Deterministic Equivalent Performance Analysis of Time-Varying Massive MIMO Systems

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

Deterministic Equivalent Performance Analysis of Time-Varying Massive MIMO Systems

Anastasios K. Papazafeiropoulos, *Member, IEEE*, and Tharmalingam Ratnarajah, *Senior Member, IEEE*

***Abstract*—Delayed channel state information at the transmitter (CSIT) due to time variation of the channel, coming from the users' relative movement with regard to the BS antennas, is an inevitable degrading performance factor in practical systems. Despite its importance, little attention has been paid to the literature of multi-cellular multiple-input massive multiple-output (MIMO) system by investigating only the maximal ratio combining (MRC) receiver and the maximum ratio transmission (MRT) precoder. Hence, the contribution of this work is designated by the performance analysis/comparison of/with more sophisticated linear techniques, i.e., a minimum-mean-square-error (MMSE) detector for the uplink and a regularized zero-forcing (RZF) precoder for the downlink are assessed. In particular, we derive the deterministic equivalents of the signal-to-interference-plus-noise ratios (SINRs), which capture the effect of delayed CSIT, and make the use of lengthy Monte Carlo simulations unnecessary. Furthermore, prediction of the current CSIT after applying a Wiener filter allows to evaluate the mitigation capabilities of MMSE and RZF. Numerical results depict that the proposed achievable SINRs (MMSE/RZF) are more efficient than simpler solutions (MRC/MRT) in delayed CSIT conditions, and yield a higher prediction at no special computational cost due to their deterministic nature. Nevertheless, it is shown that massive MIMO are preferable even in time-varying channel conditions.**

***Index Terms*—Massive MIMO, delayed CSIT, channel estimation, channel prediction, linear precoding, linear detection.**

I. Introduction

Multi-user MIMO (MU-MIMO), applied to next generation systems (e.g., 802.16m [3], LTE-Advanced [4]), is one of the core technologies promising to provide a remarkable increase in data rates. Such systems include several co-channel users communicating with a base station (BS) equipped with multiple antennas. However, the technological transition to 5G systems is expected to demand a thousand-fold higher capacity.

Massive MIMO, where the BS includes a very large number of antennas, have emerged as one of the most promising technologies towards this direction because more degrees of freedom and increased power efficiency are achieved by simplifying multi-user processing, reducing transmit power, as well as vanishing the effects of thermal noise and fast fading [5]–[16]. Along these lines, given a multi-cellular scenario, linear detectors and precoders behave nearly optimal as the number of BS antennas goes to infinity, taking into account that channel vectors tend to be orthogonal when the number of antennas is large [7]. Specifically, the author derived the asymptotic signal-to-interference-plus-noise ratios (SINRs) and inferred that the effects of fast fading, uncorrelated noise, and intra-cell interference tend to disappear for unlimited number of BS antennas, even with simple maximum-ratio combining (MRC) in the uplink and maximum-ratio transmission (MRT) in the downlink. Similarly, derivation of lower bounds in the uplink revealed that MRC and zero forcing (ZF) perform the same in the low spectral efficiency regime, while ZF outperforms at higher spectral efficiencies [9]. Moreover, application of tools from random matrix theory (RMT) in [10] led to deterministic approximations (deterministic equivalents) of the SINR for the uplink with MRC and minimum mean-square error (MMSE) receivers as well as for the downlink with beamforming (BF) or else maximal ratio transmission (MRT) and regularized zero-forcing (RZF) precoders under the assumption that the numbers of BS antennas and users go to infinity at the same rate. Interestingly, it was shown that even in the case of a moderate number of BS antennas and users, the deterministic approximation of the SINR is tight. Nevertheless, the computational requirements of conjugate and zero-forcing linear precoders have been analyzed in [13], while a cooperative transmission strategy among BSs (multi-cell) with large number of antennas has been considered in [16], in order to study the asymptotic SINR.

Notably, except of grounding and well documenting the theory of massive MIMO, system simulations and channel measurements have validated their promising benefits [17], [18]. However, the assessment of this technique has not been conducted in depth. Specifically, the exploration of its behavior by including channel imperfections is imperative, in order to understand its real limits, given that the acquisition of perfect and timely CSI is crucial. Such imperfections include pilot contamination [6]–[13], [19], [20], transceiver hardware impairments [21], [22], and phase noise drift [23]. For example, the error induced in the channel due to the pilot contamination effect constitutes a bottleneck, since it saturates the performance with the number of antennas. Being an inherent weakness of multi-cell systems, caused by the reuse of pilot sequences in adjacent cells, it has been studied thoroughly.

