frac{1}{N} \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right) \bar{\mathbf{g}}_k} \right], \quad (19)$$

where $\bar{\mathbf{H}}_k$ is obtained by removing the k -th column from the LoS-channels matrix $\bar{\mathbf{H}}$.

Using the expression (19), we can now derive $\bar{\mathbf{g}}_k$ by solving the SE optimization problem (P1). Note that the problem (P1) is the sum of decoupled positive and increasing functions, therefore, a sufficient condition to solve it is to find $\forall k$, $k = 1, \dots, K$, $\bar{\mathbf{g}}_k$ that satisfies:

$$\bar{\mathbf{g}}_k = \underset{\bar{\mathbf{g}}_k}{\text{argmax}} \frac{\bar{\mathbf{g}}_k^H \left(\frac{\bar{\mathbf{h}}_k \bar{\mathbf{h}}_k^H}{N} \right) \bar{\mathbf{g}}_k}{\bar{\mathbf{g}}_k^H \left(\frac{1}{N} \sum_{i=1}^K \mathbf{R}_i + \frac{1}{N} \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right) \bar{\mathbf{g}}_k}. \quad (\text{P1}')$$

It can be seen that (P1') is equivalent to Rayleigh quotient and thus admits the solution:

$$\bar{\mathbf{g}}_k = \left(\sum_{i=1}^K \mathbf{R}_i + \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \bar{\mathbf{h}}_k \quad (20)$$

Consequently, under assumption 1,

$$\overline{\text{SE}}_k^{\text{stat},S} = \log \left[1 + \frac{1}{N} \bar{\mathbf{h}}_k^H \left(\frac{1}{N} \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right)^{-1} \bar{\mathbf{h}}_k \right] + \mathcal{O} \left(\frac{1}{N} \right). \quad (21)$$

Furthermore, under favorable propagation conditions:

$$\overline{\text{SE}}_k^{\text{stat},S} = \log \left(1 + \frac{\rho_d}{N} \|\bar{\mathbf{h}}_k\|^2 \right) + \mathcal{O} \left(\frac{1}{N} \right). \quad (22)$$

As can be seen from (21)-(22), the UL performances generated by $\bar{\mathbf{g}}_k$ are mainly determined by the level of the specular component. That is, the proposed statistical processing is essentially beneficial in LoS-prevailing environments, thereby requiring a certain level of κ .

4) *Comparative Analysis (Case Study):* As concluded above and will be illustrated in simulations, the statistical processing is convenient when the specular component is dominant over the scattered signals. For this reason, we aspire here to find a condition on the Rician factor κ_k under which the proposed statistical processing outperforms the conventional processing. We examine a simple network setting where:

$\frac{1}{N} \bar{\mathbf{h}}_j^H \bar{\mathbf{h}}_j \xrightarrow[N \rightarrow \infty]{a.s.} \frac{\beta_i \kappa_i}{1 + \kappa_i} \delta_{ij}$, where δ_{ij} is the Kronecker delta. Specifically, we determine $\kappa_k, \forall k$, s.t.:

$$\log \left(1 + \frac{\rho_d \beta_k \kappa_k}{1 + \kappa_k} \right) \geq \left(1 - \frac{\tau}{T} \right) \log \left[1 + \rho_d \left(\frac{1}{N} \text{tr} \tilde{\mathbf{R}}_k + \frac{\beta_k \kappa_k}{1 + \kappa_k} \right) \right] \quad (\text{P3})$$

Result: Taking into account that $\tau \in [K, T)$, $\frac{1}{N} \bar{\mathbf{h}}_j^H \bar{\mathbf{h}}_j \xrightarrow[N \rightarrow \infty]{a.s.} \frac{\beta_i \kappa_i}{1 + \kappa_i} \delta_{ij}$ and under Assumption 1, it can be shown that the statistical processing outperforms the conventional channel-estimate based processing, if κ_k verifies the sufficient condition⁴:

$$\kappa_k \geq \frac{\text{tr} \Theta_k}{N} \frac{T - K}{K}. \quad (23)$$

Proof: A proof is given in Appendix C. ■

Furthermore, if the correlation matrix Θ_k follows the widely used one-ring model [32] (introduced in the next section (35)) or the exponential correlation model [33], then $\frac{1}{N} \text{tr} \Theta_k = 1, \forall k$, therefore, (23) writes : $\kappa_k \geq \frac{T-K}{K}$. These inequalities provide a lower bound on a sufficient Rician factor above which the statistical combining is more profitable than the conventional receiver. This bound is a function of the systems parameters, including the coherence length and the number of users. Moreover, as can be seen, for a fixed T , $\frac{T-K}{K}$ is a decreasing function of K . As a result, the higher is the number of users, the smaller is the required κ_k to enable the use of the proposed statistical receiver with better UL performance and as such, avoid training along with channel estimation and its associated errors.

III. PERFORMANCE ANALYSIS IN A MULTI-CELL SCENARIO

In this section, we extend our analysis of both conventional and statistical combining schemes to a multi-cell scenario. The objective is to examine the impact of inter-cell interference and pilot contamination on the overall performance of such systems. Therefore, we consider a multi-cell network with L cells having each K single-antenna users communicating with an N -antennas BS. In this line, we follow the same notations as in the single-cell scenario, except that we add a triple sub-script indication to differentiate the receiving BS from the cell where the UE is located. For example, $\mathbf{h}_{j\ell k}$ represents the channel linking the k -th UE in cell ℓ to BS $_j$. Plus, $\mathbf{R}_{j\ell k}$ refers to the correlation matrix of channel $\mathbf{h}_{j\ell k}$, etc. Accordingly, the received signal at BS $_j$ is given by :

$$\mathbf{y}_j = \sqrt{p} \sum_{\ell=1}^L \sum_{i=1}^K \mathbf{h}_{j\ell i} \mathbf{x}_{\ell i} + \mathbf{n}_j, \quad (24)$$

where \mathbf{n}_j represents a zero-mean additive Gaussian noise with variance $\sigma^2 \mathbf{I}_N$. Plus, we consider correlated Rician fading for intra-cell or local channels and correlated Rayleigh fading for channels from other cells. This is a reasonable setting for inter-cell channels, owing to the longer distances between UEs

⁴As shown in proof, for mathematical convenience, we consider a higher bound than it is necessary.

and the BSs in other cells, that would likely include scatterers and thus significantly reduce the possibility of a Line-of-Sight transmission. Specifically, the channel linking BS $_j$ to UE k located in cell ℓ is modeled as:

$$\mathbf{h}_{j\ell k} = \sqrt{\beta_{j\ell k}} \left(\sqrt{\frac{1}{1 + \kappa_{j\ell k}}} \Theta_{j\ell k}^{\frac{1}{2}} \mathbf{z}_{j\ell k} + \delta_{j\ell} \sqrt{\frac{\kappa_{j\ell k}}{1 + \kappa_{j\ell k}}} \bar{\mathbf{z}}_{j\ell k} \right), \quad (25)$$

where $\delta_{j\ell}$ is the Kronecker delta, and $\beta_{j\ell k}$ accounts for the large-scale fading. Finally, to simplify the presentation of the results, let $\mathbf{R}_{j\ell k} = \frac{\beta_{j\ell k}}{1 + \kappa_{j\ell k} \delta_{j\ell}} \Theta_{j\ell k}$, and the aggregate matrix of the LoS components in cell j denoted $\bar{\mathbf{H}}_j = [\bar{\mathbf{h}}_{j1} \bar{\mathbf{h}}_{j2} \dots \bar{\mathbf{h}}_{jK}]$, with $\bar{\mathbf{h}}_{jk} = \sqrt{\frac{\beta_{j\ell k} \kappa_{j\ell k}}{1 + \kappa_{j\ell k}}} \bar{\mathbf{z}}_{j\ell k}$.

