

Towards a Realistic Assessment of Multiple Antenna HCNs: Residual Additive Transceiver Hardware Impairments and Channel Aging

Anastasios Papazafeiropoulos and Tharm Ratnarajah

Abstract—Given the critical dependence of broadcast channels by the accuracy of channel state information at the transmitter (CSIT), we develop a general downlink model with zero-forcing (ZF) precoding, applied in realistic heterogeneous cellular systems with multiple antenna base stations (BSs). Specifically, we take into consideration imperfect CSIT due to pilot contamination, channel aging due to users relative movement, and unavoidable residual additive transceiver hardware impairments (RATHIs). Assuming that the BSs are Poisson distributed, the main contributions focus on the derivations of the upper bound of the coverage probability and the achievable user rate for this general model. We show that both the coverage probability and the user rate are dependent on the imperfect CSIT and RATHIs. More concretely, we quantify the resultant performance loss of the network due to these effects. We depict that the uplink RATHIs have equal impact, but the downlink transmit BS distortion has a greater impact than the receive hardware impairment of the user. Thus, the transmit BS hardware should be of better quality than user's receive hardware. Furthermore, we characterise both the coverage probability and user rate in terms of the time variation of the channel. It is shown that both of them decrease with increasing user mobility, but after a specific value of the normalised Doppler shift, they increase again. Actually, the time variation, following the Jakes autocorrelation function, mirrors this effect on coverage probability and user rate. Finally, we consider space division multiple access (SDMA), single user beamforming (SU-BF), and baseline single-input single-output (SISO) transmission. A comparison among these schemes reveals that the coverage by means of SU-BF outperforms SDMA in terms of coverage.

Index Terms—Channel estimation, channel aging, additive hardware impairments, coverage probability, multiple antenna heterogeneous cellular networks.

I. INTRODUCTION

The design of the future Fifth Generation (5G) networks, demanding to cover the upcoming avalanche of wireless traffic volume due to the accompanied societal development, is quite challenging. Intercell interference is considered as a key limiting factor among the next generation of wireless systems [1], [2], which include both a vector Gaussian broadcast and interference channels by means of multi-user multiple-input multiple-output (MU-MIMO) and multi-cell scenarios,

respectively. Simplistic models such as the Wyner model [3], where the intercell interference is assumed to be a constant factor of the total interference, are highly inaccurate given the dramatic variation of the signal-to-interference-plus-noise ratio (SINR) value across a cell. Nevertheless, other works consider a fixed user or a small number of interfering BSs which might provide tractable results but very pessimistic with not much insight on users' performance [4].

Fortunately, tractable and accurate models have been developed for studying the downlink coverage, taking into account the full network interference [5]–[7]. Specifically, the introduction of the randomness regarding the locations of the BSs by means of a heterogeneous Poisson point process (PPP), which allows the usage of tools from stochastic geometry, has taken place. In this technique, known as heterogeneous networks (HetNets) design, small cells are embodied into a macrocell network with main benefits being a dense coverage and ubiquitous high throughput [8]. Hence, novel results regarding the coverage probability have been derived to quantify the quality of service in next generation networks without the need for Monte-Carlo simulations. HetNets design belongs to the major technologies currently on the table for 5G. The heterogeneous cellular networks (HCNs), which are the focal point of this study, can be considered as HetNets of a single tier.

Although appealing in their concept, HCNs as any other network, are hampered by the inevitable transceiver hardware impairments [9]–[25]. The impact of hardware impairments is a major challenge because the applied compensation algorithms, which include analog and digital signal processing, cannot remove the impairments completely, since the time-varying hardware characteristics cannot be fully parameterized and estimated, and because there is a randomness induced by different sources of noise [9], [22], [26], [27]. Especially, cheap hardware components, being attractive for industrial implementation, are particularly prone to the transceiver impairments. Such impairments are originated by amplifier non-linearities, I/Q-imbalance, and quantization errors [9], [10], [14], and can mainly be modelled as residual additive impairments [9]–[13], [16]–[25]. In the literature, three basic categories are met, namely, the residual additive impairments, the multiplicative impairments, and the amplified thermal noise [26], [27]. The residual additive impairments, modelled as independent additive distortion noises at the BS as well as at the user, describe the aggregate effect from many impairments. On the other hand, there are hardware impairments multiplied

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

A. Papazafeiropoulos and T. Ratnarajah are with the Institute for Digital Communications (IDCOM), University of Edinburgh, Edinburgh, EH9 3JL, U.K., (email: a.papazafeiropoulos, t.ratnarajah@ed.ac.uk).

This work was supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under grants EP/N014073/1 and EP/N015312/1.

with the channel vector, which might cause channel attenuation and phase shifts. They appear as a multiplicative distortion that cannot be incorporated by the channel vector. Herein, our analysis focuses on the impact of the residual additive hardware impairment, while the study of multiplicative impairments is left for future work. Regarding the introduction of amplified thermal noise, this is straightforward. Moreover, it is worthwhile to mention that the adoption of the current model for the additive impairments is based on its analytical tractability and the experimental verifications. Remarkably, some of the authors have achieved to introduce the rate-splitting approach as a robust method against the RATHIs, although these impairments are residual [24], [25]. The topic of dealing with other methods and strategies to mitigate the RATHIs is left for future work. Notably, since HetsNets are a candidate solution for 5G systems [28], [29], they need to be evaluated in the presence of the detrimental transceiver hardware imperfections, in order to make realistic conclusions regarding their final implementation. It is worthwhile to mention that the research area of HCNs has not considered this kind of inevitable degradations except [21], which in turn, brings out a gold rush for new research. Note that [21] assumes perfect channel state information at the transmitter (CSIT), neglecting that, in practice, CSIT is imperfect due to several reasons.

Several works have investigated the impact of generally imperfect CSIT in different scenarios of HetNets [30]–[34]. However, except [35], no other previous work has taken into account for pilot contamination due to the re-use of the pilot sequences or users’ mobility, which both are inevitable degrading causes on system’s performance and result in imperfect CSIT [36]–[38]. In particular, pilot contamination is an inherent weakness in next generation systems aiming to employ the concept of massive MIMO which include a large number of antennas (tens or hundreds) at each BS [36]. Notably, massive MIMO are based on time-division duplexing (TDD) mode for channel estimation. With respect to users’ mobility and its induced time variation of the channel, a lack of investigation appears in the literature [37], [38]. Its practical importance becomes higher, especially, in outdoor urban environments, where the mobility of terminals is increased. Despite that numerous studies, concerning 5G systems, have been presented (see [39] and references therein), the lack of channel aging works continues. The impact of terminal mobility is often modelled by means of a stationary ergodic Gauss-Markov block fading [37], [38], [40]–[42], [42]–[45]. For example, in [37], the authors provided deterministic equivalents (DEs) for the maximal-ratio-combining (MRC) receivers in the uplink and the maximal-ratio-transmission (MRT) precoders in the downlink¹. This analysis was extended in [38], [41] by deriving DEs for the more sophisticated minimum mean-square error (MMSE) receivers (for the uplink) and regularized zero-forcing precoders (for the downlink). It is worthwhile to mention that in the analysis of this paper, we consider both the pilot contamination and the relative user mobility.

¹The deterministic equivalents are deterministic tight approximations of functionals of random matrices of finite size.

A. Motivation-Central Idea

This work relies on the observation that the promising HCNs, meant to be applied in 5G systems, have been mostly evaluated under the assumptions of perfect hardware and static environment, while these are highly unrealistic. In addition, no imperfect CSIT has been assumed regarding HCNs MIMO systems. Only in [35], the authors have studied the uplink of massive MIMO systems with pilot contamination. Hence, the fundamental question behind this work is “how the transceiver impairments, the pilot contamination, and the channel aging affect the downlink performance of realistic HCNs, when imperfect CSIT is accounted?” Motivated by recent performance analysis results, we are going to establish the theoretical framework modelling the additive residual transceiver hardware impairments (RATHIs) and imperfect CSIT, and identify the realistic potentials of HCNs before their final implementation.

The main contributions are summarised as follows.

- Contrary to existing work [21] which has studied the effect of RATHIs in the case of perfect CSIT, we calculate the estimated channel due to RATHIs, pilot contamination, and channel aging, and then, we evaluate the effect of RATHIs and imperfect CSIT on the performance of downlink realistic HCNs in terms of the coverage probability and the user rate. For the sake of comparison, we also present the results corresponding to perfect hardware and no user mobility. As far as the authors are aware, this is the most general result in the literature that accounts for practical impairments and imperfect CSIT.
- Contrary to [35], where only the uplink performance was investigated for massive MIMO, we focus on the downlink. Moreover, we present more general results, since the proposed metrics can describe scenarios with not only finite number of BS antennas, but also with large number of antennas. On the other hand, works such as [35] describe only the large number of antennas regime.
- We investigate and elaborate on the impact of RATHIs and imperfect CSIT on the downlink coverage probability and the downlink achievable user rate. Actually, without being vague regarding the sources of imperfect CSIT in HetNets such as in [30]–[34], we consider the presence of pilot contamination, channel aging, and hardware impairments. Specifically, we show that the uplink hardware impairments have an equal effect on the estimation of the channel. However, the downlink hardware impairments behave differently. The BS transmit impairments degrade more the system performance than the user receive impairments. Moreover, the higher the time variation of the channel, the higher the degradation of the system. Accordingly, a proper system design should take into account these observations.
- We focus on the design of a realistic HCN and investigate, if it is beneficial to transmit with space division multiple access (SDMA) or to employ fewer antennas under practical conditions. In fact, single-stream transmission looks better than SDMA. In other words, the claim having more antennas is always beneficial is not necessarily

correct, as it heavily depends on how the transmit antennas are used and which transmission/reception scheme is employed from a system perspective.

The remainder of this paper is structured as follows. Section II presents the basic parameters of the system model of a HCN with randomly located BSs having multiple antennas and serving multiple users. In Section III, a description of the RATHIs is provided. In Section IV, the channel estimation phase, considering pilot contamination and channel aging, is modelled under the presence of RATHIs. Next, Section V exposes the downlink transmission under RATHIs and imperfect CSIT. Subsection VI-A includes the derivation and investigation of the coverage probability in a realistic HCN with multiple antenna BSs impaired by RATHIs, when multiple users are served. Remarkably, a relative movement of the users with comparison to the BS antennas is considered as well. In Subsection VI-B, the derivation of the achievable user rate is provided under the same realistic conditions. The numerical results are placed in Section VII, while Section VIII summarises the paper.

Notation: Vectors and matrices are denoted by boldface lower and upper case symbols. $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, and $\text{tr}(\cdot)$ express the transpose, conjugate, Hermitian transpose, and trace operators, respectively. The expectation and variance operators are denoted by $\mathbb{E}[\cdot]$ and $\text{Var}[\cdot]$. The $\text{diag}\{\cdot\}$ operator generates a diagonal matrix from a given vector, and the symbol \triangleq declares definition. The notations $\mathcal{C}^{M \times 1}$ and $\mathcal{C}^{M \times N}$ refer to complex M -dimensional vectors and $M \times N$ matrices, respectively. The indicator function $\mathbb{1}(e)$ is 1 when event e holds and 0 otherwise, and $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. Moreover, $B(x, y)$ is the Beta function defined in [46, Eq. (8.380.1)], and $\Gamma(x, y)$ denotes the Gamma distribution with shape and scale parameters x and y , respectively. Furthermore, $\bigcup_{x \in A}$ denotes the union with A being an index set. Also, $\mathcal{L}_I(s)$ expressed the Laplace transform of I . Finally, $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ represents a circularly symmetric complex Gaussian variable with zero-mean and covariance matrix Σ .