Additionally to the aforementioned imperfections, delayed

Anastasios K. Papazafeiropoulos is with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, United Kingdom (e-mail:a.papazafeiropoulos@imperial.ac.uk). Tharmalingam Ratnarajah is with the Institute for Digital Communications (IDCoM), University of Edinburgh, EH9 3JL Edinburgh, UK (email: T.Ratnarajah@ed.ac.uk).

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

Parts of this work were presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Florence, Italy, May 2014 [1], and at the IEEE Wireless Communications and Networking Conference (WCNC), Istanbul, Turkey, March 2014 [2].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

CSIT appears due to the time-varying nature of real channels. In fact, the channel varies between when it is learned via estimation and when it is used for precoding or detection due to the relative movement between the antennas and the scatterers. Although it is main cause for the inevitably imperfect channel state information CSIT that should be taken seriously into account, an effort for characterizing the impact of delayed CSIT in massive MIMO has been addressed only in [24]–[26]. Especially, in [24], an application of the deterministic equivalent analysis was presented by considering linear techniques in the uplink and downlink in terms of maximum ratio combining (MRC) detector and maximum ratio transmission (MRT) precoder, respectively. Moreover, in [25], the optimal linear receiver in the case of cellular massive MIMO has been derived by exploiting the correlation between the channel estimates and the interference from other cells, while in [26], the uplink analysis of a cellular network with zero-forcing (ZF) receivers that holds for any finite as well as infinite number of BS antennas has been provided.

Given that further study of massive MIMO systems under such CSI assumptions is of paramount practical and theoretical interest, the focus of this paper is to cover the arising need for conducting realistic characterization and comparison of the corresponding ergodic achievable rates accounting for user mobility, when other types of receivers/precoders are considered. Note that this work extends substantially our conference papers [1], [2] by elaborating more on the effect of channel aging, providing an overall and thorough presentation of the analyses of both the uplink and downlink channels, and presenting the lengthy proofs, which have theoretical and practical interest regarding the large RMT. Hence, in the spirit of deterministic equivalents, developed in [10] for cellular networks, which employ the large MIMO concept, i.e., the number of users and BS antennas become larger at the same speed, we provide a deeper study in both the uplink and the downlink by considering a MMSE receiver and a RZF precoder under the assumption of imperfect CSI at the BS due to its delayed knowledge and pilot contamination as in [24]. It should be pointed out that our analysis is not just a straightforward extension of [24], which unfortunately brings too many typos, but it encompasses much more complex tools to deal with the more mathematically complex expressions of MMSE and RZF. Nevertheless, not only do we provide more details with comparison to [10], [24] in order to shed light into the mathematical analysis, but we also use a mixture of both approaches from the very start of defining the models to the end for providing the most advantageous analysis in terms of mathematical convenience and engineering insight. Note that the approximations (deterministic equivalents) are accurate even for realistic system dimensions, as simulations show. Thus, our analysis circumvents any need for lengthy Monte Carlo simulations. Specifically, the paper makes the following main contributions:

- We carry out an asymptotic performance analysis by deriving deterministic equivalent (asymptotically tight approximations) sum-rates accounting for the aforementioned channel impairments. In particular, we employ a MMSE receiver for the uplink and a RZF precoder for the downlink. Especially, the results quantify the loss in the performance due to delayed CSIT for different Doppler shifts, and illustrate the outperformance of MMSE and RZF against MRC and MRT, respectively. Nevertheless, it can be concluded that large MIMO are preferable even in time-varying channel conditions.
- We exploit the time correlation of the channel, in order to mitigate the detrimental effect of delayed CSI by predicting the current channel state from past measurements. Basically, capitalizing on the results in [24], we apply the finite impulse response (FIR) Wiener predictor under the concept of asymptotic equivalent analysis, in order to derive the corresponding realistic deterministic sum-rates for both the uplink (MMSE) and the downlink (RZF) channels. The results enable us to study explicitly how much more efficiently the predicted CSIT can overcome the inherent degradation due to delayed CSIT, if more sophisticated techniques than MRC and MRT are used.

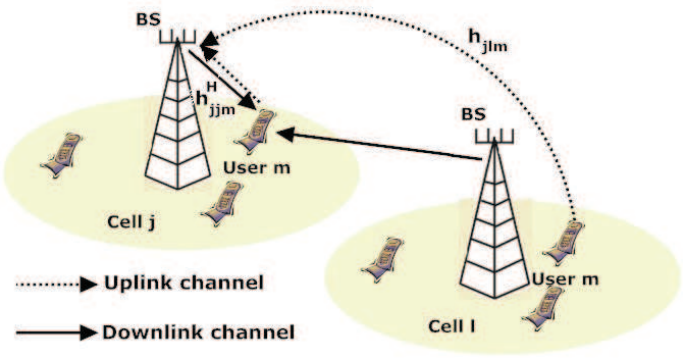

Fig. 1. Massive MIMO system with N BS antennas and K user terminals. The solid and dotted arrows represent the downlink and uplink channels, respectively. Due to TDD, channel reciprocity is considered. $\mathbf{h}^{\mathsf{H}}_{jjm}$ and $\mathbf{h}_{jlm}$ denote the desired downlink column channel vector in cell j and the interference uplink channel vector from cell l, respectively.