A. Conventional combining in Multi-Cell Systems

Similarly to the single-cell scenario, to design the conventional receiver, we consider a pre-training phase of τ symbols in each cell. On the other hand, to account for pilot contamination, we assume that the same set of pilot sequences is reused in every cell. More specifically, the same pilot is assigned to every k -th UE in each cell, and as such $\forall (j, k)$, the estimates of the channels $\mathbf{h}_{j1k}, \mathbf{h}_{j2k}, \dots, \mathbf{h}_{jLk}$, will be correlated. Accordingly, using the MMSE estimation, the estimate of \mathbf{h}_{jjk} is given by [29]:

$$\hat{\mathbf{h}}_{jjk} = \mathbf{R}_{jjk} \Phi_{jk} \left(\sum_{\ell'=1}^L \mathbf{h}_{j\ell'k} + \frac{1}{\sqrt{\tau \rho_{tr}}} \mathbf{n}_{jk}^{tr} \right) + \delta_{j\ell} \bar{\mathbf{h}}_{jk}, \quad (26)$$

where $\Phi_{jk} = \left(\sum_{\ell'=1}^L \mathbf{R}_{j\ell'k} + \frac{1}{\tau \rho_{tr}} \mathbf{I}_N \right)^{-1}$. Therefore, $\hat{\mathbf{h}}_{jjk} \sim \mathcal{CN}(\bar{\mathbf{h}}_{jk}, \tilde{\mathbf{R}}_{jjk})$, with $\tilde{\mathbf{R}}_{jjk} = \mathbf{R}_{jjk} \Phi_{jk} \mathbf{R}_{jjk}$. Plus, the estimation error $\boldsymbol{\xi}_{jjk} = \mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk}$, follows the distribution $\boldsymbol{\xi}_{jjk} \sim \mathcal{CN}(0, \mathbf{R}_{jjk} - \tilde{\mathbf{R}}_{jjk})$. Let $\mathbf{g}_{jk} \in \mathbb{C}^{N \times 1}$ denote the conventional combining vector that BS $_j$ uses to process the signal sent by its UE k . This vector is given by [4]:

$$\mathbf{g}_{jk} = \left(\sum_{i=1}^K \hat{\mathbf{h}}_{jji} \hat{\mathbf{h}}_{jji}^H + \mathbf{A}_j + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{jjk}. \quad (27)$$

where $\mathbf{A}_j \in \mathbb{C}^{N \times N}$ is an arbitrary hermitian positive semi-definite design parameter. For instance, it could contain the covariances of estimation errors and inter-cell interference as in [4]. Therefore, we shall put: $\mathbf{A}_j = \sum_{i=1}^K (\mathbf{R}_{jji} - \tilde{\mathbf{R}}_{jji}) + \sum_{\ell \neq j}^L \sum_{i=1}^K \mathbf{R}_{j\ell i}$. Accordingly, the achievable SE corresponding to this transmission, $\text{SE}_{jk}^{\text{conv}, M}$, is defined as in (28) (in the top of the next page).

Theorem 4 (Conventional combining in multi-cell systems). *Under Assumption 1, we have : $\text{SE}_{jk}^{\text{conv}, M} - \overline{\text{SE}}_{jk}^{\text{conv}, M} \xrightarrow[N \rightarrow \infty]{a.s.} 0$,*

such that $\overline{\text{SE}}_{jk}^{\text{conv}, M}$ is given in (29) (top of next page), with $\bar{\mathbf{Q}}_j = \left(\frac{1}{N} \bar{\mathbf{H}}_j^H \bar{\mathbf{H}}_j + \text{diag} \left\{ \frac{1}{N} \text{tr}(\tilde{\mathbf{R}}_{jji}) \right\}_{i=1}^K + \frac{1}{\rho_d} \mathbf{I}_K \right)^{-1}$, and $\bar{\mathbf{q}}_{j,k}$ is the k -th column of matrix $\bar{\mathbf{Q}}_j$.

Proof: The proof is given in Appendix D. ■

As can be seen from the SE approximation $\overline{\text{SE}}_{jk}^{\text{conv}, M}$ (29), the achievable UL SE in the multi-cell setting has an analogous

$$\overline{\text{SE}}_{jk}^{\text{conv},M} = \left(1 - \frac{\tau}{T}\right) \mathbb{E} \left[\log \left(1 + \frac{|\mathbf{g}_{jk}^H \hat{\mathbf{h}}_{jjk}|^2}{\mathbb{E} \left[\mathbf{g}_{jk}^H \left(\sum_{(\ell,i) \neq (j,k)} \mathbf{h}_{j\ell i} \mathbf{h}_{j\ell i}^H + \boldsymbol{\xi}_{jk} \boldsymbol{\xi}_{jk}^H + \frac{1}{\rho_d} \mathbf{I}_N \right) \mathbf{g}_{jk} \mid \hat{\mathbf{H}}_j \right]} \right) \right]. \quad (28)$$

$$\overline{\text{SE}}_{jk}^{\text{conv},M} = \left(1 - \frac{\tau}{T}\right) \log \left[1 + \frac{\left|1 - \frac{1}{\rho_d} [\overline{\mathbf{Q}}_j]_{kk}\right|^2}{\sum_{\substack{i=1 \\ i \neq k}}^K \left| \frac{1}{\rho_d} [\overline{\mathbf{Q}}_j]_{ki} \right|^2 + \frac{1}{\rho_d} \left([\overline{\mathbf{Q}}_j]_{kk} - \frac{1}{\rho_d} [\overline{\mathbf{Q}}_j]_{kk}^2 \right) + \sum_{\substack{\ell=1 \\ \ell \neq j}}^L \sum_{i=1}^K \left| \frac{1}{N} [\overline{\mathbf{Q}}_j]_{ki} \text{tr}(\mathbf{R}_{j\ell i} \boldsymbol{\Phi}_{ji} \mathbf{R}_{jj i}) \right|^2} \right]. \quad (29)$$

expression to the single-cell scenario (10), apart from the inter-cell interference represented by the last term of the denominator of $\overline{\text{SE}}_{jk}^{\text{conv},M}$ (29). As a result, most conclusions provided in Theorem 1 (*conventional combining in multi-cell systems*) hold true in the multi-cell setup, including the cancellation of the estimation errors as N grows large without bound. In addition, intra-cell interference is also generated by the inner products between the LoS components, and as such, dissipates under favorable propagation conditions.

We now move on to investigating the effect of inter-cell interference on this combining approach. In this line, by expanding the inter-cell interference approximation, $\overline{\text{SE}}_{jk}^{\text{conv},M}$ writes as in (30) (top of next page). Expression (30) separates the pilot contamination induced interference from the remaining inter-cell interference, which we refer to as “uncorrelated” interference. Clearly, inter-cell interference limits the overall performances even at the infinite-antenna limit. Nevertheless, note that its impact can be alleviated through the mitigation of the uncorrelated interference by observing that this latter is eliminated when $\overline{\mathbf{Q}}_j$ becomes diagonal. In fact, this is achieved in favorable propagation conditions, wherein the performances will attain:

Corollary 2 (Favorable propagation in multi-cell). *if $\frac{1}{N} \overline{\mathbf{h}}_{ji}^H \overline{\mathbf{h}}_{jk} \xrightarrow[N \rightarrow \infty]{a.s.} 0$, for $i \neq k$, we have:*

$$\overline{\text{SE}}_{jk}^{\text{conv},M} = \left(1 - \frac{\tau}{T}\right) \log \left(1 + \frac{\left(\frac{\rho_d}{N} (\text{tr} \tilde{\mathbf{R}}_{jjk} + \|\overline{\mathbf{h}}_{jk}\|^2) \right)^2}{\frac{\rho_d}{N} (\text{tr} \tilde{\mathbf{R}}_{jjk} + \|\overline{\mathbf{h}}_{jk}\|^2) + \underbrace{\sum_{\ell \neq j}^L \left(\frac{\rho_d}{N} \text{tr}(\mathbf{R}_{j\ell k} \boldsymbol{\Phi}_{jk} \mathbf{R}_{jj k}) \right)^2}_{\text{induced by pilot contamination}}} \right). \quad (31)$$

Another important outcome from Theorem 4 and corollary 2 lies in the interplay between the interference emanating from pilot-contamination and the LoS signals. In fact, consider the quantity $\frac{1}{N} \text{tr}(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell k} \boldsymbol{\Phi}_{jk} \mathbf{R}_{jj k})$ in (30) and (31) which represents this type of correlated interference. As shown in the following proof, this term is a decreasing function of the Rician factor κ_{jk} . Actually, in the limiting case $\kappa_{jk} \rightarrow \infty$, we have $\frac{1}{N} \text{tr}(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell k} \boldsymbol{\Phi}_{jk} \mathbf{R}_{jj k}) \rightarrow 0$. Consequently, we can state that in such multi-cell systems, another advantage of having stronger LoS components is to reduce the adverse

effects of pilot contamination which is known to be a limiting performance factor in massive MIMO systems [1], [4].