II. SYSTEM MODEL

In this paper, we consider a network cellular MU system with a BS per cell that has multiple antennas and serves multiple users. The locations of the BSs are drawn according to an independent PPP Φ_B with density λ_B . In other words, we refer to the formulation of MU-MIMO HCNs. Herein, we note that the locations of the users are also modelled by an independent PPP Φ_k with a sufficiently high density λ_k . In addition, the BS in the l th cell, deployed with a number of BS antennas M_l , communicates with its associated users whose number is K_l such that $K_l \leq M_l$. Hence, many degrees of freedoms are shared². Actually, we assume that each cell is large enough, i.e., the l th macrocell can accommodate K users

²Given this model, it is realistic that the downlink transmit power p_l , the number of BS antennas M_l , and the number of users served by each BS K_l differ across cells. However, for the sake of simplicity, we assume universal parameters in this setting, i.e., $p_l = p$, the number of BS antennas $M_l = M$, and the number of users served by each BS $K = K_l \leq M_l = M$.

connected with the nearest BS. In other words, K users, that are independently distributed, belong to the Voronoi cell of this BS, while a Voronoi tessellation is structured by the set of all these cells. Exploiting Slivnyak's theorem we are able to conduct the analysis by focusing on a typical user found at the origin [47]. Basically, we focus on the downlink scenario of communication between the BS and the associated users. During the uplink, channel estimation takes place, while we elaborate further on the downlink data transmission, where we derive the coverage probability and study its realistic behaviour.

Regarding HCNs, their evolution, started from the downlink single-input single-output (SISO) systems [5], has enabled the coexistence of multiple antenna strategies [48], [49] such as beamforming and SDMA. In fact, HCNs and multiple antenna strategies coexist and complement each other. Thus, they should not be studied in isolation, as happened in the premature literature in this area. For this reason, we focus on the impact of RATHIs on multiple antenna HCNs.

All point-to-point channels are characterised by independent and identically distributed (i.i.d.) Rayleigh block fading models with unit mean. Moreover, the same time-frequency resources are shared by the users across all cells. Note that we assume that the channel coherence time is T_c . At the same time, aiming to provide a realistic analysis, we assume imperfect CSIT due to pilot contamination, channel aging, and RATHIs. In other words, the BS is aware of the estimated channel, obtained during the training phase and having duration τ symbols, while the downlink transmission phase has a duration of $T_c - \tau$ symbols.

III. PRESENTATION OF RATHIS

In practice, both the users and the BSs are affected by certain additive impairments [9], [10]. Although mitigation schemes are implemented in both the transmitter and receiver, these are not perfect. Therefore, RATHIs still emerge by means of residual additive distortion noises [9], [10]. More concretely, at the transmitter side, an impairment emerges that causes a mismatch between the intended signal and what is actually transmitted during the transmit processing, while at the receiver side the received signal appears a distortion.

The study of the impact of RATHIs has originated from conventional wireless systems and has continued to 5G networks such as massive MIMO systems [9]–[13], [16], [17], [26], [50], [51]. Unfortunately, the majority of HCNs literature, except [21] relies on the assumption of perfect transceiver hardware, although hardware imperfections exist. Reasonably, it is conjectured that by following the same path will increase the gap between theory and practice.

Mathematically speaking, given the channel realisations, the conditional transmitter and receiver distortion noises for the i th link are modelled as Gaussian distributed, where their average power is proportional to the average signal power, as shown by measurement results [10].

Let us denote the transmit and receive nodes as nodes i and j , where $i = \text{UE}$ and $j = \text{BS}$ for the uplink, while $i = \text{BS}$ and $j = \text{UE}$ for the downlink. In addition, let T_i be the number of transmit antennas, i.e., $T_{\text{UE}} = 1$ for the uplink and $T_{\text{BS}} = M$

for the downlink. $\mathbf{Q}_i[n]$ is the transmit covariance matrix at time instance n of the corresponding node with diagonal elements $q_{i_1}[n], \dots, q_{T_i}[n]$, e.g., if the transmitter node is the UE, $\mathbf{Q}_i[n]$ degenerates to a scalar $Q_{\text{UE}}[n]$. Specifically, the RATHIs at the transmitter and the receiver are given by

$$\boldsymbol{\eta}_{t,n}^i \sim \mathcal{CN}(0, \boldsymbol{\Lambda}_n^i) \quad (1)$$

$$\boldsymbol{\eta}_{r,n}^j \sim \mathcal{CN}(0, \boldsymbol{\Upsilon}_n^j), \quad (2)$$

where $\boldsymbol{\Lambda}_n^i = \kappa_{t_i}^2 \text{diag}(q_1[n], \dots, q_{T_i}[n])$ and $\boldsymbol{\Upsilon}_n^j = \kappa_{r_j}^2 \|x_{k,n}\|^{-\alpha} \sum_{k=1}^{|j|} \mathbf{h}_k^H[n] \mathbf{Q}_k[n] \mathbf{h}_k[n]$. Note that if $j = \text{UE}$, then $|j| = 1$. The circularly-symmetric complex Gaussianity can be justified by the aggregate contribution of many impairments³. The proportionality parameters $\kappa_{t_i}^2$ and $\kappa_{r_j}^2$ describe the severity of the residual impairments at the transmitter and the receiver side. In applications, these parameters are met as the error vector magnitudes (EVM) at each transceiver side [52]. The procedure for obtaining the knowledge of the estimated channel is described in the following section.

IV. CHANNEL ESTIMATION

The transmit signal by each user to its BS is attenuated with distance r as $r^{-\alpha}$, where α is the path-loss exponent parameter. During the channel estimation, an effect, known as pilot contamination due to the re-use of the pilot sequences, might arise. It is worthwhile to mention that both the system model and the proposed expressions below can describe any number of BS antennas, e.g., both small and large number of antennas. Indeed, the results are quite general, and will be derived by means of a common analysis followed in the study of HetNets. In other words, contrary to works for massive MIMO that can describe only the large antenna regime, the proposed expressions can describe the whole spectrum in terms of the number of antennas (from small to large number). Notably, the pilot contamination concerns systems with any number of antennas. However, it is negligible for small number of antennas, while its presence is observable for large MIMO cellular systems [36]. Hence, since our model is able to describe any number of antennas (even a large number), pilot contamination is meaningful in the current work.

A. Pilot Contamination

Each BS estimates the CSI during an uplink training phase, where the sharing of the same band of frequencies leads to degradation of the performance of the system due to pilot contamination. If the subscript tr denotes the training stage, the noisy observation of the channel vector from the user k (typical user at the origin) at time instance n , transmitting each pilot symbol with average power $\rho_{\text{up}}^{\text{UE}}$ ($p_{\text{tr}} = \tau \rho_{\text{up}}^{\text{UE}}$) to

its associated BS in the presence of RATHIs, is given by

$$\begin{aligned} \tilde{\mathbf{y}}_{k,\text{tr}}[n] = & \underbrace{\mathbf{h}_k[n] \|x_{k,n}\|^{-\alpha/2}}_{\text{Desired signal}} + \underbrace{\mathbf{h}_k[n] \boldsymbol{\eta}_{t,n}^{\text{UE}} \|x_{k,n}\|^{-\alpha/2}}_{\text{UE transmit impairment}} \\ & + \underbrace{\boldsymbol{\eta}_{r,n}^{\text{BS}}}_{\text{BS receive impairment}} + \underbrace{\sum_{l \in \Phi_B/b_0} \mathbf{g}_{lk}[n] \|y_{lk,n}\|^{-\alpha/2}}_{\text{Interference}} + \underbrace{\frac{1}{\sqrt{p_{\text{tr}}}} \mathbf{N}_{\text{tr}}[n] \boldsymbol{\psi}_k^H}_{\text{noise}} \end{aligned} \quad (3)$$

where $\mathbf{h}_k[n] \in \mathcal{C}^{M \times 1}$ is the desired channel vector from the k th user in the current cell (located at the origin) to its associated BS located at $x_{k,n}$, while the interference term $\mathbf{g}_{lk}[n] \in \mathcal{C}^{M \times 1}$ is the channel vector corresponding to the link from the k th user of the l th cell located at $y_{lk,n} \in \mathcal{R}^2$. The vector $\boldsymbol{\psi}_k \in \mathcal{C}^{\tau \times 1}$ denotes the training sequence of the k th user with $\boldsymbol{\psi}_k \boldsymbol{\psi}_k^H = \mathbf{I}$ and $\mathbf{N}_{\text{tr}}[n] \in \mathcal{C}^{M \times \tau}$ is a spatially white additive Gaussian noise matrix with i.i.d entries at the current base station during the training stage, which are distributed as $\mathcal{CN}(0, 1)$. Similarly, the channel vectors $\mathbf{g}_{lk,n} \forall l, n$ are Gaussian distributed as $\mathcal{CN}(0, \mathbf{I}_M)$, and they are independent across cells and user distances. Based on (1) and (2) for $i = \text{UE}$ and $j = \text{BS}$, respectively, and after making the appropriate substitutions, the hardware impairments at the uplink are written such that their variances are given by

$$\begin{aligned} \Lambda_n^{\text{UE}} &= \kappa_{\text{tUE}}^2 \rho_{\text{up}}^{\text{UE}} \\ \Upsilon_n^{\text{BS}} &= \kappa_{\text{rBS}}^2 \rho_{\text{up}}^{\text{UE}} \|x_{k,n}\|^{-\alpha} \sum_{k=1}^K \text{diag}(|h_k^1[n]|^2, \dots, |h_k^M[n]|^2), \end{aligned}$$

where $\mathbf{Q}_{\text{UE}}[n]$ is now deterministic and $\rho_{\text{up}}^{\text{UE}} = \text{tr}(\mathbf{Q}_{\text{UE}}[n])$. After applying the minimum mean square error (MMSE) estimation to (3), in order to estimate $\mathbf{h}_k[n]$ conditioned on the distance $\|x_{k,n}\|$ from the k th user, the current BS obtains

$$\begin{aligned} \hat{\mathbf{h}}_k[n] \Big|_{\|x_{k,n}\|} &= \mathbb{E}[\mathbf{h}_k[n] \tilde{\mathbf{y}}_{k,\text{tr}}^H[n]] \mathbb{E}^{-1}[\tilde{\mathbf{y}}_{k,\text{tr}}[n] \tilde{\mathbf{y}}_{k,\text{tr}}^H[n]] \tilde{\mathbf{y}}_{k,\text{tr}}[n] \\ &= \frac{\sqrt{p_{\text{tr}}} \|x_{k,n}\|^{-\alpha/2}}{p_{\text{tr}} \|x_{k,n}\|^{-\alpha} (\kappa_{\text{r,BS}}^2 + \kappa_{\text{t,UE}}^2) + p_{\text{tr}} \text{Var}(I_n) + 1} \tilde{\mathbf{y}}_{k,\text{tr}}[n], \end{aligned} \quad (4)$$

where $I_n = \sum_{l \in \Phi_B/b_0} \mathbf{g}_{lk}[n] \|y_{lk,n}\|^{-\alpha/2}$ represents the interference contribution at time n and $\text{Var}(\cdot)$ is the variance operator conditioned on $\|x_{k,n}\|$. Specifically, conditioning on $\|x_{k,n}\|$, $\text{Var}(I_n)$ can be derived by using Campbell's theorem [53] as

$$\begin{aligned} \text{Var}(I_n) &= 2\mathbb{E}(\|\mathbf{g}_{lk}[n]\|^2) \pi \lambda_B \int_{\|x_{k,n}\|}^{\infty} r^{1-\alpha} dr \\ &= 2M\pi \lambda_B \frac{\|x_{k,n}\|^{2-\alpha}}{\alpha - 2}, \end{aligned} \quad (5)$$

where we have used that the square of $\mathbf{g}_{lk}[n]$ follows a $\Gamma(M, 1)$ distribution with mean value equal to M . Clearly, increase of the path-loss exponent and BSs' density, increases the variance of the interference.