Notation: Vectors and matrices are represented by boldface lower and upper case symbols. $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $\mathrm{tr}\,(\cdot)$ denote the transpose, Hermitian transpose, and trace operators, respectively. The expectation operation and the spectral norm of a matrix are denoted by $\mathbb{E}\{\cdot\}$ and $\|\cdot\|$, respectively. The symbol $\otimes$ stands for the Kronecker product. Finally, $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes a circularly symmetric complex Gaussian vector $\mathbf{b}$ with zero-mean and covariance matrix $\mathbf{\Sigma}$.

II. System Model

A wireless cellular network is considered with L cells, as illustrated in Fig. 1. Following the common cellular architecture, we assume that each cell has one BS equipped with N antennas and K distributed active users with a single antenna each, served simultaneously. Note that both N, K increase by keeping their ratio fixed, in order to preserve the characteristic property of large MIMO. In other words, the number of antennas of each BS grows at the same speed with the number of the user terminals in every cell. Perfect synchronization occurs in time and frequency by all nodes (BSs and users), which operate in a time division duplex (TDD) protocol. Staying

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

complied with 3GPP LTE/LTE-Advanced standards [27], a universal frequency reuse of one is assumed, although other frequency reuse factors can also be implemented.

We consider flat-fading channels, where transmissions take place over a single frequency or subcarrier; Extensions to frequency selective channels follow similarly. In addition, a quasi-static fading model is assumed with coherence bandwidth much larger than the channel bandwidth, where the channel coefficients vary from symbol to symbol, but they are constant within one symbol.

The channel vector $\mathbf{h}_{jlm}[n] \in \mathbb{C}^N$ from user m in cell l to BS j at the n-th time slot is modelled as

$$\mathbf{h}_{jlm}[n] = \mathbf{R}_{jlm}^{1/2}\mathbf{q}_{jlm}[n], \tag{1}$$

where $\mathbf{R}_{jlm} \in \mathbb{C}^{N\times N}$ is a deterministic Hermitian-symmetric positive definite matrix and $\mathbf{q}_{jlm}[n] \in \mathbb{C}^N$ is an uncorrelated fast fading Gaussian channel vector with elements having zero mean and unit variance, i.e. $\mathbf{q}_{jlm}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. Note the independence of the deterministic matrix $\mathbf{R}_{jlm} = \mathbb{E}\left[\mathbf{h}_{jlm}[n]\mathbf{h}_{jlm}^{\mathsf{H}}[n]\right]$ of symbol index n because, in typical systems, their parameters change on a much larger time scale compared with the coherence interval and stay constant during several OFDM symbols [10], [12], [24]. Interestingly, many effects such as path loss dependent on the distance of the users from the BS, spatial correlation due to lack of limited antenna spacing, and different antenna patterns can be characterized by means of $\mathbf{R}_{jlm}$. Hereafter, the indices r and f denote the uplink (reverse) and the downlink (forward) channels, respectively.

A. Uplink Transmission

During the uplink phase, a simultaneous transmission of data occurs by the user terminals to their corresponding BS. The baseband signal $\mathbf{y}_{\mathrm{r},j} \in \mathbb{C}^N$, received at BS j at time instant n, is given as

$$\mathbf{y}_{\mathrm{r},j}[n] = \sqrt{p_{\mathrm{r}}}\sum_{l=1}^{L}\mathbf{H}_{jl}[n]\mathbf{x}_{\mathrm{r},l}[n] + \mathbf{z}_{\mathrm{r},j}[n], \tag{2}$$

where $p_{\mathrm{r}} > 0$ denotes the average transmit signal-to-noise ratio (SNR), $\mathbf{H}_{jl}[n] = \left[\mathbf{h}_{jl1}[n], \mathbf{h}_{jl2}[n], \ldots, \mathbf{h}_{jlK}[n]\right] \in \mathbb{C}^{N\times K}$ is the channel matrix from the users in cell l to BS j, $\mathbf{x}_{\mathrm{r},l}[n] = [x_{\mathrm{r},l1}[n], \ldots, x_{\mathrm{r},lK}[n]]^{\intercal} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ is the transmit symbol vector from the users in cell l, and $\mathbf{z}_{\mathrm{r},j}[n] \sim \mathcal{CN}\left(\mathbf{0}, \sigma_j^2\mathbf{I}_N\right)$ is the noise vector. It is worthwhile to mention that the transmit symbols $x_{\mathrm{r},lm}[n]$ are mutually independent.