Proof: The proof is given in Appendix E ■

B. Statistical Combining in Multi-Cell Systems

In this section, we propose to investigate whether the statistical receiver defined in the single-cell scenario can still be beneficial in a multi-cell system with LoS-prevailing transmissions. In other words, we are interested in investigating the resilience of this combining scheme when it is subject to inter-cell interference. In this line, let $\overline{\mathbf{g}}_{jk}$ indicate the statistical combining vector associated with the communication between UE k and its BS j , defined as:

$$\overline{\mathbf{g}}_{jk} = \left(\sum_{i=1}^K \mathbf{R}_{jji} + \overline{\mathbf{H}}_{j,/k} \overline{\mathbf{H}}_{j,/k}^H + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \overline{\mathbf{h}}_{jk}, \quad (32)$$

where $\overline{\mathbf{H}}_{j,/k}$ is matrix $\overline{\mathbf{H}}_j$ without the k -th column. Furthermore, since utilizing this receiver allows to circumvent training and estimation, the corresponding UL SE attains:

$$\text{SE}_{jk}^{\text{stat},M} = \mathbb{E} \left[\log \left(1 + \frac{|\overline{\mathbf{g}}_{jk}^H \overline{\mathbf{h}}_{jk}|^2}{\mathbb{E} \left[\overline{\mathbf{g}}_{jk}^H \left(\sum_{\ell=1}^K \sum_{i=1}^K \mathbf{h}_{j\ell i} \mathbf{h}_{j\ell i}^H - \overline{\mathbf{h}}_{jk} \overline{\mathbf{h}}_{jk}^H + \frac{N}{\rho_d} \mathbf{I}_N \right) \overline{\mathbf{g}}_{jk} \right]} \right) \right]. \quad (33)$$

Next, we provide an asymptotic approximation of the achievable SE generated by $\overline{\mathbf{g}}_{jk}$. This constitutes the last main result of this work.

Theorem 5 (Statistical combining in multi-cell systems). *Under Assumption 1, $\text{SE}_{jk}^{\text{stat},M} - \overline{\text{SE}}_{jk}^{\text{stat},M} \xrightarrow[N \rightarrow \infty]{a.s.} 0$, with*

$$\overline{\text{SE}}_{jk}^{\text{stat},M} = \log \left[1 + \overline{\mathbf{h}}_{jk}^H \left(\overline{\mathbf{H}}_{j,/k} \overline{\mathbf{H}}_{j,/k}^H + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \overline{\mathbf{h}}_{jk} \right], \quad (34)$$

Proof: Under assumption 1, since $\overline{\mathbf{g}}_{jk}$ is deterministic, a direct application of the convergence of quadratic forms lemma [27] and the continuous mapping Theorem [31] yields $\overline{\text{SE}}_{jk}^{\text{stat},M}$. ■

First of all, we emphasize once again that the receiver, $\overline{\mathbf{g}}_{jk}$, is purposely designed for environments with strong LoS

$$\overline{\text{SE}}_{jk}^{\text{conv},M} = \left(1 - \frac{\tau}{T}\right) \log \left(1 + \frac{\left|\frac{\rho_d}{[\mathbf{Q}_j]_{kk}} - 1\right|^2}{\frac{\rho_d}{[\mathbf{Q}_j]_{kk}} - 1 + \underbrace{\sum_{\ell \neq j} \left(\frac{\rho_d}{N} \text{tr}(\mathbf{R}_{j\ell k} \Phi_{jk} \mathbf{R}_{j\ell k})\right)^2}_{\text{induced by pilot contamination}} + \underbrace{\sum_{i \neq k} \frac{\rho_d}{N} \frac{[\mathbf{Q}_j]_{ki}}{[\mathbf{Q}_j]_{kk}} \text{tr}(\mathbf{R}_{j\ell i} \Phi_{ji} \mathbf{R}_{j\ell i})}_{\text{uncorrelated interference}}}\right)}. \quad (30)$$

components. Second, in such environments, comparing the above multi-cell SE $\overline{\text{SE}}_{jk}^{\text{stat},M}$ (34) with the single-cell one $\overline{\text{SE}}_k^{\text{stat},S}$ (21), reveals that employing the statistical combining scheme entails a similar asymptotic performance gain for both network settings, *i.e.*, $\overline{\text{SE}}_{jk}^{\text{stat},M} - \overline{\text{SE}}_k^{\text{stat},S} \xrightarrow[N \rightarrow \infty]{a.s.} 0$. Therefore, we can conclude that, for LoS-prevailing communications, inter-cell interference can be mitigated through the use of $\overline{\mathbf{g}}_{jk}$ and as such, does not constitute a limitation of the achievable capacity when this processing approach is employed. This outcome is explained by the fact that the underlying premise behind the statistical receiver is to bypass training and thereby, prevent pilot-contamination and its ensuing undesirable effects. It is also important to note that this desirable feature comes in contrast to conventional combining, whose UL SE remains limited by pilot-contamination-induced interference, as previously shown in expressions (30)-(31), *i.e.*, $\overline{\text{SE}}_{jk}^{\text{conv},M} \leq \overline{\text{SE}}_k^{\text{conv},S}$.

IV. NUMERICAL RESULTS

In this section, we carry out MonteCarlo simulations over 1000 channel realizations to validate, for finite system dimensions, the asymptotic results for both single-cell and mutli-cell settings, given in Sections II and III.

A. Single-cell scenario

For this scenario, we consider a single-cell massive MIMO having one BS with $N = 150$ antennas, $K = 20$ users, and a coherence length $T_c = 500$ symbols. The inner cell-radius is 150m and the users are uniformly distributed around the BS at an arrival angle θ_k . Furthermore, the pathloss is given $\beta_k = \frac{1}{x_k^\alpha}$, where x_k is the distance between UE k and the BS and $\alpha = 2.5$. The specular component $\overline{\mathbf{z}}_k$ follows the model $[\overline{\mathbf{z}}_k]_n = e^{-j(n-1)\pi \sin(\theta_k)}$. Moreover, to ensure distinct Rician factors among the users, we assume throughout all the simulations that $\forall k, \kappa_k \sim \mathcal{U}[0, \kappa_{\max}]$. As a result, varying κ_{\max} yields specular signals with different levels of strength which ultimately enables to epitomize both NLoS and LoS prevailing environments. Finally, the elements of the correlation matrix Θ_k of channel \mathbf{h}_k are given by the one ring model [32]:

$$[\Theta_k]_{uv} = \frac{1}{\theta_{k,max} - \theta_{k,min}} \int_{\theta_{k,min}}^{\theta_{k,max}} e^{j \frac{2\pi}{\lambda} a_{uv} \cos(\theta)} d\theta, \quad (35)$$

where λ denotes the signal's wavelength and a_{uv} is the distance between receive antennas u and v . We also choose $\frac{a_{uv}}{\lambda} = 0.5|u - v|$, $\theta_{k,min} = -\pi$ and $\theta_{k,max} = \theta_k - \pi$.

We first illustrate the effects of the LoS presence and the length of the training sequence on the performances when using the conventional combining to, ultimately, validate the

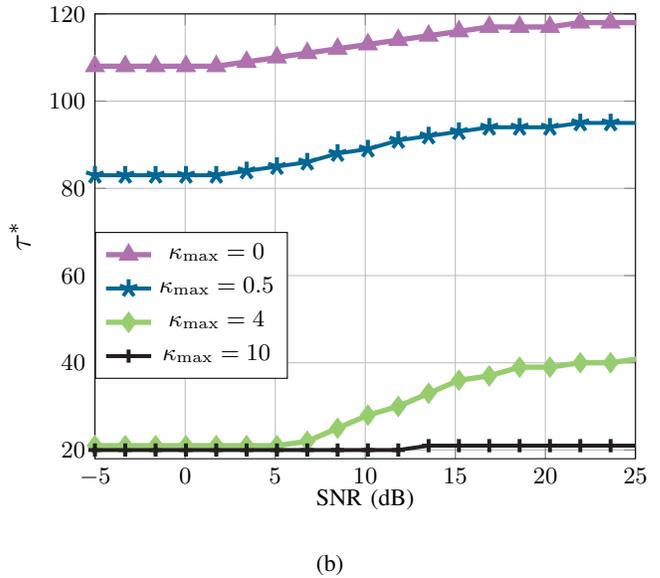
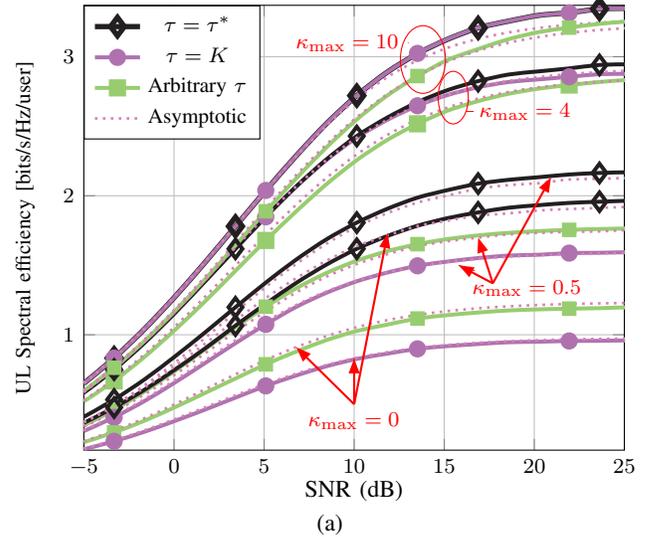