As a result, the estimated channel $\hat{\mathbf{h}}_k$ and estimation error vectors $\mathbf{e}_k[n] = \mathbf{h}_k[n] - \hat{\mathbf{h}}_k[n]$ at time instance n are

³The additive distortions are time-dependent because they take new realisations for each new data signal.

uncorrelated and distributed as $\hat{\mathbf{h}}_k[n] \sim \mathcal{CN}(\mathbf{0}, \sigma_{\hat{\mathbf{h}}_k}^2 \mathbf{I}_M)$ and $\tilde{\mathbf{e}}_k[n] \sim \mathcal{CN}(\mathbf{0}, \sigma_{\tilde{\mathbf{e}}_k}^2 \mathbf{I}_M)$ with

$$\sigma_{\hat{\mathbf{h}}_k}^2 = \frac{(\alpha - 2)}{\left(\kappa_{r,BS}^2 + \kappa_{t,UE}^2 + 1\right)(\alpha - 2) + 2M\pi\lambda_B\|x_k\|^2 + \|x_k\|^\alpha} \quad (6)$$

and

$$\sigma_{\tilde{\mathbf{e}}_k}^2 = \frac{\left(\kappa_{r,BS}^2 + \kappa_{t,UE}^2\right)(\alpha - 2) + 2M\pi\lambda_B\|x_k\|^2}{\left(\kappa_{r,BS}^2 + \kappa_{t,UE}^2 + 1\right)(\alpha - 2) + 2M\pi\lambda_B\|x_k\|^2 + \|x_k\|^\alpha}. \quad (7)$$

Remark 1: If we set the distortion parameters in (4) or in any other expression including the RATHIs equal to zero, we result in the expressions corresponding to the ideal uplink model, which does not account for RATHIs.

Remark 2: Interestingly, in the interference-limited regime, the covariances of the estimated channel and the estimation error do not depend on the uplink transmit power p_{tr} . Moreover, in these expressions and under these conditions, the distance from the typical user is not raised to the path-loss parameter α . Thus, the covariances are not environment-dependent.

Remark 3: The uplink RATHIs, $\kappa_{t,UE}^2$ and $\kappa_{r,BS}^2$, have the same effect on the accuracy of the estimated channel. Therefore, equal caution should be given in the quality of the user transmit and BS receive hardware.

B. Channel Aging

Another additive reason with negative effect that necessitates the estimation of the channel is imposed by the relative movement of the users with comparison to the BS antennas. In such case, the description of the channel can be made by a time-varying model. In mathematical terms, the relationship of the current sample of the channel with its past samples, i.e., the time varying-model at time-slot n , is represented by an autoregressive-model of order 1 that includes the second order statistics of the channel in terms of its autocorrelation function being generally a function of velocity of the user, the propagation geometry, and the antenna characteristics. Basically, this is a Gauss-Markov model of low order (1) for reasons of computational complexity and tractability. Actually, the current channel between the BS and the typical user belonging to its cell is modelled as

$$\mathbf{h}_k[n] = \delta \mathbf{h}_k[n-1] + \mathbf{e}_k[n], \quad (8)$$

where $\mathbf{h}_k[n-1]$ is the channel in the previous symbol duration and $\mathbf{e}_k[n] \in \mathbb{C}^N$ is an uncorrelated channel error due to the channel variation modelled as a stationary Gaussian random process with i.i.d. entries and distribution $\mathcal{CN}(\mathbf{0}, (1 - \alpha^2)\mathbf{I}_N)$ [54].

Given that the variation of the channel is described by means of its second order statistics, we employ the autocorrelation function of the channel, which is an appropriate measure. We choose the Jakes model for representing the autocorrelation function, which is widely accepted due to its generality and simplicity [40]. The Jakes model describes a propagation medium with two-dimensional isotropic scattering and a monopole antenna at the receiver [55]. Mathematically, the

normalised discrete-time autocorrelation function of the fading channel is expressed by

$$r_h[k] = J_0(2\pi f_D T_s |k|). \quad (9)$$

Specifically, f_D and T_s are the maximum Doppler shift and the channel sampling period, respectively. Regarding the maximum Doppler shift f_D , it can be expressed in terms of the relative velocity of the v , i.e., $f_D = \frac{vf_c}{c}$, where $c = 3 \times 10^8$ m/s is the speed of light and f_c is the carrier frequency. Note that we assume that the base stations have perfect knowledge of $\delta = r_h[1]$.

Interestingly, we are able to combine both effects of pilot contamination and time-variation of the channel according to [37], [38]. Specifically, we have that the fading channel at time slot n can be expressed by

$$\begin{aligned} \mathbf{h}_k[n] &= \delta \mathbf{h}_k[n-1] + \mathbf{e}_k[n] \\ &= \delta \hat{\mathbf{h}}_k[n-1] + \tilde{\mathbf{e}}_k[n], \end{aligned} \quad (10)$$

where $\hat{\mathbf{h}}_k[n-1]$ and $\tilde{\mathbf{e}}_k[n] = \delta \tilde{\mathbf{h}}_k[n-1] + \mathbf{e}_k[n] \sim \mathcal{CN}(\mathbf{0}, \sigma_{\tilde{\mathbf{e}}_k}^2 \mathbf{I}_M)$ with $\sigma_{\tilde{\mathbf{e}}_k}^2 = \left(1 - \delta^2 \sigma_{\hat{\mathbf{h}}_k}^2\right)$ are mutually independent. In other words, the estimated channel at time n is now $\hat{\mathbf{h}}_k[n] = \delta \hat{\mathbf{h}}_k[n-1]$.

Remark 4: Imperfect CSIT is the result of three different sources, namely, i) pilot contamination, ii) channel aging, and iii) RATHIs.

V. DOWNLINK TRANSMISSION

The channel power distribution of a link depends upon its physical representation, i.e., if it describes the desired or the interference contribution to the received signal or we address to a multi-antenna BS deployment with single or multi-user transmission. However, in order to arrive at the stage of the statistical distribution of the power, we have to model the downlink transmission. Note that if $\delta = 0$, we obtain a static environment with no user mobility.

This paper assumes linear precoding by the matrix $\mathbf{W}[n] = [\mathbf{w}_1, \dots, \mathbf{w}_K] \in \mathbb{C}^{M \times K}$, employed by the associated BS, which multiplies the data signal vector $\mathbf{d}[n] = [d_1[n], \dots, d_K[n]]^T \in \mathbb{C}^M \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ for all users in that cell. In our case, we employ ZF precoding that has the form

$$\hat{\mathbf{W}}[n] = \bar{\mathbf{H}}[n] (\bar{\mathbf{H}}^H[n] \bar{\mathbf{H}}[n])^{-1} \quad (11)$$

$$= \delta^{-1} \bar{\mathbf{H}}^\dagger[n-1] = \delta^{-1} \hat{\mathbf{W}}[n-1], \quad (12)$$

where $\bar{\mathbf{H}}[n]$ is the normalised version of $\hat{\mathbf{H}}[n]$ and is related to $\hat{\mathbf{H}}[n]$ as defined in Appendix A, while (\cdot) denotes the pseudo-inverse of a matrix. Thus, the precoder is normalised and the average transmit power per user of the associated BS is constrained to p , i.e., $\mathbb{E}[\text{tr}(\hat{\mathbf{W}}[n] \hat{\mathbf{W}}^H[n])] = 1$. Note that $\hat{\mathbf{W}}[n]$ denotes the ZF precoder of the associated BS at time n as well as in (12) we have introduced the user mobility effect for the k th user by means of $\hat{\mathbf{h}}_k[n] = \delta \hat{\mathbf{h}}_k[n-1]$.

Thus, the received signal from the associated BS to the typical user at $x_{k,n}$ during the n th time-slot after applying the ZF precoder, accounting for a quasi-static block fading

model with frequency-flat fading channels varying for symbol to symbol and RATHIs, can be expressed as⁴

$$y_k[n] = \mathbf{h}_k^H[n] \|x_{k,n}\|^{-\alpha/2} (\mathbf{s}[n] + \boldsymbol{\eta}_{t,n}^{\text{BS}}) + \eta_{r,n}^{\text{UE}} + \sum_{l \in \Phi_B/x_{k,n}} \mathbf{g}_{lk}^H[n] \mathbf{s}_{lk}[n] \|y_{lk,n}\|^{-\alpha/2}, \quad (13)$$

where $\mathbf{s}[n] = \mathbf{W}[n] \mathbf{d}[n] \in \mathbb{C}^{M \times 1}$ is the transmit signal vector for the k th user with covariance matrix $\mathbf{Q}_{\text{BS}}[n] = \mathbb{E}[\mathbf{s}[n] \mathbf{s}^H[n]]$ and $p = \text{tr}(\mathbf{Q}_{\text{BS}}[n])$ is the associated average transmit power. The channel vector $\mathbf{h}_k[n] \in \mathbb{C}^{M \times 1}$ denotes the desired channel vector between the associated BS located at $x_{k,n} \in \mathbb{R}^2$ and the typical user at time-instance n . Similarly, $\mathbf{g}_{lk}[n] \in \mathbb{C}^{M \times 1}$ expresses the interference channel vector from the BSs found at $y_{lk,n} \in \mathbb{R}^2$ far from the typical user at time-instance n . Especially, in the case of Rayleigh fading, the channel power distributions of both the direct and the interfering links follow the Gamma distribution [56].