Taking into consideration the application of linear detectors, the symbol $x_{\mathrm{r},lm}$ of user terminal m is obtained by the jth BS after applying a linear filter $\mathbf{W}_j[n] \in \mathbb{C}^{N\times K}$ to $\mathbf{y}_{\mathrm{r},j}[n]$. In particular, the received signal after the detector is (scaled by $1/\sqrt{p_{\mathrm{r}}}$)

$$\tilde{y}_{\mathrm{r},jm}[n] = \underbrace{\mathbf{w}_{jm}^{\mathsf{H}}[n]\mathbf{h}_{jjm}[n]x_{\mathrm{r},jm}[n]}_{\text{desired signal}} + \underbrace{\frac{1}{\sqrt{p_{\mathrm{r}}}}\tilde{z}_{\mathrm{r},jm}[n]}_{\text{noise}} + \underbrace{\sum_{(l,k)\neq(j,m)}\mathbf{w}_{jm}^{\mathsf{H}}[n]\mathbf{h}_{jlk}[n]x_{\mathrm{r},lk}[n]}_{\text{interference}}, \tag{3}$$

where $\tilde{z}_{\mathrm{r},jm} = \mathbf{w}_{\mathrm{r},jm}^{\mathsf{H}}[n]\mathbf{z}_{\mathrm{r},jm}[n]$ is spatially filtered Gaussian noise and $\mathbf{w}_{\mathrm{r},jm}[n]$ is the mth column of $\mathbf{W}_j[n]$.

B. Downlink Transmission

The simultaneous transmission of data from the base stations forms an interfering broadcast channel, which makes necessary the use of a precoding matrix. Due to use of TDD, the downlink channel is related with the uplink channel by applying the Hermitian transpose operation. Hence, the received signal $y_{\mathrm{f},jm} \in \mathbb{C}$ by the mth user in the jth cell is

$$y_{\mathrm{f},jm}[n] = \underbrace{\sqrt{p_{\mathrm{f}}}\sqrt{\lambda_j}\mathbf{h}_{jjm}^{\mathsf{H}}[n]\mathbf{f}_{jm}[n]x_{\mathrm{f},jm}[n]}_{\text{desired signal}} + \underbrace{z_{\mathrm{f},jm}[n]}_{\text{noise}} + \underbrace{\sum_{(l,k)\neq(j,m)}\sqrt{p_{\mathrm{f}}}\sqrt{\lambda_l}\mathbf{h}_{ljm}^{\mathsf{H}}[n]\mathbf{f}_{lk}[n]x_{\mathrm{f},lk}[n]}_{\text{interference}}, \tag{4}$$

where we have assumed that all BSs send the same transmit power p_{f}, $\mathbf{s}_l[n] = \sqrt{\lambda_j}\mathbf{F}_j[n]\mathbf{x}_{\mathrm{f},j}[n]$ constitutes the transmit signal vector by the jth BS and $z_{\mathrm{f},jm}[n] \sim \mathcal{CN}(0, \sigma_{jm}^2)$ is complex Gaussian noise at user m. Note that $\mathbf{F}_j[n] \in \mathbb{C}^{N\times K}$ is the linear precoding matrix that BS j uses, in order to transmit the data symbol vector $\mathbf{x}_{\mathrm{f},j}[n] = \left[x_{\mathrm{f},j1}[n],\ x_{\mathrm{f},j2}[n], \ldots,\ x_{\mathrm{f},jK}[n]\right]^{\intercal} \in \mathbb{C}^K \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ to its K serving users. The normalization parameter λ_l constrains the average transmit power per user of BS l to p_{f}, i.e., $\mathbb{E}\left[\frac{p_{\mathrm{f}}}{K}\mathbf{s}_l^{\mathsf{H}}[n]\mathbf{s}_l[n]\right]$, which gives

$$\lambda_l = \frac{1}{\mathbb{E}\left[\frac{1}{K}\operatorname{tr}\mathbf{F}_l[n]\mathbf{F}_l^{\mathsf{H}}[n]\right]} \tag{5}$$

III. Channel Impairments

The promised theoretical gains of multi-user channels, concerning power and multiplexing gains, require the knowledge of perfect current CSIT. Unfortunately, certain phenomena appear in realistic cases that degrade the quality of the CSIT. In this section, we present certain basic impairments that limit the performance by starting with the pilot contamination due to the interference from adjacent cells [7]. The description of this effect leads to inaccurate CSIT which can be rendered known by means of minimum mean square error (MMSE) estimation. Next and most importantly, we explain how delayed CSIT arises as a main detrimental effect. In fact, we model the delayed CSIT and we achieve to predict the current state by exploiting the time correlation of the channel by means of a simple linear channel predictor.