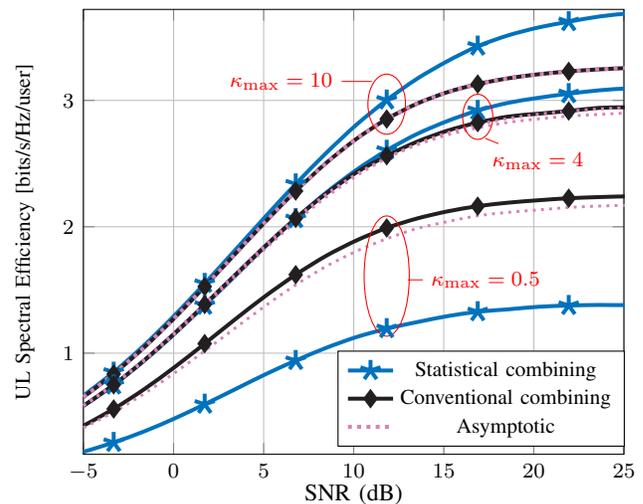
Fig. 1: Single-cell setting: (a) Impact of τ on the UL SE using conventional combining (4), for different levels of Rician factor, *s.t.* $\kappa_k \sim \mathcal{U}(0, \kappa_{\max})$. (b) Optimal number of training symbols τ^* (14) for different levels of κ_k .

conclusions of Theorems 1 and 2. To this end, we plot in Fig.1a the achievable SE $\overline{\text{SE}}_k^{\text{conv},S}$ (6) for different levels of the Rician factors, including, $\kappa_{\max} = 0$ (corresponding to Rayleigh fading), 0.5, 4 and 10. Moreover, to manifest the importance of the number of training symbols, we represent in

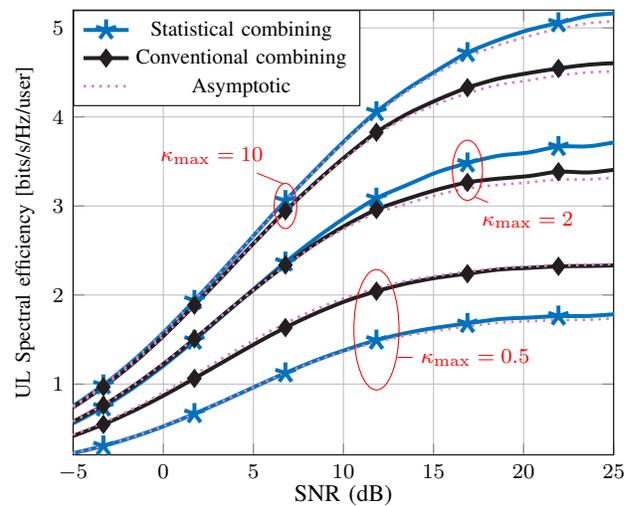
the same figure $SE_k^{\text{conv},S}$ with different values of τ , namely: the minimum K , the optimal τ^* (14), and another arbitrary value ($\neq \tau^*$ nor K). Solid and dotted lines represent empirical and asymptotic SEs, respectively. In addition, for each scenario, we plot in Fig.1b the obtained optimal values τ^* (13)-(14) with respect to the SNR for the various levels of κ_{\max} .

Overall, as expected, the LoS has a beneficial impact since increasing the Rician factor enables higher SEs, for any value of τ . As to this latter, it can be seen that the best performances are clearly obtained when the optimal number of symbols given in Theorem 2 (*i.e.* $\tau = \tau^*$) is considered (Fig.1a, diamond-marked curves). Furthermore, note that the gap between the settings $\tau = \tau^*$ and $\tau = K$ is particularly noteworthy at small values of Rician factors, (Fig.1a, $\kappa_{\max} = 0$ and 0.5, *i.e.* $0 \leq \kappa_k \leq 0.5, \forall k$). However, this difference in performance becomes less significant as κ_k takes higher values (Fig.1a, curves $\kappa_{\max} = 4$ and 10, *i.e.* $0 \leq \kappa_k \leq 10$). These simulation results confirm that as κ takes higher values, $\tau^* \rightarrow K$ as also displayed in Fig.1b. Indeed, Fig.1b clearly asserts that, for small Rician factors ($\kappa_{\max} = 0$ and 0.5), τ^* takes increasingly higher levels with the increase of the SNR. Conversely, $\tau^* \rightarrow K$ when $\kappa_{\max} = 10$. Evidently, since $\tau^* \rightarrow K$ for these scenarios, the SEs corresponding to $\tau = K$ and $\tau = \tau^*$ are almost identical (Fig.1a, overlapping diamond and circle-marked curves); whereas, interestingly, for $\tau > K$ (represented by the square-marked curve), we observe a decrease in the SE. Consequently, the plots in Fig.1a and Fig.1b validate the conclusions of Section II-C and the case study indicating the importance of assigning the optimal number of training symbols to attain the best performance. Additionally, as another important result, these simulations manifest that above a certain level of κ_{\max} , *i.e.* as the LoS get stronger, investing in longer training sequences to enhance the spectral efficiency is actually counterproductive. This interesting outcome inspired the proposed statistical combining that is a more opportune approach in such environments, as was previously demonstrated in Section II, and is illustrated in what follows.

Under the aforementioned network setting and for different values of κ , we compare in Fig.2 the UL SE $SE_k^{\text{stat},S}$ (7) achieved using the statistical receiver with the one attained by the conventional technique $SE_k^{\text{conv},S}$ (6) assuming optimal training, *i.e.* $\tau = \tau^*$. Plus, we represent ordinary and favorable propagation conditions in Fig.2a and Fig.2b, respectively. First, comparing Fig.2a with Fig.2b reveals that favorable propagation enable better performances for both combining techniques. This consequence, as explained in Theorem 1, is due to the cancellation of LoS induced intra-cell interference when the specular signals are mutually orthogonal. Second, as can be seen in both propagation conditions, conventional LMMSE is more beneficial than the statistical combiner at low ranges of the Rician factor, (Fig.2, $\kappa_{\max} = 0.5$). Nonetheless, as the LoS component becomes stronger, $SE_k^{\text{stat},S}$ progressively approaches $SE_k^{\text{conv},S}$, up to generating exceeding gains starting at $\kappa_{\max} = 4$ in ordinary conditions, and $\kappa_{\max} = 1.5$ in favorable propagation. This consequence can be justified by the expression $SE_k^{\text{stat},S}$ that clearly demonstrates that the statistical receiver's performance is mainly determined by the strength



(a) Ordinary propagation



(b) Favorable propagation

Fig. 2: Single-cell setting: UL SE using conventional combining (4) with optimal training τ^* and statistical combining (20) for different levels of κ , with (a) ordinary and (b) favorable propagation conditions.

of the LoS components. Therefore, these results confirm our single-cell analysis by substantiating the existence of a $\bar{\kappa}$ above which the statistical processing outperforms the conventional one. In the same line, this threshold value is fairly lower in favorable propagation compared to ordinary propagation environments. It is also important to note that they extend the outcome in Section II-C4 to a more realistic scenario that accounts for different per-user correlations and Rician factors. Finally, Fig.1 and Fig.2 validate, for finite system dimensions, the accuracy of the asymptotic approximations derived in Theorems 1, 2 and 3.