Given that we have assumed the realistic scenario of RATHIs as well as imperfect CSI due to pilot contamination and time-variation of the channel (channel aging) (see (10) and (12)), the received signal by user k can be written as^{5, 6}

$$y_k[n] = \hat{\mathbf{h}}_k^H[n-1] \hat{\mathbf{w}}_k[n-1] \mathbf{d}[n] \|x_{k,n}\|^{-\alpha/2} + \delta^{-1} \tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{W}}[n-1] \mathbf{d}[n] \|x_{k,n}\|^{-\alpha/2} + \mathbf{h}_k^H[n] \|x_{k,n}\|^{-\alpha/2} \boldsymbol{\eta}_{t,n}^{\text{BS}} + \eta_{r,n}^{\text{UE}} + \sum_{l \in \Phi_B/x_{k,n}} \mathbf{g}_{lk}^H[n] \hat{\mathbf{W}}_l[n] \mathbf{d}_l[n] \|y_l\|^{-\alpha/2}, \quad (14)$$

where we have used (10) for replacing the current desired channel⁷. In addition, we have used (12) to express the current precoder in terms of its delayed instance known at the BS. The first term in (14) represents the desired signal contribution in the typical current cell, while the second term describes the estimation error effect. Further, the third term describes the contribution due to the transmit BS impairment. The fourth term represents the receive impairment part because of the hardware impairments at the user side. Moreover, the next term characterises the inter-cell interference coming from BSs in different cells.

$\boldsymbol{\eta}_{t,n}^{\text{BS}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Lambda}_n^{\text{BS}})$ and $\eta_{r,n}^{\text{UE}} \sim \mathcal{CN}(\mathbf{0}, \Upsilon_n^{\text{UE}})$ are the residual downlink additive Gaussian distortions at the BS and the UE, which are given by (1) and (2) for $i = \text{BS}$ and $j = \text{UE}$, respectively. Specifically, we have

$$\boldsymbol{\Lambda}_n^{\text{BS}} = \kappa_{t,\text{BS}}^2 \text{diag}(q_1[n], \dots, q_M[n]) \quad (15)$$

$$\Upsilon_n^{\text{UE}} = \kappa_{r,\text{UE}}^2 \|x_{k,n}\|^{-\alpha} \mathbf{h}_k^H[n] \mathbf{Q}_{\text{BS}}[n] \mathbf{h}_k[n]. \quad (16)$$

Remark 5: The receive distortion at user k includes the path-loss coming from the associated BS.

⁴Since we focus on the investigation of the interference-limited SIR, its expression does not depend on the transmit power.

⁵We assume that the RATHIs from other BSs are negligible due to the increased path-loss, while the transmit hardware impairment depends only on the transmit signal power from the tagged BS and not from the path-loss.

⁶Herein, we assume that the thermal noise is negligible as compared to the distortion noises and interference from the other cells as shown by simulations and previously known results. Hence, its omission is reasonable. However, it can be included in the proposed analysis with no extra special effort.

⁷The replacement concerns only the current desired channel because the interference part is not of direct interest and can be kept in the initial form.

The achievable downlink SIR γ_k from the associated BS to the k th user can be defined as in (19) at the top of the next page. Note that we have assumed encoding of the message over many realisations of all sources of randomness in the model including noise, channel estimate error, and RATHIs. Based on this remark, the hardware impairments are written as $\boldsymbol{\eta}_{t,n}^{\text{BS}} \sim \mathcal{CN}(\mathbf{0}, \kappa_{t,\text{BS}}^2 \mathbf{I}_K)$ and $\eta_{r,n}^{\text{UE}} \sim \mathcal{CN}(\mathbf{0}, \kappa_{r,\text{UE}}^2 \|x_{k,n}\|^{-\alpha} \|\mathbf{h}_k[n]\|^2)$.

Remark 6: If the distortion parameters in (14) or in any other expression including the downlink RATHIs are set to zero, we result in the expressions corresponding to the ideal downlink model, which does not account for RATHIs [49].

In order to develop this general model, we assume that the desired channel power from the associated BS at time $n = \tau + 1, \dots, T_c$ located at $x_{k,n} \in \mathbb{R}^2$ to the typical user k , found at the origin, is given by $Z_k[n]$, while the interfering link from other BSs, i.e., the inter-cell interference from other cells' BSs (located at $y_{lk,n} \in \mathbb{R}^2$) is denoted by $I_l[n]$.

Proposition 1: The SIR of the downlink transmission from the associated BS to the typical user, accounting for RATHIs and imperfect CSI due to pilot contamination and time variation of the channel due to user mobility, can be represented as in (19) where the probability density function (PDF) of the desired signal power is obtained by

$$p_{Z_k[n]}(z) = \frac{e^{-z/\sigma_{\mathbf{g}_k}^2}}{(M-K)! \sigma_{\mathbf{h}_k}^2} \left(\frac{z}{\sigma_{\mathbf{h}_k}^2} \right)^{M-K}, \quad z \geq 0 \quad (20)$$

and the various terms are given by

$$Z_k[n] = |\hat{\mathbf{h}}_k^H[n-1] \hat{\mathbf{w}}_k[n-1]|^2 \quad (21)$$

$$E_k[n] = \delta^{-2} (1 + \kappa_{t,\text{BS}}^2) \|\tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{W}}[n-1]\|^2 \quad (22)$$

$$I_{\eta_t}[n] = \kappa_{t,\text{BS}}^2 |\hat{\mathbf{h}}_k^H[n-1] \hat{\mathbf{w}}_k[n-1]|^2 \quad (23)$$

$$I_{\eta_r}[n] = \kappa_{r,\text{UE}}^2 \|\mathbf{h}_k[n]\|^2 \quad (24)$$

$$I_l[n] = \sum_{l \in \Phi_B/x_{k,n}} |\mathbf{g}_{lk}[n] \hat{\mathbf{W}}_l[n]|^2 \|y_l\|^{-\alpha} = \sum_{l \in \Phi_B/x_{k,n}} g_{lk,n} \|y_l\|^{-\alpha}. \quad (25)$$

Note that \mathbf{w}_k is the k th column of \mathbf{W}_k . In addition, we have $I_{\eta_t}[n] \sim p_k \kappa_{t,\text{BS}}^2 \Gamma(\Delta, \sigma_{\mathbf{h}_k}^2)$, and $I_{\eta_r}[n] \sim p_k \kappa_{r,\text{UE}}^2 \Gamma(M, 1)$ while the total interference from all the other base stations found at a distance $\|y_{lk,n}\|$ from the typical user is $I_l[n] = \sum_{l \in \Phi_B/x_{k,n}} p_l g_{lk}[n] \|y_{lk,n}\|^{-\alpha}$, where $g_{lk}[n] = |\mathbf{g}_{lk}[n] \hat{\mathbf{W}}_l[n]|^2 \sim \Gamma(K, 1)$.

Proof: See Appendix A. \blacksquare

It is worthwhile to mention that the distance $\|x_{k,n}\|$ in the uplink and the downlink expresses two different variables and no confusion should arise. Specifically, during the uplink, the covariance of the estimated channel that includes $\|x_{k,n}\|^{-\alpha}$ is calculated for a given distance. Thus, it is deterministic, while its instance in the downlink is randomly distributed.

VI. MAIN RESULTS

Herein, we present the coverage probability and the achievable rate of the typical user, which are the main results of this work.

$$\gamma_k = \frac{Z_k[n] \|x_{k,n}\|^{-\alpha}}{(E_k[n] + I_{\eta_t}[n] + I_{\eta_r}[n]) \|x_{k,n}\|^{-\alpha} + I_l[n]}, \quad (19)$$

A. Coverage Probability

The focal point of this subsection is the derivation of an upper bound of the downlink coverage probability of the typical user in a multiple antenna HCN under the presence of RATHIs as well as imperfect CSIT due to pilot contamination and channel aging. Please note that the derivation of a lower bound, which is quite meaningful, is left for future work. Specifically, this subsection first presents a formal definition of the coverage probability with random BS locations drawn from a PPP. Next, the main result is provided with the technical derivation given in Appendix B. Remarkably, although we start from an abstract definition, we result in the most general and versatile expression known in the literature towards a more realistic assessment of a network with randomly located BSs. Interestingly, we present a more general result than [21], since now imperfect CSIT is assumed, when the BSs are randomly located. It is based on the calculation of the Laplace transforms provided by means of Proposition 2 and Lemma 1 provided below.

Definition 1 ([21], [49]): A typical user is in coverage if its effective downlink SIR from at least one of the randomly located BSs in the network is higher than the corresponding target T . In general, we have

$$p_c(T, \lambda_B, \alpha, \delta, q) \triangleq \mathbb{E} \left[\mathbf{1} \left(\bigcup_{x \in \Phi_B} \text{SIR}(x) > T \right) \right], \quad (26)$$

where q defines a set of parameters. Specifically, we define $q \triangleq \{\kappa_t^{\text{UE}}, \kappa_r^{\text{UE}}, \kappa_t^{\text{BS}}, \kappa_r^{\text{BS}}\}$.

Theorem 1: The upper bound of the downlink probability of coverage $p_c(T, \lambda_B, \alpha, \delta, q)$ in a general cellular network with randomly distributed multiple antenna BSs, accounting for RATHIs and imperfect CSIT due to pilot contamination and channel aging, is given by

$$\begin{aligned} p_c(T, \lambda_B, \alpha, \delta, q) &\leq \lambda_B \int_{l \in \mathbb{R}^2} \sum_{i=0}^{\Delta-1} \sum_{k=0}^{i-u_1-u_2-u_3=i-k} \binom{i}{k} \\ &\times \binom{i-k}{u_1+u_2+u_3} \frac{(-1)^i s^k}{i!} \frac{d^{u_1}}{ds^{u_1}} \mathcal{L}_{E_k}(l^{-\alpha} s) \\ &\times \frac{d^{u_2}}{ds^{u_2}} \mathcal{L}_{I_{\eta_t}}(l^{-\alpha} s) \frac{d^{u_3}}{ds^{u_3}} \mathcal{L}_{I_{\eta_r}}(l^{-\alpha} s) \frac{d^k}{ds^k} \mathcal{L}_{I_l}(s) dl, \quad (27) \end{aligned}$$

where $l = \|x\|$ and $s = \frac{T}{\sigma_{\hat{\mathbf{g}}_k}^2} l^\alpha$, while $\mathcal{L}_{I_{\eta_r}}(s)$, $\mathcal{L}_{I_{\eta_t}}(s)$, and $\mathcal{L}_{I_l}(s)$ are the Laplace transforms of the powers of the receive distortion, transmit distortion, and interference coming from other BSs.

Proof: See Appendix B⁸. ■

Proposition 2: The Laplace transform of the interference power of a general cellular network with randomly distributed multiple antenna BSs having RATHIs and imperfect CSIT is

given by

$$\mathcal{L}_{I_l}(s) = \exp\left(-s^{\frac{2}{\alpha}} \mathcal{C}(\alpha, K)\right), \quad (28)$$

where $\mathcal{C}(\alpha, K) = \frac{2\pi\lambda_B}{\alpha} \sum_{m=0}^K \binom{K}{m} \text{B}\left(K-m+\frac{2}{\alpha}, m-\frac{2}{\alpha}\right)$.