A. Pilot Contamination

Pilot Contamination, appearing in practical systems, is observed during channel estimation, which takes place at the training phase. In this phase, the user terminals send sequences (pilots) that are going to allow the BS j estimate its local channel $\hat{\mathbf{H}}_{jj}$ [10]. We assume that the length of the training period is τ and that all cells share the same set of orthogonal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

pilots $\mathbf{\Psi} = [\boldsymbol{\psi}_1; \ldots; \boldsymbol{\psi}_K] \in \mathbb{C}^{K\times\tau}$ with $\mathbf{\Psi}$ normalized, i.e., $\mathbf{\Psi}\mathbf{\Psi}^{\mathsf{H}} = \mathbf{I}_K$. The received signal reads as

$$\mathbf{Y}_{\mathrm{p},j}[n] = \sqrt{p_{\mathrm{p}}\tau}\Big(\sum_{l=1}^{L}\mathbf{H}_{jl}[n]\Big)\mathbf{\Psi} + \mathbf{Z}_{\mathrm{p},j}[n], \tag{6}$$

where p_{p} is the common average transmit power for all users and $\mathbf{Z}_{\mathrm{p},j}[n] \in \mathbb{C}^{N\times\tau}$ is spatially white additive Gaussian noise matrix at BS j during this phase. Correlation of the received signal with the training sequence $\frac{1}{\sqrt{p_{\mathrm{p}}\tau}}\boldsymbol{\psi}_m^{\mathsf{H}}$ of user m provides at BS j

$$\tilde{\mathbf{y}}_{\mathrm{p},jm}[n] = \underbrace{\mathbf{h}_{jjm}[n]}_{\text{desired}} + \underbrace{\sum_{l\neq j}\mathbf{h}_{jlm}[n]}_{\text{interference}} + \underbrace{\frac{1}{\sqrt{p_{\mathrm{p}}\tau}}\mathbf{Z}_{\mathrm{p},j}[n]\boldsymbol{\psi}_m^{\mathsf{H}}}_{\text{noise}}, \tag{7}$$

where $\tilde{\mathbf{z}}_{\mathrm{p},j}[n] = \mathbf{Z}_{\mathrm{p},j}[n]\boldsymbol{\psi}_m^{\mathsf{H}} \sim \mathcal{CN}(\mathbf{0}, \sigma_b^2\mathbf{I}_N)$. The interference term coming from users using the same sequence but belonging to other cells is known as pilot contamination, while the last term corresponds to noise contamination. The effect of pilot contamination, leading to a noisy observation of the received signal, makes necessary the use of MMSE estimation, in order to obtain the estimated channel $\hat{\mathbf{h}}_{jjm}[n]$, distributed as $\hat{\mathbf{h}}_{jjm}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_{jjm})$, according to [28]

$$\begin{aligned}\hat{\mathbf{h}}_{jjm}[n] &= \mathbf{R}_{jjm}\mathbf{Q}_{jm}\tilde{\mathbf{y}}_{\mathrm{p},jm}[n] \\ &= \mathbf{R}_{jjm}\mathbf{Q}_{jm}\Big(\sum_{l}\mathbf{h}_{jlm}[n] + \frac{1}{\sqrt{p_{\mathrm{p}}\tau}}\tilde{\mathbf{z}}_{\mathrm{p},b}[n]\Big),\end{aligned} \tag{8}$$

where

$$\mathbf{Q}_{jm} = \left(\frac{\sigma_j^2}{p_{\mathrm{p}}\tau}\mathbf{I}_N + \sum_{l}\mathbf{R}_{jlm}\right)^{-1}, \text{ and } \mathbf{\Phi}_{jlm} = \mathbf{R}_{jjm}\mathbf{Q}_{jm}\mathbf{R}_{jlm}.$$

It has to be mentioned that $\mathbf{R}_{jlm}$, $\mathbf{Q}_{jm}$, and $\mathbf{\Phi}_{jlm}$ are independent of n $\forall j$, l, and m. The orthogonality property of MMSE estimation allows decomposing the current CSIT as

$$\mathbf{h}_{jjm}[n] = \hat{\mathbf{h}}_{jjm}[n] + \tilde{\mathbf{h}}_{jjm}[n], \tag{9}$$

where $\tilde{\mathbf{h}}_{jjm}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{jjm} - \mathbf{\Phi}_{jjm})$ is the channel estimation error. Note that $\hat{\mathbf{h}}_{jjm}[n]$ and $\tilde{\mathbf{h}}_{jjm}[n]$ are statistically independent because they are uncorrelated and jointly Gaussian.