B. Multi-cell scenario

For the multi-cell scenario, we consider $L = 3$ adjacent cells having the same parameters as defined in the single-cell section. Besides, for each cell, we consider cell-edge users as shown in Fig.3. Deploying the users in such a configuration

generates high levels of inter-cell interference, and the close angles of arrival ensures considerable intra-LoS interference.

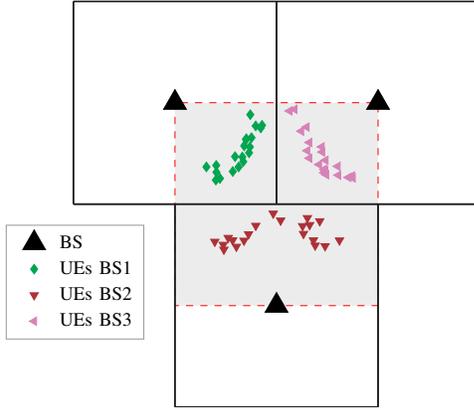


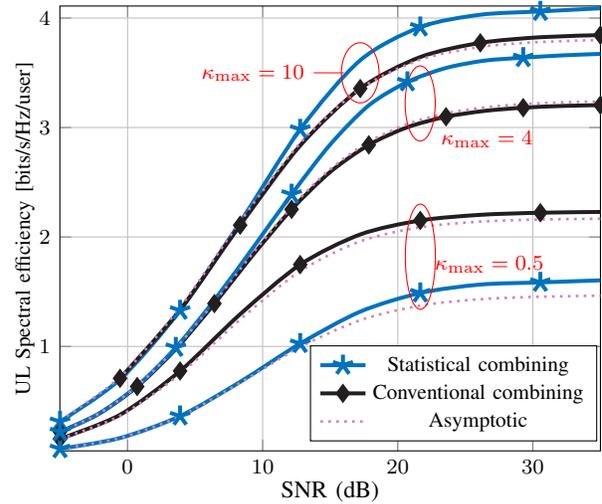
Fig. 3: Multi-cell network setup with $L=3$ cells and $K=20$ cell-edge users.

We illustrate in Fig.4 the achievable SEs for both statistical and conventional combining schemes considering pilot contamination and correlated Rician fading. Fig.4a and Fig.4b account for ordinary and favorable propagation conditions, respectively. In accordance with the discussion in section III, these figures confirm that the multi-cell UL spectral efficiencies follow the same pattern as those observed for a single-cell system. That is, firstly, higher Rician factors entail increasingly better performances. Secondly, favorable propagation conditions further enhance the SE, due to the cancellation of the uncorrelated inter-cell interference for the conventional combining, and the intra-LoS interference for both receivers, as analytically demonstrated in Theorems 4 and 5.

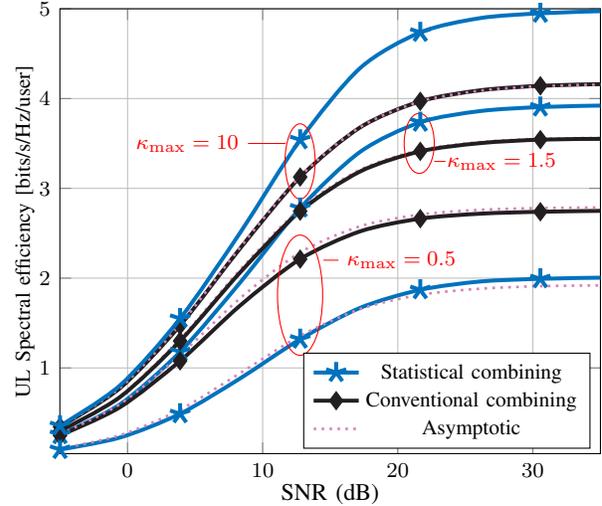
Next, to highlight the impact of inter-cell interference on the performance of the receivers, we consider cell 1 from the network setup in Fig.3 as a cell of interest, and propose to compare its achievable UL SE in the cases where it is deployed:

- (1) in the multi-cell setting of Fig.3,
- (2) in a single-cell setting having the same system and channel parameters as in (1).

Accordingly, we represent in Fig.5 the SEs corresponding to these cases for different levels of κ_{\max} . Dashed and solid lines (and arrows) correspond to cases (1) and (2), respectively. As indicated for $\kappa_{\max} = 2$, in the multi-cell setting, the statistical combining achieves a 33% SE gain over the conventional one; whereas in the single-cell case, the observed increase is by 8% only. As for smaller Rician factors (*i.e.*, $\kappa_{\max} = 1$), we can see that in the multi-cell plot, the statistical receiver outperforms the conventional one for a lower SNR (starting 17 dB) than it is the case for the single-cell scheme, wherein this is only achieved for SNRs above 25dB. Consequently, Fig.5 validates that, in the infinite antenna-limit, compared to conventional combining, employing the statistical receiver works even better in a multi-cell network. As demonstrated in SectionIII, this is explained by the fact that since it mitigates the inter-cell interference, this processing technique actually engenders a



(a) Ordinary propagation



(b) Favorable propagation

Fig. 4: Multi-Cell setting: UL SE using multi-cell conventional combining (27) with $\tau = K$ and statistical processing (32) with different levels of Rician factor, in (a) ordinary and (b) favorable propagation conditions.

similar multi-cell SE, $\overline{SE}_k^{\text{stat},M}$, to when used in a single-cell system, $\overline{SE}_k^{\text{stat},S}$. To summarize, Fig.5 asserts that for the same cell j : $\overline{SE}_k^{\text{stat},S} - \overline{SE}_{jk}^{\text{stat},M} \xrightarrow[N \rightarrow \infty]{a.s.} 0$, whereas, due to pilot contamination, $\overline{SE}_k^{\text{conv},S} \geq \overline{SE}_{jk}^{\text{conv},M}$. Nonetheless, note that the gap observed in Fig.5 between $\overline{SE}_k^{\text{stat},S}$ and $\overline{SE}_k^{\text{stat},S}$ is due to the finite system dimension considered in the simulations which will be reduced as N grows larger.

V. CONCLUSION

We studied in this work the UL performances of single and multi-cell massive MIMO systems underlying spatially correlated Rician channels, with the assumption of imperfect channel estimates. Considering the large-antenna limit, we derived closed-form approximations of the spectral efficiencies achieved by the LMMSE conventional receiver and proposed a

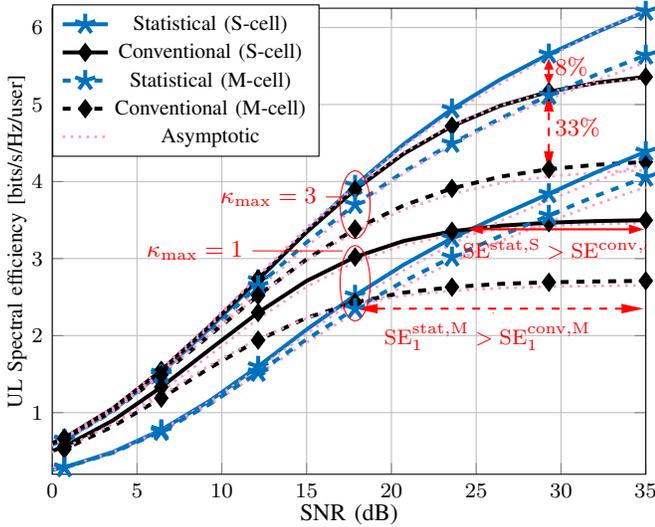


Fig. 5: UL SEs of conventional and statistical combining in both single-cell ($L = 1$) and multi-cell ($L = 3$) settings, for different levels of Rician factor, and $K = 20$ cell-edge users. Solid and dashed lines (and arrows) correspond to the single-cell and multi-cell cases, respectively. Dotted lines represent the asymptotic approximations given in Theorems 1 - 5.

novel statistical combining scheme. For the former, the approximations were exploited to determine an explicit expression of the optimal number of training symbols which was shown to be particularly important for small Rician factors. Conversely, the study reveals that, in LoS-prevailing environments, investing in longer training sequences to enhance the SE is ineffective. This result, led us to propose the statistical receiver that is more beneficial for systems with strong LoS components. The multi-cell analysis unveiled that conventional processing is limited by pilot contamination, even under favorable propagation; yet, it demonstrated that stronger LoS signals reduce this correlated interference. On another note, the asymptotic derivations indicated that statistical combining allows to mitigate inter-cell interference, and as such, outperforms the conventional receiver in a multi-cell system to an even higher extent. Finally, the approximations given in this work can be applied for realistic scenarios involving different correlation matrix models, Rician factors, CSI errors. In essence, they provide a general framework that can be harnessed to perform further analysis of similar networks.