Proof: See Appendix C. ■

Lemma 1: The Laplace transforms of the parts, describing the RATHIs I_{η_t} and I_{η_r} as well as the estimation error, are given by

$$\mathcal{L}_{I_{\eta_t}}(s) = \frac{1}{(1 + \kappa_{\text{tBS}}^2 s)^\Delta}, \quad (29)$$

$$\mathcal{L}_{I_{\eta_r}}(s) = \frac{1}{(1 + \kappa_{\text{rUE}}^2 s)^M}, \quad (30)$$

$$\mathcal{L}_{E_k}(s) = \frac{1}{(1 + \delta^{-2} (1 + \kappa_{\text{tBS}}^2) \sigma_{\hat{\mathbf{e}}_k}^2 s)^\Delta}. \quad (31)$$

Proof: The first two Laplace transforms are easily obtained, since I_{η_t} and I_{η_r} follow the scaled Gamma distributions with scaled parameters κ_{tBS}^2 and κ_{rUE}^2 , as mentioned in Appendix A. In the same appendix, it is shown that the estimation error is a scaled Gamma distribution. ■

Remark 7: The expressions corresponding to the Laplace transforms of the distortion noises reveal that the downlink RATHIs show a different behaviour than the uplink impairments. Specifically, κ_{tBS}^2 has a greater impact than κ_{rUE}^2 , since $M > \Delta$. As a result, the quality of the transmit BS impairments should be kept above a certain standard that would not allow significant distortion of the system performance. Moreover, the Laplace transform of the transmit impairments depends on both the number of BS antennas and users, while the Laplace transform of the receive impairments depends only on the number of BS antennas. In other words, the higher the number of users is, the smaller the impact of the transmit impairments becomes, since the corresponding Laplace transform increases. Furthermore, the Laplace transform of the estimation error depends on the channel aging by means of δ and both transmit and receive hardware impairments. In fact, the more severe the channel aging is by means of smaller δ (higher users' velocity), the smaller $\mathcal{L}_{E_k}(s)$ becomes.

Lemma 2: The PDF of $g_{lk,n}$, describing the interfering marks, is $\Gamma(K, 1)$ distributed.

Proof: Given that the precoding matrices $\hat{\mathbf{W}}_l[n]$ have unit-norm columns and each instance does not depend on $\mathbf{g}_{lk}[n]$ and $\hat{\mathbf{g}}_{lk}[n]$ (the normalised version of $\mathbf{g}_{lk}[n]$), $\hat{\mathbf{W}}_l[n]$ are independent isotropic unit-norm random vectors. Hence, the sum of $\mathbf{g}_{lk}[n]$ expresses the squared modulus of a linear combination of K complex normal random variables, i.e., $g_{lk,n} \sim \Gamma(K, 1)$. ■

Corollary 1: When $M = K$, i.e., in the special case of full SDMA, the upper bound of the coverage probability with RATHIs and channel aging, described by Theorem 1, is given

⁸Although the expressions [21]-[25] could be characterised in the asymptotic regime, this action would result in deterministic expressions that could not be manipulated statistically, in order to derive the coverage probability.

by

$$p_c(T, \lambda_B, \alpha, \delta, q) \leq \lambda_B \int_{l \in \mathbb{R}^2} \mathcal{L}_{E_k}(l^{-\alpha} s) \mathcal{L}_{I_{n_t}}(l^{-\alpha} s) \times \mathcal{L}_{I_{n_r}}(l^{-\alpha} s) \mathcal{L}_{I_l}(s) dl. \quad (32)$$

B. Average Achievable Rate

This subsection presents the mean achievable data rate for a typical user under the proposed general system model with RATHIs and imperfect CSIT. Given that the analysis and some definitions are similar to Section VI-A, the presentation is more concise. Actually, the following theorem is one of the main results of this paper, being unique in the research area of practical systems with hardware impairments, when the BSs are randomly positioned.

Theorem 2: The downlink average achievable user rate of a HCN in the presence of RATHIs and imperfect CSIT due to pilot contamination and channel aging is

$$R_k(T, \lambda_B, \alpha, \delta, q) = \int_{x \in \mathbb{R}^2} \int_{t > 0} \sum_{i=0}^{\Delta-1} \sum_{k=0}^i \sum_{u_1+u_2+u_3=i-k} \binom{i}{k} \times \binom{i-k}{u_1+u_2+u_3} \frac{(-1)^i s^k}{i!} \frac{d^{u_1}}{ds^{u_1}} \mathcal{L}_{E_k}(l^{-\alpha} s (e^t - 1)) \times \frac{d^{u_2}}{ds^{u_2}} \mathcal{L}_{I_{n_t}}(l^{-\alpha} s (e^t - 1)) \frac{d^{u_3}}{ds^{u_3}} \mathcal{L}_{I_{n_r}}(l^{-\alpha} s (e^t - 1)) \times \frac{d^k}{ds^k} \mathcal{L}_{I_l}(s (e^t - 1)) dt dl, \quad (33)$$

where the various parameters are also defined in Theorem 1.

Proof: See Appendix D. ■

VII. NUMERICAL RESULTS

This section aims at investigating the impact of the various parameters such as number of users and antennas on the general expressions corresponding to the coverage probability and the user rates provided by Theorems 1 and 2, respectively. Given that the typical user lies at the origin, we choose a sufficiently large area of 50 km × 50 km, where the locations of the BSs are simulated as realisations of a PPP with given density $\lambda_B = 0.01$. The users' PPP density is considered to be $\lambda_k = 60\lambda_B$ as in [35].

Coverage of the user means that the received SIR from at least one of the BSs exceeds a certain target. We are interested in the calculation of the desired signal strength and interference power. Note that the received SIR is obtained from each BS. The desired estimate of the coverage probability comes by repeating this scheme an adequate number of times. Thus, this procedure allows us not only to validate our model, but also to demonstrate the effect of the various parameters on the coverage probability.

Herein, we consider a setup, where the number of antennas per BS and the number of users are $M = 5$ and $K = 3$, respectively. The path-loss is set to $\alpha = 3$, while the uplink and downlink transmit powers are $\rho_{\text{up}}^{\text{UE}} = 5$ dB and $p = 15$ dB. Note that $\tau = K$, hence $p_{\text{tr}} = 15$ dB. Moreover, we assume that the distance between the user and the BS during the uplink phase is $\|x_k\| = 15$ m, and it is known by the BS. In the

figures, the proposed analytical expressions of the coverage probability $p_c(T, \lambda_B, \alpha, \delta, q)$ and the achievable user rate $R_k(T, \lambda_B, \alpha, \delta, q)$ are plotted along with the corresponding simulated results. The “dot” and “solid” lines illustrate the analytical results with specific RATHIs/user mobility and no transceiver impairments or no relative user movement. In a similar way, the bullets designate the simulation results. Obviously, the inevitable presence of RATHIs or the time variation of the channel result in the worsening of the system performance. Actually, the more severe these effects are, the higher the degradation of the system performance becomes.

A. Impact of RATHIs

In order to investigate how the RATHIs affect the coverage probability, we assume no user mobility, i.e., $\delta = 0$. First, we focus on the uplink RATHIs κ_{rBS}^2 and κ_{tUE}^2 , while we have assumed that the downlink RATHIs are zero. Specifically, in Fig 1, we show the variation of the coverage probability for different values of uplink hardware impairments. Based on this model, the transmit impairment of the user and the receive impairment at the BS contribute the same at the coverage probability, which is in agreement with theory. For example, we set $\kappa_{\text{tUE}}^2 = 0.08$ and $\kappa_{\text{rBS}}^2 = 0$ as well as $\kappa_{\text{tUE}}^2 = 0$ and $\kappa_{\text{rBS}}^2 = 0.08$ and the coverage probability does not change. Thus, as a design plan, we could keep κ_{tUE}^2 constant and play with the quality of the BS hardware for obtaining a certain coverage probability.

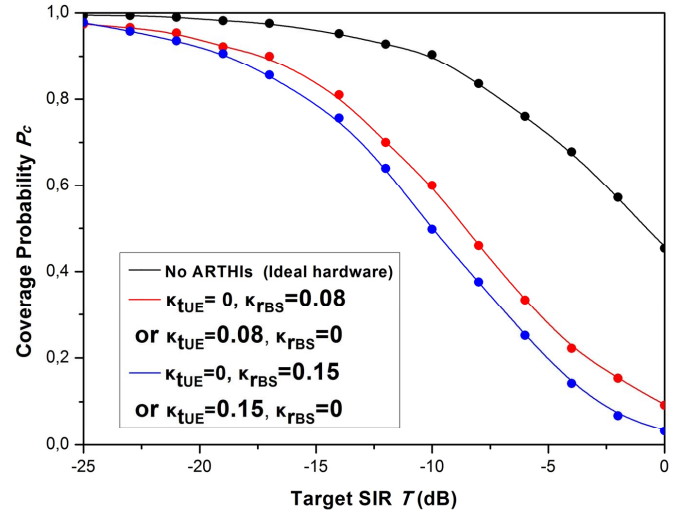


Fig. 1. Coverage probability of a multiple-antenna HCN for varying uplink RATHIs versus ρ ($M = 5$, $K = 3$, $\alpha = 3$, $\kappa_{\text{tBS}} = \kappa_{\text{tUE}} = 0$, $\delta = 0$).

In Fig. 2, we plot the coverage probability as a function of the target SIR for different values of the downlink RATHIs. Clearly, here the transmit hardware impairment κ_{tBS} exhibits higher impact on $p_c(q)$ than κ_{rUE} , which coincides with the theoretical observations. For this reason, we should be more careful with the quality of the transmit impairments at the BS. These nominal values of RATHIs are quite reasonable according to [51]. Moreover, in the same figure, we have depicted the simulated and theoretical results corresponding to

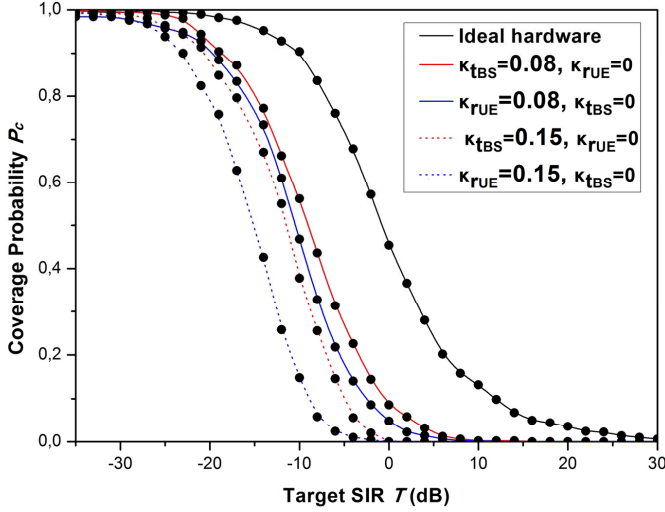


Fig. 2. Coverage probability of a multiple-antenna HCN for varying downlink RATHIs versus ρ ($M = 5$, $K = 3$, $\alpha = 3$, $\kappa_{tUE} = \kappa_{rBS} = 0$, $\delta = 0$).