B. Delayed CSIT

The transmitter (BS) obtains CSI indirectly by assuming channel reciprocity. However, this reverse-channel information requires the forward and reverse links to occur at the same time. Unfortunately, a delay, being inherent in real channels, appears between the time the channel is learned via estimation and when it is applied for precoding or detection due to its time variation coming from the relative movement between the antennas and the scatterers.

A new model, capable of modelling the time variation of the channel, is presented, in order to study the effects of delayed CSIT on very large MIMO channels. Without loss of generality, we assume that all users move with the same velocity. As a result, the time variation does not depend on the user index. While this seems not realistic, we stay very near to the practical case by considering the worst-case scenario where we set all users with the velocity corresponding to the most varying user.

It is known that the variation of the channel is described by means of the second order statistics of the channel. An appropriate measure is the autocorrelation function of the channel, which is dependent on many parameters such as the characteristics of the propagation medium and the velocity of the user [29]. Although many models, describing the autocorrelation function, could be used, we choose the Jakes model, which is widely accepted due to its generality and simplicity. It refers to a propagation medium with two-dimensional isotropic scattering and a monopole antenna at the receiver [30]. According to this model, the normalized discrete-time autocorrelation function of the fading channel is given by

$$r_h[i] = J_0(2\pi f_D T_s|i|), \tag{10}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, T_s is the channel sampling duration, $|i|$ is the delay in terms of the number of symbols, and $f_D = \frac{vf_c}{c}$ is the maximum Doppler shift (v is the velocity of the user in $^m/_s$, $c = 3\times10^8$ $^m/_s$ is the speed of light, and f_c is the carrier frequency). Increase of the argument of the Bessel function, i.e., higher velocity of the user or more delay, results to decrease of the magnitude to zero but with some ripples in the meanwhile. In other words, as the delay $|i|$ increases or the user moves faster, the autocorrelation $r_h[i]$ decreases in magnitude to zero though not monotonically, since there are some ripples.

In our analysis, an appropriate method for relating the past samples of the fading channel with its current state can be given by the application of an autoregressive model of order L, denoted as AR(L) [29]. Its expression is

$$h[n] = \sum_{\ell=1}^{L} a_\ell h[n-\ell] + z[n], \tag{11}$$

where $\{a_\ell\}_{\ell=1}^{L}$ are the AR coefficients and $z[n]$ is temporally uncorrelated complex white Gaussian noise process with zero mean and variance $\sigma^2_{z,(L)}$. Although the order L of the model improves the accuracy of the current channel, we assume that $L = 1$ to keep complexity in logical levels without sacrificing enough accuracy. Otherwise, the design of predictors would be prohibitive, since many parameters should be estimated. Thus, the current channel between the BS j and the mth user, belonging to the lth cell, is modelled as

$$\mathbf{h}_{jlm}[n] = \alpha\mathbf{h}_{jlm}[n-1] + \mathbf{e}_{jlm}[n], \tag{12}$$

where $\mathbf{h}_{jlm}[n-1]$ is the channel in the previous symbol duration and $\mathbf{e}_{jlm}[n] \in \mathbb{C}^N$ is an uncorrelated channel error due to the channel variation modelled as a stationary Gaussian random process with i.i.d. entries and distribution $\mathcal{CN}(\mathbf{0}, (1-\alpha^2)\mathbf{R}_{jlm})$ [31]. It is assumed that the BSs know perfectly $\alpha = r_h[1] = J_0(2\pi f_D T_s)$. Nevertheless, we have from (12) that

$$\mathbb{E}[\mathbf{h}_{jlm}[n]\mathbf{h}_{jlm}^{\mathsf{H}}[n-1]] = \alpha\mathbf{R}_{jlm}. \tag{13}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