APPENDIX A
PROOF OF THEOREM 1

We demonstrate in this section the results of Theorem 1. As the derivations rely most often on the same arguments, we mention the pertinent steps to derive the asymptotic approximation (10). First, define $\tilde{\mathbf{Q}} = \left(\frac{1}{N} \hat{\mathbf{H}}^H \mathbf{Z} \hat{\mathbf{H}} + \frac{1}{\rho_d} \mathbf{I}_K \right)^{-1}$, with $\mathbf{Z}^{-1} = \frac{\rho_d}{N} \sum_{i=1}^K (\mathbf{R}_i - \tilde{\mathbf{R}}_i) + \mathbf{I}_N$. Second, using the Woodbury matrix identity enables to express all the signals constituting $\text{SE}_k^{\text{conv},S}$ in terms of the elements of matrix $\tilde{\mathbf{Q}}$. For instance, the signal term $|\mathbf{g}_k^H \hat{\mathbf{h}}_k|^2$, can be written as:

$$\left| \frac{1}{N} \hat{\mathbf{h}}_k^H \left(\frac{\hat{\mathbf{H}} \hat{\mathbf{H}}^H}{N} + \frac{1}{N} \sum_{i=1}^K (\mathbf{R}_i - \tilde{\mathbf{R}}_i) + \frac{1}{\rho_d} \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_k \right|^2 =$$

$$\left| 1 - \frac{1}{\rho_d} [\tilde{\mathbf{Q}}]_{kk} \right|^2. \quad (36)$$

Accordingly, $\text{SE}_k^{\text{conv},S}$ in (6) can be rewritten as (37): (top of next page). In fact, putting the spectral efficiency $\text{SE}_k^{\text{conv},S}$ in this form (37) facilitates the derivation of the deterministic equivalent of this latter since we can simply use the LLN as follows:

- Under assumption 1, the LLN allows us to put $\frac{1}{N} [\hat{\mathbf{H}}^H \hat{\mathbf{H}}]_{ij} - \frac{1}{N} \mathbb{E} [\hat{\mathbf{h}}_i^H \hat{\mathbf{h}}_j] \xrightarrow[N \rightarrow \infty]{a.s.} 0$. Therefore, under Assumption 1, using the continuous mapping theorem [31], we have :

$$[\tilde{\mathbf{Q}}]_{ij} - [\mathbf{Q}]_{ij} \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (38)$$

where the matrix \mathbf{Q} is given in (8).

Thus, a direct application of (38) with the continuous mapping theorem [31] enables us to find asymptotic approximations of most of the terms in (37). For instance, for the signal term, we find: $\left| 1 - \frac{1}{\rho_d} [\tilde{\mathbf{Q}}]_{kk} \right|^2 - \left| 1 - \frac{1}{\rho_d} [\mathbf{Q}]_{kk} \right|^2 \xrightarrow[N \rightarrow \infty]{a.s.} 0$. Likewise, the same steps allow to derive approximations of the intra-cell interference term and processed noise. As to the estimation error term $\mathcal{E}_k = \frac{1}{N^2} \sum_{i=1}^K \tilde{\mathbf{q}}_k^H \hat{\mathbf{H}}^H \xi_i \xi_i^H \hat{\mathbf{H}} \tilde{\mathbf{q}}_k$, we mainly adopt the same reasoning except for the following step. Indeed, in order to find a deterministic equivalent for \mathcal{E}_k , we first exploit the orthogonality property of LMMSE channel estimation by observing that $\forall \{k, i\}$, $\hat{\mathbf{h}}_k$ and ξ_i are independent. After that, applying the convergence of quadratic forms lemma [27] yields:

$$\mathcal{E}_k - \left\{ \frac{1}{N^2} \sum_{i=1}^K \mathbf{q}_k^H \left(\hat{\mathbf{H}}^H \left(\frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \right) \bar{\mathbf{H}} + \text{diag} \left\{ \text{tr} \left(\tilde{\mathbf{R}}_{\ell} \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \right) \right\}_{\ell=1}^K \right) \mathbf{q}_k \right\} \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (39)$$

Finally, putting all the above terms together yields the asymptotic approximation of the spectral efficiency in Theorem 1 and as such, concludes the proof.

APPENDIX B
PROOF OF THEOREM 2

Denote $\partial_{\tau} F$ and $\partial_{\tau}^2 F$ as the first and second derivatives of any function $F(\tau)$ with respect to τ . To prove the results of Theorem 2, we use the following approach: First, we show that $\forall \tau$, $\partial_{\tau} \text{SE}^{\text{conv}}$ is monotonically decreasing and that $\exists \tau_0$, such that $\partial_{\tau} \text{SE}^{\text{conv}}|_{\tau=\tau_0} = 0$. Therefore, τ_0 is unique and $\tau_0 = \text{argmax}_{\tau} \text{SE}^{\text{conv}}$. Accordingly, if this step is verified, considering the constraint in (P2), finding τ^* is simply obtained as:

- If $\tau_0 \leq K$, then $\tau^* = K$, which is depicted by the first solution (13).
- On the other hand, if $\tau_0 \in [K, T[$, then $\tau^* = \tau_0$ which is represented by the solution (14).

Proof that $\forall \tau$, $\partial_{\tau} \text{SE}^{\text{conv}}$ is monotonically decreasing: To establish this, a sufficient condition is to have: $\forall \tau \in [K, T[$, $\partial_{\tau}^2 \text{SE}_k^{\text{conv},S} < 0$. In this line, using (11), we have :

$$\partial_{\tau} \overline{\text{SE}}^{\text{conv}} =$$

$$\text{SE}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log \left(1 + \frac{\left|1 - \frac{1}{\rho_d} [\tilde{\mathbf{Q}}]_{kk}\right|^2}{\sum_{i=1, i \neq k}^K \left|\frac{1}{\rho_d} [\tilde{\mathbf{Q}}]_{ki}\right|^2 + \frac{1}{N^2} \sum_{i=1}^K \tilde{\mathbf{q}}_k^H \hat{\mathbf{\Gamma}}^H \boldsymbol{\xi}_i \boldsymbol{\xi}_i^H \hat{\mathbf{\Gamma}} \tilde{\mathbf{q}}_k + \frac{1}{\rho_d} \left([\tilde{\mathbf{Q}}]_{kk} - \frac{1}{\rho_d} [\tilde{\mathbf{Q}}^2]_{kk}\right)}\right). \quad (37)$$

$$\frac{1}{K} \sum_{k=1}^K \left[\left(1 - \frac{\tau}{T}\right) \frac{\gamma'_k(\tau)}{1 + \gamma_k(\tau)} - \frac{1}{T} \log(1 + \gamma_k(\tau)) \right], \quad (40)$$

$$\begin{aligned} \partial_\tau^2 \overline{\text{SE}}^{\text{conv}} &= \frac{1}{K} \sum_{k=1}^K \left[-\frac{2}{T} \frac{\gamma'_k(\tau)}{1 + \gamma_k(\tau)} \right. \\ &\quad \left. + \left(1 - \frac{\tau}{T}\right) \frac{\gamma''_k(\tau)(1 + \gamma_k(\tau)) - (\gamma'_k(\tau))^2}{(1 + \gamma_k(\tau))^2} \right], \quad (41) \end{aligned}$$

with $\gamma'_k(\tau)$ and $\gamma''_k(\tau)$ given by :

$$\gamma'_k(\tau) = \rho_d \frac{\mathbf{q}_k^H \overline{\mathbf{D}}_2 \mathbf{q}_k}{([\mathbf{Q}]_{kk})^2}, \quad (42)$$

$$\gamma''_k(\tau) = -2\rho_d \frac{\mathbf{q}_k^H \overline{\mathbf{D}}_2 (\mathbf{Q}[\mathbf{Q}]_{kk} - \mathbf{q}_k \mathbf{q}_k^H) \overline{\mathbf{D}}_2 \mathbf{q}_k - [\mathbf{Q}]_{kk} \mathbf{q}_k^H \overline{\mathbf{D}}_3 \mathbf{q}_k}{([\mathbf{Q}]_{kk})^3}, \quad (43)$$