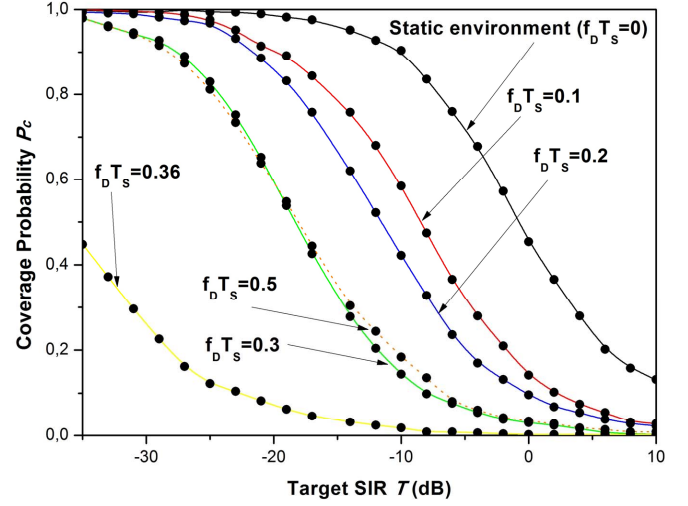


Fig. 4. Achievable user rate of a HCN versus ρ for varying user mobility ($M = 5$, $K = 3$, $\alpha = 3$, $\kappa_{tUE} = \kappa_{rBS} = \kappa_{rUE} = \kappa_{rUE} = 0$, $\delta = 0$).

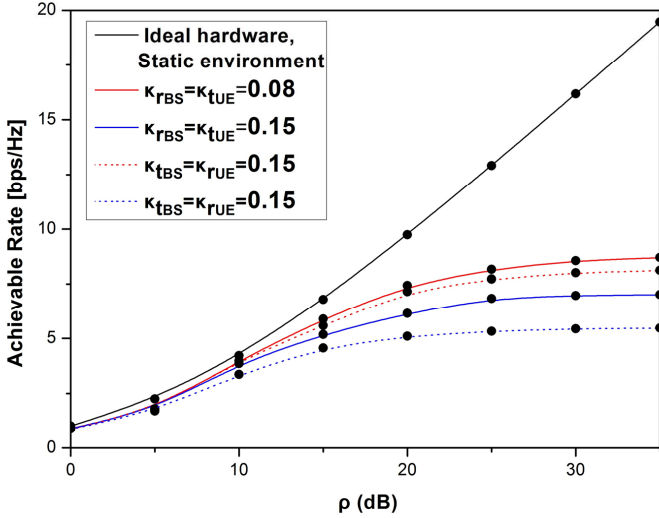


Fig. 3. Achievable user rate of a multiple-antenna HCN for varying uplink and downlink RATHIs versus ρ ($M = 5$, $K = 3$, $\alpha = 3$, $\delta = 0$).

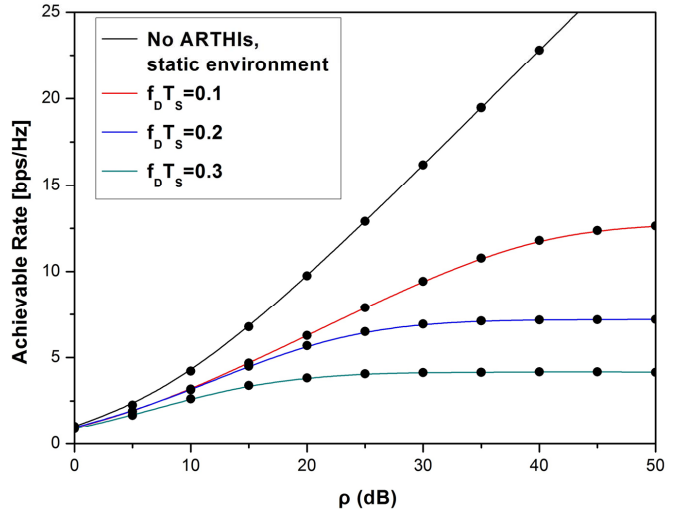


Fig. 5. Coverage probability of a multiple-antenna HCN versus ρ for varying Doppler shift ($M = 5$, $K = 3$, $\alpha = 3$, $\kappa_{tUE} = \kappa_{rBS} = \kappa_{rUE} = \kappa_{rUE} = 0$, $\delta = 0$).

ideal hardware for the sake of comparison. Similar conclusions can be extracted by observing the rate versus ρ for varying downlink RATHIs in Fig 3.

B. Impact of User Mobility

We study the effect of time-variation of the channel due to user mobility on the coverage probability of a cellular network in Fig. 4. These observations are consistent with the coverage probability and user rate results derived in Sections VI-A and VI-B, respectively. Note that in order to focus only on the effect of time-variation of the channel, we assume that the hardware is ideal, i.e., both in the downlink and the uplink the RATHIs are set to zero. Clearly, it is shown that an increase of the time variation of the channel results in a decrease of $p_c(T, \lambda_B, \alpha, \delta, q)$. This behaviour continues till the normalised Doppler shift $f_D T_s$ becomes $f_D T_s \approx 0.36$. Then,

increasing $f_D T_s$, the coverage probability is improved. Actually, behind this behaviour is hidden the variation of the Jakes autocorrelation function with $f_D T_s$. In Fig. 5, we illustrate the achievable rate versus ρ . It is shown that the rate decreases with increasing time-variation, as expected.

C. Comparison between Transmission Techniques

We elaborate on the following three transmission setups: i) SISO with $M = K = 1$, i.e., we have a single transmit antenna per BS and one user ii) full SDMA with $M = K = 5$, i.e., 5 transmit antennas per BS serving 5 users iii) SU-BF transmission, in order to shed light on the realistic performance of HCNs, as far as the system parameters M and K change. Furthermore, we choose a set of parameters of RATHIs and time variation, representing a realistic scenario, to be $q = \{0.03, 0.03, 0.08, 0.08\}$ and $\delta = 0.1$.

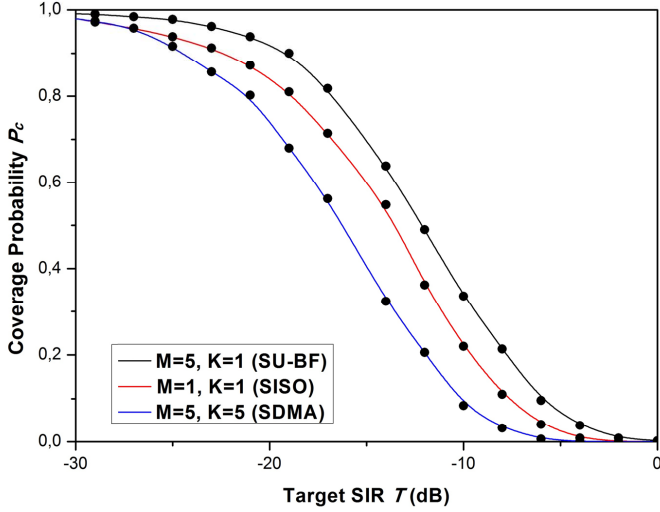


Fig. 6. Coverage probability of a multiple-antenna HCN versus ρ for various combinations of multi-antenna techniques ($\alpha = 3$, $\kappa_{\text{tUE}} = \kappa_{\text{rBS}} = \kappa_{\text{rUE}} = \kappa_{\text{tBS}} = 0.08$, $\delta = 0.1$).

In Fig. 6, we observe that the SU-BF transmission achieves higher coverage probability with comparison to SISO. Note that the latter is better than SDMA. In other words, we verify that it is better to serve a single user in each resource block, either by SISO or SU-BF, instead of serving multiple users. However, herein and with comparison to [49], we illustrate how, given this property, the coverage probability varies with SIR in the presence of the RATHIs. Similar observations have been mentioned in [21] and in [49] for the cases of RATHIs and perfect CSIT with ideal hardware, respectively. Especially, we illustrate that SU-BF transmission proves to be better with comparison to SISO in some regimes/situations because, in addition to the proximity gain enjoyed by the SISO due to extreme densification, the SU-BF transmission presents an additional beamforming gain. Moreover, SISO outperforms SDMA. More concretely, instead of serving multiple users, it is preferable to serve a single user in each resource block by means of SISO or SU-BF. Herein, it is worthwhile to mention that in interference-limited networks, the use of more antennas for multi-stream transmission (SDMA) is not always beneficial, but it depends on which transmission/reception scheme is employed. Also, this claim depends on the performance metric we study. Interestingly, having this property in mind, we depict how the presence of the RATHIs affect the performance of the system in terms of the degradation of both coverage probability and user rate.

VIII. CONCLUSION

In this paper, we presented a new framework for the downlink cellular network analysis with imperfect CSIT. Specifically, we proposed a multiple antenna HCN that generalises the state of the art by accounting for imperfect CSIT due to pilot contamination, channel aging, and RATHIs. Furthermore, we quantified the performance loss due to imperfect CSIT. We showed that the uplink RATHIs have equivalent contribution to both metrics under investigation, while the downlink RATHIs

exhibit different impact. Specifically, the quality of the BS transmitter has greater impact than the receive hardware of the user. Moreover, we demonstrated the loss due to time variation in both coverage probability and user rate. In addition, we demonstrated the outperformance of SU-BF in terms of coverage and user-rate against both SISO and SDMA in the case of a more realistic analysis than previous works did. In other words, we confirmed that, given a total number of transmit antennas, it is better to share them across many BSs than collecting them in a few BSs. Actually, the proposed framework is of high interest because it allows us to explore how various multiple antenna techniques affect the coverage and the rate in realistic conditions.

APPENDIX A PROOF OF PROPOSITION 1

Initially, we assume that the columns of the precoding matrix $\hat{\mathbf{W}}[n]$ with unit norm equal the normalised columns of $\hat{\mathbf{H}}^H[n] \left(\hat{\mathbf{H}}[n] \hat{\mathbf{H}}^H[n] \right)^{-1}$. In other words, $\hat{\mathbf{W}}[n] = \bar{\mathbf{H}}^H[n] \left(\bar{\mathbf{H}}[n] \bar{\mathbf{H}}^H[n] \right)^{-1}$, where $\bar{\mathbf{H}}[n] = [\bar{\mathbf{h}}_1[n], \dots, \bar{\mathbf{h}}_K[n]] \in \mathbb{C}^{(M \times K)}$ with columns $\bar{\mathbf{h}}[n] = \frac{\hat{\mathbf{h}}[n]}{\|\hat{\mathbf{h}}[n]\|}$. As a result, the numerator of the SIR in (19), describing the desired signal power, is $\Gamma(\Delta, \sigma_{\hat{\mathbf{h}}_k}^2)$ distributed with $\Delta = M - K + 1$ (or else it follows the Erlang distribution with shape and scale parameters Δ and $\sigma_{\hat{\mathbf{h}}_k}^2$, respectively), since $Z_k[n] = |\bar{\mathbf{h}}_k^H[n] \mathbf{w}_k[n]|^2 \cdot \|\mathbf{h}_k[n]\|^2$. Actually, $Z_k[n]$ can be written as the product of two independent random variables distributed as $B(M - K + 1, K - 1)$ and $\Gamma(M, \sigma_{\hat{\mathbf{h}}_k}^2)$, respectively [57]⁹. Note that ZF beamforming has been applied. Therefore, the resultant PDF of the product follows the $\Gamma(\Delta, \sigma_{\hat{\mathbf{h}}_k}^2)$ distribution. The term in the denominator, concerning the error, is expressed in terms of a sum of K independent random variables as

$$\begin{aligned} E_k[n] &= \delta^{-2} (1 + \kappa_{\text{tBS}}^2) \|\tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{W}}[n-1]\|^2 \\ &= \delta^{-2} (1 + \kappa_{\text{tBS}}^2) \sum_{i=1}^K |\tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{w}}_i[n-1]|^2. \end{aligned}$$