Basically, we can incorporate the effect of channel estimation, presented previously, to the autoregressive model of (12). As a result, the jth BS can be able to estimate its channel with the mth user in the local cell $\hat{\mathbf{h}}_{jjm}[n+1]$ by encountering at the same time both pilot contamination and delayed CSIT. This gives the ability to the BS to design the decoder $\mathbf{W}_j[n+T]$ or the precoder $\mathbf{F}_j[n+T]$ at time instance $(n+T)$ for the uplink and downlink, respectively. At time $(n+1)$, i.e., when the delay lasts for one channel sample $(T=1)$, the proposed model takes the form

$$\begin{aligned}\mathbf{h}_{jjm}[n+1] &= \alpha\mathbf{h}_{jjm}[n]+\mathbf{e}_{jjm}[n+1]\\ &= \alpha\hat{\mathbf{h}}_{jjm}[n]+\alpha\tilde{\mathbf{h}}_{jjm}[n]+\mathbf{e}_{jjm}[n+1] \quad (14)\\ &= \alpha\hat{\mathbf{h}}_{jjm}[n]+\tilde{\mathbf{e}}_{jjm}[n+1], \quad (15)\end{aligned}$$

where $\tilde{\mathbf{e}}_{jjm}[n+1]\sim\mathcal{CN}(\mathbf{0},\mathbf{R}_{jjm}-\alpha^2\mathbf{\Phi}_{jjm})$ and $\hat{\mathbf{h}}_{jjm}[n]$ are mutually independent. Following this way, the description of the channel with regards to larger delays is obvious. Obviously, the combined error $\tilde{\mathbf{e}}_{jjm}[n+1]$ depends on both the pilot contamination and delayed CSIT effects, allowing the export of interesting outcomes during the following analysis.

C. Channel Prediction

The standard approach, confronting and circumventing channel uncertainties, is to apply a channel prediction method, which obtains the current channel state from delayed measurements[1]. In other words, our target is to predict the autoregressive random process, given in (15), that accounts for pilot contamination and delayed CSIT. Interference from other BSs can be treated as uncorrelated additive Gaussian noise with zero mean and constant variance during the training phase. However, these channels change over time as the user moves. Suppose that we apply the p-th order linear Wiener predictor $\mathbf{V}_{jjm}=[\mathbf{V}_{jjm,0}\ \mathbf{V}_{jjm,1}\ \dots\ \mathbf{V}_{jjm,p}]\in\mathbb{C}^{N\times N(p+1)}$. This predictor provides us the predicted channel

$$\bar{\mathbf{h}}_{jjm}[n+1]=\sum_{q=0}^{p}\mathbf{V}_{jjm,q}\tilde{\mathbf{y}}_{\mathrm{p},jm}[n-q]=\mathbf{V}_{jjm}\bar{\mathbf{y}}_{\mathrm{p},jm}[n], \quad (16)$$

where $\bar{\mathbf{y}}_{\mathrm{p},jm}[n]=\left[\tilde{\mathbf{y}}^{\mathsf{H}}_{\mathrm{p},jm}[n]\ \tilde{\mathbf{y}}^{\mathsf{H}}_{\mathrm{p},jm}[n-1]\ \dots\ \tilde{\mathbf{y}}^{\mathsf{H}}_{\mathrm{p},jm}[n-p]\right]^{\mathsf{H}}\in\mathbb{C}^{N(p+1)\times 1}$.

The p-th order linear Wiener predictor is obtained in [24] as

$$\mathbf{V}_{jjm}=\alpha[\delta(p,\alpha)\otimes\mathbf{R}_{jjm}]\mathbf{T}_{jm}(p,\alpha), \quad (17)$$

Certain definitions have to be given regarding (17). Specifically, we have

$$\delta(p,\alpha)=[1\ \alpha\ \dots\ \alpha^p], \quad (18)$$

$$\Delta(p,\alpha)=\begin{pmatrix}1 & \alpha & \cdots & \alpha^p\\ \alpha & 1 & \cdots & \alpha^{p-1}\\ \vdots & \vdots & \ddots & \vdots\\ \alpha^p & \alpha^{p-1} & \cdots & 1\end{pmatrix}, \quad (19)$$

$$\mathbf{T}_{jm}(p,\alpha)=\left[\Delta(p,\alpha)\otimes\sum_{l=1}^{L}\mathbf{R}_{jlm}+\frac{\sigma_b^2}{p_{\mathrm{p}}\tau}\mathbf{I}_{N(p+1)}\right]^{-1}. \quad (20)$$

[1]It is worthwhile to mention that channel prediction techniques are used for downlink frequency division duplex (FDD) massive MIMO as well [32]–[34].