and $\overline{\mathbf{D}}_\alpha = \frac{(-1)^\alpha}{\rho_{tr} \tau^\alpha} \text{diag} \left\{ \frac{1}{N} \text{tr}(\mathbf{R}_\ell^\alpha \boldsymbol{\Phi}_\ell^\alpha) \right\}_{\ell=1}^K$, with α being an integer. Note that $\overline{\mathbf{D}}_\alpha$ is a positive semi-definite matrix for all even values of α , and negative semi-definite otherwise. With this in mind, from (42), we can see that $\gamma'_k(\tau) \geq 0$, $\forall \tau \in [K, T]$. Plus, since $\forall \tau$, $\gamma_k(\tau)$ is evidently positive, it can be seen from (41) that $\gamma''_k(\tau) \leq 0$, $\forall \tau \in [K, T]$ is a sufficient condition to obtain : $\partial_\tau^2 \overline{\text{SE}}^{\text{conv}} \leq 0$. With this in mind, using the fact that $\mathbf{a}^H \mathbf{a} \mathbf{I}_M - \mathbf{a} \mathbf{a}^H$ is a positive semi-definite matrix, we can easily show that $\gamma''_k(\tau) \leq 0$, $\forall \tau$. This concludes the proof that $\partial_\tau \text{SE}^{\text{conv}}$ is monotonically decreasing with respect to τ , and validates the results given in Theorem 2.

APPENDIX C PROOF OF COROLLARY II-C4

For simplicity, note that the index “ k ” will be dropped in the sequel. Accordingly, denoting α_i the i -th eigenvalue of $\boldsymbol{\Theta}$, our objective is to find κ such that :

$$\begin{aligned} &\left(1 - \frac{\tau}{T}\right) \log \left[1 + \frac{\beta \rho_d}{1 + \kappa} \left(\frac{1}{N} \sum_{i=1}^N \frac{\alpha_i^2}{\alpha_i + \frac{1 + \kappa}{\rho_{tr} \beta \tau}} + \kappa\right)\right] \\ &\leq \log \left(1 + \rho_d \beta \frac{\kappa}{1 + \kappa}\right). \quad (\text{P3}') \end{aligned}$$

Since “log” is an increasing function and $\alpha_i \geq 0$, $\forall i$, we consider the upper bound: $\frac{\alpha_i^2}{\alpha_i + \frac{1 + \kappa}{\rho_{tr} \beta \tau}} \leq \alpha_i$, for all positive values of κ , τ and ρ_{tr} . Therefore, (P3') is satisfied whenever κ verifies:

$$\begin{aligned} &\left(1 - \frac{\tau}{T}\right) \log \left[1 + \frac{\beta \rho_d}{1 + \kappa} \left(\frac{1}{N} \text{tr} \boldsymbol{\Theta} + \kappa\right)\right] \\ &\leq \log \left(1 + \rho_d \beta \frac{\kappa}{1 + \kappa}\right). \quad (44) \end{aligned}$$

Applying “exp” on both sides of (44) yields the lower bound $\kappa \geq f(\kappa)$, with:

$$f(\kappa) = -1 + \frac{\frac{1}{N} \text{tr} \boldsymbol{\Theta}}{-\frac{\kappa}{\kappa+1} + \frac{1}{\beta \rho_d} \left(-1 + \left(1 + \beta \rho_d \frac{\kappa}{1 + \kappa}\right)^{\frac{T}{T-\tau}}\right)}. \quad (45)$$

Next, we use the following result : *Let $x, \alpha \in \mathbb{R}$, and consider $y(x)$, a positive increasing function of all $x \geq 0$ with $y(0) = 0$. If $\alpha > 1$, the following inequality holds:*

$$(1 + y(x))^\alpha \geq (1 + \alpha y(x)), \quad (46)$$

Remark. This result can be proved by showing that the function $g(x) = (1 + y(x))^\alpha - (1 + \alpha y(x))$ verifies, $\forall x \geq 0$: $g(0) = 0$ and $g'(x) \geq 0$, thus yielding: $g(x) \geq 0$, $\forall x \geq 0$.

Accordingly, applying (46) on (45) allows us to obtain the following condition on κ :

$$\kappa \geq -1 + \frac{\text{tr} \boldsymbol{\Theta} (T - \tau) (1 + \kappa)}{N \tau \kappa}.$$

This latter admits the solution: $\kappa \geq \frac{\text{tr} \boldsymbol{\Theta} (T - \tau)}{N \tau}$. Finally, since the function $\frac{T - \tau}{\tau}$ is decreasing in τ , plus the fact that $\tau \in [K, T]$, we can simply consider the lower bound given in (23), therefore concluding the proof.

APPENDIX D PROOF OF THEOREM 4

The same steps and arguments given in appendix A can be used to find the asymptotic approximation $\overline{\text{SE}}_{jk}^{\text{conv},M}$ in Theorem 4, and are thus omitted due to space limitations. Nevertheless, the main difference lies in the inter-cell interference term, where it is imperative to take into account the correlation between the estimates and the interfering channels that share the same pilot, s.t., $\forall \ell \neq j$: $\frac{1}{N} \mathbb{E} \left[\hat{\mathbf{h}}_{jj}^H \hat{\mathbf{h}}_{j\ell} \right] - \frac{1}{N} \text{tr}(\mathbf{R}_{j\ell} \boldsymbol{\Phi}_{jk} \mathbf{R}_{jj}) \xrightarrow[N \rightarrow \infty]{a.s.} 0$.

APPENDIX E

PROOF THAT PILOT CONTAMINATION IS DECREASING WITH RESPECT TO κ_{jk}

First we need the following preliminary results.

Lemma 1. [34] *For any positive semi-definite (PSD) $N \times N$ matrices \mathbf{A} and \mathbf{B} , the matrices $\mathbf{A}\mathbf{B}\mathbf{A}$, $\mathbf{B}\mathbf{A}\mathbf{B}$ and $\mathbf{A} + \mathbf{B}$ are positive semi-definite, and $\text{tr}(\mathbf{A}\mathbf{B}) \geq 0$. Plus, if $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$, then $\mathbf{A}\mathbf{B}$ is also PSD. Finally, if \mathbf{A} is positive definite (PD), then \mathbf{A}^{-1} is also PD.*

Second, let $\kappa = \frac{1}{N} \text{tr}(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell} \boldsymbol{\Phi}_{jk} \mathbf{R}_{jj})$. A straightforward differentiation of $f(\kappa_{jk})$ yields:

$$f'(\kappa_{jk}) = \frac{-1}{1 + \kappa_{jk}} \frac{1}{N} \text{tr} \left(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell} \boldsymbol{\Phi}_{jk} (\mathbf{R}_{jjk} - \tilde{\mathbf{R}}_{jjk}) \right). \quad (47)$$

Accordingly, we prove in what follows that $\text{tr}\left(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell k} \Phi_{jk}(\mathbf{R}_{jjk} - \tilde{\mathbf{R}}_{jjk})\right) \geq 0$ which eventually leads to $f'(\kappa_{jk}) \leq 0$. In this line, we assume that the correlation matrix \mathbf{R}_{jjk} , is positive definite. That is, we add the ‘perturbation’ $\epsilon \mathbf{I}_N$, $\epsilon > 0$ as follows. Denote $\underline{\mathbf{R}}_{jjk} = \mathbf{R}_{jjk} + \epsilon \mathbf{I}_N$, and $\underline{\Phi}_{jk} = \left(\underline{\mathbf{R}}_{jjk} + \sum_{\ell \neq j} \mathbf{R}_{j\ell k} + \frac{1}{\rho_{tr}} \mathbf{I}_N\right)^{-1}$. Therefore, under this assumption, let $g(\kappa_{jk}, \epsilon) = \text{tr}\left(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell k} \underline{\Phi}_{jk}(\underline{\mathbf{R}}_{jjk} - \underline{\mathbf{R}}_{jjk} \underline{\Phi}_{jk} \underline{\mathbf{R}}_{jjk})\right)$. Note that $g(\kappa_{jk}, 0) = \text{tr}\left(\sum_{\ell \neq j}^L \mathbf{R}_{j\ell k} \Phi_{jk}(\mathbf{R}_{jjk} - \tilde{\mathbf{R}}_{jjk})\right)$. Plus, since $g(\kappa_{jk}, \epsilon)$ is continuous, if $g(\kappa_{jk}, \epsilon) > 0$, $\forall \epsilon > 0$, then $\lim_{\epsilon \rightarrow 0} g(\kappa_{jk}, \epsilon) \geq 0$.