Since $|\tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{w}}_i[n-1]|^2$ is the squared modulus of a linear combination of M complex random variables with $\hat{\mathbf{w}}_i[n-1]$ being independent and having unit norm, it is distributed as $\Gamma(1, \sigma_{\tilde{\mathbf{e}}_k}^2)$. Thus, $\|\tilde{\mathbf{e}}_k^H[n] \hat{\mathbf{W}}[n-1]\|^2$ is $\Gamma(K, \sigma_{\tilde{\mathbf{e}}_k}^2)$ distributed. Taking the expectation over the transmit distortion noise of the tagged BS, then $I_{\eta_{\text{tBS}}^{\text{BS}}}[n] = p\kappa_{\text{tBS}}^2 \|\mathbf{h}_k^H[n]\|^2$, which follows a scaled $\Gamma(M, \sigma_{\hat{\mathbf{h}}_k}^2)$ distribution. Following a similar procedure, we take the expectation over the receive distortion noise $\eta_{\text{r},n}^{\text{BS}}$. We obtain $I_{\eta_{\text{r},n}^{\text{UE}}}[n] = p\kappa_{\text{rUE}}^2 \|\mathbf{h}_k[n]\|^2$ following a scaled $\Gamma(M, 1)$. Regarding the other term in the denominator, it represents the interference from other BSs, $I_l[n]$. Especially, $g_{lk,n} = |\mathbf{g}_{lk}^H[n] \mathbf{W}_l[n]|^2 \sim \Gamma(K, 1)$ because $\mathbf{W}_l[n-1]$ being the precoding matrices from other BSs have unit-norm and are independent from the normalised $\bar{\mathbf{g}}_{lk}[n]$. On this ground, $g_{lk,n}$

⁹The random variable $\|\hat{\mathbf{h}}_k^H[n]\|^2$ is the linear combination of M i.i.d. exponential random variables each with variance $\sigma_{\hat{\mathbf{h}}_k}^2$, i.e., $\|\hat{\mathbf{h}}_k^H[n]\|^2 \stackrel{d}{\sim} \Gamma[M, \sigma_{\hat{\mathbf{h}}_k}^2]$.

is a linear combination of K independent complex normal random variables with unit variance, i.e., $g_{lk,n} \sim \Gamma(K, 1)$.

APPENDIX B PROOF OF THEOREM 1

Starting from the definition of $p_c(T, \lambda_B, \alpha, \delta, q)$ and by means of appropriate substitution of the SIR γ_k , we have

$$p_c(T, \lambda_B, \alpha, \delta, q) = \mathbb{E} \left[\mathbf{1} \left(\bigcup_{x_k, n \in \Phi_B} \text{SIR}(x_k, n) > T \right) \right] \quad (34)$$

$$\leq \mathbb{E} \left[\bigcup_{x_k, n \in \Phi_B} \mathbf{1}(\text{SIR}) > T \right] \quad (35)$$

$$= \mathbb{E} \left[\sum_{x_k, n \in \Phi_B} \mathbb{P}[\text{SIR} > T | l] \right] \quad (36)$$

$$= \lambda_B \int_{x \in \mathbb{R}^2} \mathbb{E}[\mathbb{P}[Z_k > T(E_k[n] + I_{\eta_t}[n] + I_{\eta_r}[n]) + Tl^\alpha I_l | l]] dx, \quad (37)$$

where in (35), we have employed the union bound or else the Boole's inequality, and in (37), we have used the Campbell-Mecke Theorem [47]. Given that $Z_k[n]$ is Gamma distributed, its PDF is provided by (20). Below, we have removed the index n for the ease of exposition. We turn our attention to the integrable part of (37). This can be expressed as

$$\begin{aligned} \mathbb{P}[Z_k > T(E_k + I_{\eta_t} + I_{\eta_r}) + Tl^\alpha I_l | l] &= e^{-\frac{T(E_k + I_{\eta_t} + I_{\eta_r})}{\sigma_{\mathbf{g}_k}^2}} \\ &\times e^{-\frac{Tl^\alpha I_l}{\sigma_{\mathbf{g}_k}^2} \sum_{i=0}^{\Delta-1} \sum_{k=0}^i \binom{i}{k} \frac{(T(E_k + I_{\eta_t} + I_{\eta_r}))^{i-k} (Tl^\alpha I_l)^k}{i! (\sigma_{\mathbf{g}_k}^2)^i}}, \end{aligned} \quad (38)$$

where in (38), we have applied the Binomial theorem. In addition, if we apply the expectation operator, we obtain

$$\begin{aligned} \mathbb{E}[\mathbb{P}[Z_k > T(I_{\eta_t} + I_{\eta_r}) + Tl^\alpha I_l | l]] &= \sum_{i=0}^{\Delta-1} \sum_{k=0}^i \sum_{u_1+u_2+u_3=i-k} \binom{i}{k} \\ &\times \binom{i-k}{u_1+u_2+u_3} \frac{(-1)^i s^k}{i!} \frac{d^{u_1}}{ds^{u_1}} \mathcal{L}_{E_k}(s) \frac{d^{u_2}}{ds^{u_2}} \mathcal{L}_{I_{\eta_t}}(s) \\ &\times \frac{d^{u_3}}{ds^{u_3}} \mathcal{L}_{I_{\eta_r}}(s) \frac{d^k}{ds^k} \mathcal{L}_{I_l}(s), \end{aligned} \quad (39)$$

where we have set $s = \frac{T}{\sigma_{\mathbf{g}_k}^2} l^\alpha$. We result in (39), after using the Multinomial theorem. Note that the inner sum is taken over all combinations of nonnegative integer indices u_1 through u_3 such that the sum $u_1 + u_2 + u_3$ is $i - k$. Moreover, we have used the definition of the Laplace Transform $\mathbb{E}_I[e^{-sI}(sI)^i] = s^i \mathcal{L}\{t^i g_I(t)\}(s)$ and the Laplace identity $t^i g_I(t) \longleftrightarrow (-1)^i \frac{d^i}{ds^i} \mathcal{L}\{g_I(t)\}(s)$. Regarding the Laplace transform $\mathcal{L}_{I_l}(s)$, it is obtained by means of Proposition 2, while the transforms $\mathcal{L}_{I_{\eta_t}}$ and $\mathcal{L}_{I_{\eta_r}}(s)$ are provided by Lemma 1. Finally, the proof is concluded after substituting (39) into (37).

APPENDIX C PROOF OF PROPOSITION 2

The Laplace transform of the interference power relies on its PDF. Specifically, setting $g_l \triangleq g_{lk,n}$, accounting for the power of the interference channel, has identical distribution for all l , $\mathcal{L}_{I_l}(s)$ can be obtained as

$$\begin{aligned} \mathcal{L}_{I_l}(s) &= \mathbb{E}_{I_l}[e^{-sI_l}] = \mathbb{E}_{I_l} \left[e^{-s \sum_{l \in \Phi_B \setminus x} g_l y^{-\alpha}} \right] \\ &= \mathbb{E}_{\Phi_B, g_l} \left[\prod_{l \in \Phi_B \setminus x} e^{-s g_l y^{-\alpha}} \right] \end{aligned} \quad (40)$$

$$= \mathbb{E}_{\Phi_B} \left[\prod_{l \in \Phi_B \setminus x} \mathcal{L}_{g_l}(s y^{-\alpha}) \right] \quad (41)$$

$$= \exp \left(-\lambda_B \int_{\mathbb{R}^2} (1 - \mathcal{L}_{g_l}(s y^{-\alpha})) dy \right) \quad (42)$$

$$= \exp \left(-2\pi \lambda_B \int_0^\infty \left(\frac{\sum_{m=1}^K \binom{K}{m} (s r^{-\alpha})^m}{(1 + s r^{-\alpha})^K} r dr \right) \right) \quad (43)$$

$$= \exp \left(-\frac{2\pi \lambda_B s^{\frac{2}{\alpha}}}{\alpha} \sum_{m=1}^K \binom{K}{m} B \left(K - m + \frac{2}{\alpha}, m - \frac{2}{\alpha} \right) \right).$$

In particular, (40) results from the independence among the locations of the BSs. Moreover (40) holds due to the independence between the spatial and the fading distributions. The property of the probability generating functional (PGFL) regarding the PPP [53] is used to obtain (42). Furthermore, we substitute the Laplace transform of g_l following a Gamma distribution. Note that in the next step, application of the Binomial theorem takes place, while we continue with conversion of the Cartesian coordinates to polar coordinates in (43). Finally, we conclude the proof by calculating the integral in the following way. We make the substitution $(1 + r^{-\alpha})^{-1} \rightarrow t$, and after many algebraic manipulations, and the use of [46, Eq. (8.380.1)] we lead to the desired result.

APPENDIX D PROOF OF THEOREM 2

The achievable rate of the typical user at the origin is derived as

$$R_k(T, \lambda_B, \alpha, \delta, q) = \mathbb{E}[\ln(1 + \text{SIR})] \quad (44)$$

$$= \int_{x \in \mathbb{R}^2} \mathbb{E}[\ln(1 + \text{SIR})] dl \quad (45)$$

$$= \int_{x \in \mathbb{R}^2} \int_{t>0} \mathbb{P}[\ln(1 + \text{SIR})] dt dl \quad (46)$$

$$= \int_{x \in \mathbb{R}^2} \int_{t>0} \mathbb{P}[h_k > (T(I_{\eta_t} + I_{\eta_r}) + Tl^\alpha I_l)(e^t - 1)] dt dl \quad (47)$$

$$\begin{aligned} &= \int_{x \in \mathbb{R}^2} \int_{t>0} \mathbb{E} \left[e^{-\frac{T}{\sigma_{\mathbf{g}_k}^2} (E_k + I_{\eta_t} + I_{\eta_r})(e^t - 1) - \frac{T}{\sigma_{\mathbf{g}_k}^2} l^\alpha I_l (e^t - 1)} \right. \\ &\times \left. \sum_{i=0}^{\Delta-1} \sum_{k=0}^i \binom{i}{k} \frac{(T(E_k + I_{\eta_t} + I_{\eta_r})(e^t - 1))^{i-k} (Tl^\alpha I_l (e^t - 1))^k}{i! (\sigma_{\mathbf{g}_k}^2)^i} \right] dt dl \end{aligned} \quad (48)$$

$$\begin{aligned}
&= \int_{x \in \mathbb{R}^2} \int_{t > 0} \sum_{i=0}^{\Delta-1} \sum_{k=0}^i \sum_{u_1+u_2+u_3=i-k} \binom{i}{k} \binom{i-k}{u_1+u_2+u_3} \frac{(-1)^i s^{ik}}{i!} \\
&\times \frac{d^{u_1}}{ds^{u_1}} \mathcal{L}_{E_k}(l^{-\alpha} s (e^t - 1)) \frac{d^{u_2}}{ds^{u_2}} \mathcal{L}_{I_{n_t}}(l^{-\alpha} s (e^t - 1)) \\
&\times \frac{d^{u_3}}{ds^{u_3}} \mathcal{L}_{I_{n_r}}(l^{-\alpha} s (e^t - 1)) \frac{d^k}{ds^k} \mathcal{L}_{I_l}(s (e^t - 1)) dt dl, \quad (49)
\end{aligned}$$

where the expectation in (44) is taken over the fading distribution and the spatial randomness described by a PPP. In (45), we have applied the property, concerning a positive random variable x , $\mathbb{E}[x] = \int_{t>0} \mathbb{P}(x > t) dt$. Next, we consider that h_k follows a Gamma distribution with shape and rate parameters, Δ and 1, respectively.