$$\bullet \forall \epsilon > 0, \text{ let: } \mathbf{B}_{jk} = \sum_{\ell \neq j}^L \mathbf{R}_{j\ell k}, \mathbf{A}_{jk} = \left(\mathbf{B}_{jk} + \frac{1}{\tau \rho_{tr}} \mathbf{I}_N\right)^{-1}.$$

Therefore $\underline{\Phi}_{jk} = \left(\underline{\mathbf{R}}_{jjk} + \mathbf{A}_{jk}^{-1}\right)^{-1}$. Now, using the Woodbury matrix identity, we can put: $\underline{\mathbf{R}}_{jjk} \underline{\Phi}_{jk} \underline{\mathbf{R}}_{jjk} = \left(\underline{\mathbf{R}}_{jjk}^{-1} + \mathbf{A}_{jk}\right)^{-1}$. Therefore:

$$g(\kappa_{jk}, \epsilon) = \text{tr}\left(\mathbf{B}_{jk} \left(\underline{\mathbf{R}}_{jjk} + \mathbf{A}_{jk}^{-1}\right)^{-1} \left(\underline{\mathbf{R}}_{jjk}^{-1} + \mathbf{A}_{jk}\right)^{-1}\right) \\ = \text{tr}\left(\mathbf{B}_{jk} \mathbf{A}_{jk} \underbrace{\left(2\mathbf{A}_{jk} + \underline{\mathbf{R}}_{jjk}^{-1} + \mathbf{A}_{jk} \underline{\mathbf{R}}_{jjk} \mathbf{A}_{jk}\right)^{-1}}_{\text{PD due to lemma 1}}\right). \quad (48)$$

Plus, observing that $\mathbf{B}_{jk} \mathbf{A}_{jk} = \mathbf{A}_{jk} \mathbf{B}_{jk}$, the product $\mathbf{B}_{jk} \mathbf{A}_{jk}$ is also a PSD matrix, based on lemma 1. Consequently, we find that (48) amounts to the trace of the product of two PSD matrices which is always a positive quantity, thus concluding the proof.

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] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [3] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct 2014.
- [4] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, February 2013.
- [5] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, “Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback,” *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.
- [6] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, “Low-complexity linear precoding for multi-cell massive MIMO systems,” in *22nd European Signal Processing Conference (EUSIPCO)*, Sept 2014, pp. 2150–2154.
- [7] I. Boukhedimi, A. Kammoun, and M. S. Alouini, “Coordinated SLNR based precoding in large-scale heterogeneous networks,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 11, no. 3, pp. 534–548, April 2017.
- [8] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, “MIMO for millimeter-wave wireless communications: beamforming, spatial multiplexing, or both?” *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, December 2014.
- [9] T. E. Bogale and L. B. Le, “Massive MIMO and mmwave for 5G wireless hetnet: Potential benefits and challenges,” *IEEE Vehicular Technology Magazine*, vol. 11, no. 1, pp. 64–75, March 2016.
- [10] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, February 2014.
- [11] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, “Millimeter-wave massive MIMO: the next wireless revolution?” *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, September 2014.
- [12] Q. Zhang, S. Jin, K. K. Wong, H. Zhu, and M. Matthaiou, “Power scaling of uplink massive MIMO systems with arbitrary-rank channel means,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 966–981, Oct 2014.
- [13] W. Tan, S. Jin, J. Wang, and M. Matthaiou, “Achievable sum-rate of multiuser massive MIMO downlink in Rician fading channels,” in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 1453–1458.
- [14] H. Tataria, P. J. Smith, L. J. Greenstein, P. A. Dmochowski, and M. Shafi, “Performance and analysis of downlink multiuser MIMO systems with regularized zero-forcing precoding in Rician fading channels,” in *IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–7.
- [15] C. Kong, C. Zhong, M. Matthaiou, and Z. Zhang, “Performance of downlink massive MIMO in Rician fading channels with zf precoder,” in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 1776–1782.
- [16] H. Falconet, L. Sanguinetti, A. Kammoun, and M. Debbah, “Asymptotic analysis of downlink MISO systems over Rician fading channels,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 3926–3930.
- [17] S. Ghacham, M. Benjillali, and Z. Guennoun, “Low-complexity detection for massive MIMO systems over correlated Rician fading,” in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, June 2017, pp. 1677–1682.
- [18] J. Zhang, L. Dai, X. Zhang, E. Björnson, and Z. Wang, “Achievable rate of Rician large-scale MIMO channels with transceiver hardware impairments,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8800–8806, Oct 2016.
- [19] L. Sanguinetti, A. Kammoun, and M. Debbah, “Asymptotic analysis of multicell massive MIMO over Rician fading channels,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 3539–3543.
- [20] H. Tataria, P. J. Smith, L. J. Greenstein, P. A. Dmochowski, and M. Matthaiou, “Impact of line-of-sight and unequal spatial correlation on uplink MU-MIMO systems,” *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 634–637, Oct 2017.
- [21] J. Hoydis, M. Kobayashi, and M. Debbah, “Optimal channel training in uplink network MIMO systems,” *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2824–2833, June 2011.
- [22] J. K. N. Nyarko, R. Yao, and C. A. Mbom, “Area performance of multi-cell massive MIMO system under Rician fading,” in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Oct 2017, pp. 1–6.
- [23] I. Boukhedimi, A. Kammoun, and M. S. Alouini, “Line-of-Sight and pilot contamination effects on correlated multi-cell massive MIMO systems,” in *IEEE Global Communications Conference, to appear*, December 2018.
- [24] G. Dong, H. Zhang, and D. Yuan, “Downlink achievable rate of massive MIMO enabled swipt systems over Rician channels,” *IEEE Communications Letters*, vol. 22, no. 3, pp. 578–581, March 2018.
- [25] D. W. Yue, Y. Zhang, and Y. Jia, “Beamforming based on specular component for massive MIMO systems in Rician fading,” *IEEE Wireless Communications Letters*, vol. 4, no. 2, pp. 197–200, April 2015.
- [26] Ö. Özdogan, E. Björnson, and E. G. Larsson, “Massive MIMO with spatially correlated Rician fading channels,” *submitted to IEEE Transactions on Communications*, May 2018. [Online]. Available: <http://arxiv.org/abs/1805.07972>
- [27] D. Paul and J. W. Silverstein, “No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix,” *Journal of Multivariate Analysis*, vol. 100, no. 1, pp. 37–57, 2009.
- [28] I. Boukhedimi, A. Kammoun, and M. S. Alouini, “On the uplink of large-scale MIMO systems with correlated Rician fading channels,” in *2018 IEEE International Conference on Communications (ICC)*, to appear, May 2018.
- [29] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice Hall, 1997.

- [30] T. Marzetta, E. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*, ser. Fundamentals of Massive MIMO. Cambridge University Press, 2016.
- [31] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley, 1995.
- [32] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. Wiley-IEEE Press, 1994.
- [33] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, Sept 2001.
- [34] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York: Cambridge University press, 2013.



Ikram Boukhedimi was born in Blida, Algeria. She received the Electrical Engineering degree from the École Polytechnique d'Alger, Algeria, in 2013 and a M.S. in Mathematics, Control and Electrical Engineering from Université de Poitiers, France in 2014. Currently, she is working towards the Ph.D in Electrical Engineering at King Abdullah University of Science and Technology, Saudi Arabia. Her research interests include random matrix theory approaches applied to wireless communication systems.

Abla Kammoun was born in Sfax, Tunisia. She received the engineering degree in signal and systems from the Tunisia Polytechnic School, La Marsa, and the Master's degree and the Ph.D. degree in digital communications from Télécom Paris Tech [then École Nationale Supérieure des Télécommunications (ENST)]. From June 2010 to April 2012, she has been a Postdoctoral Researcher in the TSI Department, Télécom Paris Tech. Then she has been at Supélec at the Alcatel-Lucent Chair on Flexible Radio until December 2013. Currently, she is a research scientist at KAUST university. Her research interests include performance analysis, random matrix theory, and semi-blind channel estimation. She is the recipient of the SAM 2014 second Prize Best paper award.



Mohamed-Slim Alouini (S'94, M'98, SM'03, F09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009.

His current research interests include the modeling, design, and performance analysis of wireless communication systems.