REFERENCES

- [1] J. Thompson *et al.*, "5G wireless communication systems: Prospects and challenges [guest editorial]," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 62–64, 2014.
- [2] V. W. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems*. Cambridge university press, 2017.
- [3] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels—parts I & II," *IEEE Trans. on Inf. Theory*, vol. 43, no. 6, pp. 1877–1911, 1997.
- [4] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.
- [5] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. on Commun.*, vol. 59, no. 11, pp. 3122–3134, 2011.
- [6] H. S. Dhillon *et al.*, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Sel. Areas in Commun.*, vol. 30, no. 3, pp. 550–560, 2012.
- [7] H.-S. Jo *et al.*, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, 2012.
- [8] J. G. Andrews *et al.*, "Femtocells: Past, present, and future," *IEEE Journal on Sel. Areas in Commun.*, vol. 30, no. 3, pp. 497–508, 2012.
- [9] T. Schenk, *RF imperfections in high-rate wireless systems: Impact and digital compensation*. Springer Science & Business Media, 2008.
- [10] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *ITG/IEEE Work. Smart Ant. (WSA)*, 2010, pp. 189–196.
- [11] B. Goransson *et al.*, "Effect of transmitter and receiver impairments on the performance of MIMO in HSDPA," in *IEEE 9th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2008, pp. 496–500.
- [12] E. Björnson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *IEEE Global Commun. Conf. (GLOBECOM)*, 2012, Dec 2012, pp. 4775–4780.
- [13] E. Björnson *et al.*, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, 2013.
- [14] J. Qi and S. Aïssa, "On the power amplifier nonlinearity in MIMO transmit beamforming systems," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 876–887, 2012.
- [15] H. Mehrpouyan *et al.*, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Proc.*, vol. 60, no. 9, pp. 4790–4807, Sept 2012.
- [16] A. K. Papazafeiropoulos, S. K. Sharma, and S. Chatzinotas, "Impact of transceiver impairments on the capacity of dual-hop relay massive MIMO systems," in *2015 IEEE Globecom Workshops (GC Wkshps)*, Dec 2015, pp. 1–6.
- [17] A. Papazafeiropoulos, S. K. Sharma, and S. Chatzinotas, "MMSE filtering performance of DH-AF massive MIMO relay systems with residual transceiver impairments," in *IEEE Intern. Conf. on Commun. (ICC 2016), Kuala Lumpur, Malaysia*, May 2016.
- [18] A. Papazafeiropoulos, S. K. Sharma, S. Chatzinotas, and B. Ottersten, "Ergodic capacity analysis of amplify-and-forward dual-hop MIMO relay systems with residual transceiver hardware impairments: Conventional and large system limits," *IEEE Trans. on Veh. Tech.*, 2017.
- [19] A. Papazafeiropoulos, S. K. Sharma, R. T. Chatzinotas, S., and B. Ottersten, "Impact of transceiver hardware impairments on the ergodic channel capacity for Rayleigh-product MIMO channels," in *IEEE Signal Processing Advances in Wireless Communications (SPAWC 2016), Edinburgh, U.K.*, July 2016, pp. 1–6.
- [20] A. Papazafeiropoulos, S. K. Sharma, S. Chatzinotas, and T. Ratnarajah, "Impact of residual transceiver impairments on MMSE filtering performance of Rayleigh-product MIMO channels," in *IEEE Signal Processing Advances in Wireless Communications (SPAWC 2017), Japan*, July 2017.
- [21] A. K. Papazafeiropoulos and T. Ratnarajah, "Downlink MIMO HCNs with residual transceiver hardware impairments," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2023–2026, 2016.
- [22] A. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. on Veh. Tech.*, 2016. [Online]. Available: <http://arxiv.org/abs/1605.07661>
- [23] A. K. Papazafeiropoulos, "Downlink performance of massive MIMO under general channel aging conditions," in *Global Communications Conference (GLOBECOM)*, 2015 IEEE. IEEE, 2015, pp. 1–6.
- [24] A. Papazafeiropoulos, B. Clerckx, and T. Ratnarajah, "Rate-splitting to mitigate residual transceiver hardware impairments in massive MIMO systems," *IEEE Trans. on Veh. Tech.*, 2017.
- [25] —, "Mitigation of phase noise in massive MIMO systems: A rate-splitting approach," in *Proc. of IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.
- [26] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. no.8, pp. 4353–4368, Aug. 2015.
- [27] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 2001–2016, 2017.
- [28] J. G. Andrews *et al.*, "What will 5G be?" *IEEE Journal on Sel. in Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [29] K. Zheng *et al.*, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 3, pp. 1738–1760, 2015.
- [30] J. Li, E. Björnson, T. Svensson, T. Eriksson, and M. Debbah, "Joint precoding and load balancing optimization for energy-efficient heterogeneous networks," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 10, pp. 5810–5822, 2015.
- [31] G. Geraci, H. S. Dhillon, J. G. Andrews, J. Yuan, and I. B. Collings, "Physical layer security in downlink multi-antenna cellular networks," *IEEE Trans. on Commun.*, vol. 62, no. 6, pp. 2006–2021, 2014.
- [32] C. Li, J. Zhang, M. Haenggi, and K. B. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. on Commun.*, vol. 63, no. 4, pp. 1419–1431, 2015.
- [33] H. H. Yang, G. Geraci, T. Q. Quek, and J. G. Andrews, "Cell-edge-aware precoding for downlink massive MIMO cellular networks," *arXiv preprint arXiv:1607.01896*, 2016.
- [34] Y. Dhungana and C. Tellambura, "Performance analysis of SDMA with inter-tier interference nulling in HetNets," *IEEE Trans. on Wireless Commun.*, 2017.
- [35] T. Bai and R. W. Heath, "Analyzing uplink SINR and rate in massive MIMO systems using stochastic geometry," *IEEE Trans. on Commun.*, vol. 64, no. 11, pp. 4592–4606, 2016.
- [36] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.
- [37] K. Truong and R. Heath, "Effects of channel aging in massive MIMO systems," *IEEE/KICS Journal of Communications and Networks, Special Issue on massive MIMO*, vol. 15, no. 4, pp. 338–351, Aug 2013.
- [38] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 10, pp. 5795–5809, Oct 2015.
- [39] "The 5G infrastructure public private partnership," 2015. [Online]. Available: <http://5g-ppp.eu/>
- [40] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. on Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, 2005.
- [41] A. Papazafeiropoulos and T. Ratnarajah, "Linear precoding for downlink massive MIMO with delayed CSIT and channel prediction," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, April 2014, pp. 809–914.
- [42] A. K. Papazafeiropoulos, H. Q. Ngo, M. Matthaiou, and T. Ratnarajah, "Uplink performance of conventional and massive MIMO cellular systems with delayed CSIT," in *IEEE 25th Annual International Symposium on*

Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014. IEEE, 2014, pp. 601–606.

- [43] C. Kong *et al.*, “Sum-rate and power scaling of massive MIMO systems with channel aging,” *IEEE Trans. on Commun.*, vol. 63, no. 12, pp. 4879–4893, 2015.
- [44] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, “Effect of channel aging on the sum rate of uplink massive MIMO systems,” in *IEEE International Symposium on Information Theory (ISIT)*, 2015. IEEE, 2015, pp. 1222–1226.
- [45] A. Papazafeiropoulos, H. Ngo, and T. Ratnarajah, “Performance of massive mimo uplink with zero-forcing receivers under delayed channels,” *IEEE Transactions on Vehicular Technology*, 2016.
- [46] I. S. Gradshteyn and I. M. Ryzhik, “Table of integrals, series, and products,” *Alan Jeffrey and Daniel Zwillinger (eds.), Seventh edition (Feb 2007)*, vol. 885, 2007.
- [47] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.
- [48] M. Kountouris and J. G. Andrews, “Downlink SDMA with limited feedback in interference-limited wireless networks,” *IEEE Trans. on Wireless Commun.*, vol. 11, no. 8, pp. 2730–2741, 2012.
- [49] H. S. Dhillon, M. Kountouris, and J. G. Andrews, “Downlink MIMO HetNets: Modeling, ordering results and performance analysis,” *IEEE Trans. on Wireless Commun.*, vol. 12, no. 10, pp. 5208–5222, 2013.
- [50] X. Zhang *et al.*, “On the MIMO capacity with residual transceiver hardware impairments,” in *Proc. IEEE Int. Conf. Commun.*, 2014, pp. 5299–5305.
- [51] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, “Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits,” *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov 2014.
- [52] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, Wiley, Ed., 2011.
- [53] S. N. Chiu *et al.*, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.
- [54] M. Vu and A. Paulraj, “On the capacity of MIMO wireless channels with dynamic CSIT,” *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1269–1283, September 2007.
- [55] W. C. Jakes, *Microwave mobile communications*. New York: Wiley, 1974.
- [56] H. Huang, C. B. Papadias, and S. Venkatesan, *MIMO Communication for Cellular Networks*. Springer Science & Business Media, 2011.
- [57] N. Jindal, “MIMO broadcast channels with finite-rate feedback,” *IEEE Trans. on Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.



Tharmalingam Ratnarajah [A’96-M’05-SM’05] is currently with the Institute for Digital Communications, University of Edinburgh, Edinburgh, UK, as a Professor in Digital Communications and Signal Processing and the Head of Institute for Digital Communications. His research interests include signal processing and information theoretic aspects of 5G and beyond wireless networks, full-duplex radio, mmWave communications, random matrices theory, interference alignment, statistical and array signal processing and quantum information theory. He has published over 300 publications in these areas and holds four U.S. patents. He was the coordinator of the FP7 projects ADEL (3.7M€) in the area of licensed shared access for 5G wireless networks and HARP (3.2M€) in the area of highly distributed MIMO and FP7 Future and Emerging Technologies projects HIATUS (2.7M€) in the area of interference alignment and CROWN (2.3M€) in the area of cognitive radio networks. Dr Ratnarajah is a Fellow of Higher Education Academy (FHEA), U.K..



Anastasios Papazafeiropoulos [S’06-M’10] is currently a Research Fellow in IDCOM at the University of Edinburgh, U.K. He obtained the B.Sc in Physics and the M.Sc. in Electronics and Computers science both with distinction from the University of Patras, Greece in 2003 and 2005, respectively. He then received the Ph.D. degree from the same university in 2010. From November 2011 through December 2012 he was with the Institute for Digital Communications (IDCOM) at the University of Edinburgh, U.K. working as a postdoctoral Research Fellow, while

during 2012–2014 he was a Marie Curie Fellow at Imperial College London, U.K. Dr. Papazafeiropoulos has been involved in several EPSRC and EU FP7 HIATUS and HARP projects. His research interests span massive MIMO, 5G wireless networks, full-duplex radio, mmWave communications, random matrices theory, signal processing for wireless communications, hardware-constrained communications, and performance analysis of fading channels.