Note that this predictor is obtained by means of the principle of orthogonality [35, Eq. 3.41]

$$\mathbb{E}\Big[\big(\mathbf{h}_{jjm}[n+1]-\mathbf{V}_{jjm}\bar{\mathbf{y}}_{\mathrm{p},jm}[n]\big)\bar{\mathbf{y}}^{\mathsf{H}}_{\mathrm{p},jm}[n]\Big]=\mathbf{0}. \quad (21)$$

The predicted channel $\bar{\mathbf{h}}_{jjm}[n+1]$ appears to have a covariance matrix equal to $\alpha^2\mathbf{\Theta}_{jjm}(p,\alpha)$, where

$$\mathbf{\Theta}_{jlm}(p,\alpha)=[\delta(p,\alpha)\otimes\mathbf{R}_{jjm}]\,\mathbf{T}_{jm}(p,\alpha)^{\mathsf{H}}\,[\delta(p,\alpha)\otimes\mathbf{R}_{jlm}]^{\mathsf{H}}.$$

Thus, the current channel can be derived from its predicted version as

$$\mathbf{h}_{jjm}[n+1]=\bar{\mathbf{h}}_{jjm}[n+1]+\breve{\mathbf{h}}_{jjm}[n+1], \quad (22)$$

where $\breve{\mathbf{h}}_{jjm}[n+1]$ is the uncorrelated channel prediction error vector with covariance matrix $\mathbf{R}_{jjm}-\alpha^2\mathbf{\Theta}_{jlm}(p,\alpha)$. Note that the *zero*th order filter provides $\bar{\mathbf{h}}_{jjm}[n+1]=\alpha\hat{\mathbf{h}}_{jjm}[n]$, i.e., no channel prediction as expected.

IV. Asymptotic Performance Analysis

This section provides the characterization of the uplink and downlink channels after taking into consideration the effects of pilot contamination and delayed CSIT, as well as application of the Wiener predictor. The theory of deterministic equivalents offers the derivation of asymptotic results as $K,N\to\infty$, while keeping a finite ratio K/N. In fact, simulations, performed in the next section, corroborate that the proposed results are of practical value, since they are tight even for moderate system dimensions. Specifically, we start with general expressions for the SINR regarding both the uplink and downlink. In the sequel, we achieve to present deterministic equivalent SINRs, where the MMSE decoder on the uplink and the RZF precoder on the downlink are applied. However, the proofs corresponding to the downlink are omitted due to limited space. The most interesting parts are i) the demonstration of the dependence of these expressions by delayed CSIT, which is inherent in real channels, and ii) the description of these expression after having applied a linear predictor, useful to mitigate any delayed CSIT effects.

We have to mention that the results rest to certain assumptions. Specifically, let the maximum eigenvalue of any correlation matrix and the inter-cell interference matrix including the channel estimation errors to be finite. Moreover, all spatial correlation matrices must not have zero energy, especially, during the simulations.

We consider that BS j has knowledge of CSI $\mathbf{g}_{jjm}[n+1]$, which differentiates among specific conditions. Thus, in case of knowledge of current CSI at time $n+1$, where the channel is not perfect but estimated due to pilot contamination, we have by means of (9) that

$$\mathbf{g}_{jjm}[n+1]=\hat{\mathbf{h}}_{jjm}[n+1]. \quad (23)$$

More concretely, the last expression has considered the knowledge of imperfect current CSI because of pilot contamination without any use of delayed CSI. When, the effect of delay is taken into account, the expression of CSI reads due to (15) as

$$\mathbf{g}_{jjm}[n+1]=\alpha\hat{\mathbf{h}}_{jjm}[n]. \quad (24)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

In a similar way, after applying channel prediction, the expression for CSI according to (22) becomes

$$\mathbf{g}_{jjm}[n+1] = \bar{\mathbf{h}}_{jjm}[n+1]. \quad (25)$$

A. Uplink Channel Analysis

During the uplink phase, the jth BS detects the transmit symbol $x_{\mathrm{r},jm}[n+1]$ of user m belonging to the same cell by applying a linear filter $\mathbf{w}_{jm}[n+1] \in \mathbb{C}^N$ to the received signal $y_{\mathrm{r},jm}[n+1]$, as shown in (3). The detected signal $\tilde{y}_{\mathrm{r},jm}[n+1]$ can be written in a different way as

$$\begin{aligned}\tilde{y}_{\mathrm{r},jm}[n+1] &= \mathbf{w}_{jm}^{\mathsf{H}}[n+1]\mathbf{g}_{jjm}[n+1]x_{\mathrm{r},jm}[n+1] \\ &+ \mathbf{w}_{jm}^{\mathsf{H}}[n+1](\mathbf{h}_{jjm}[n+1]-\mathbf{g}_{jjm}[n+1])x_{\mathrm{r},jm}[n+1] \\ &+ \sum_{(l,k)\neq(j,m)} \mathbf{w}_{jm}^{\mathsf{H}}[n+1]\mathbf{h}_{jlk}[n+1]x_{\mathrm{r},lk}[n+1] + \frac{\tilde{z}_{\mathrm{r},jm}[n+1]}{\sqrt{p_{\mathrm{r}}}}, \quad (26)\end{aligned}$$

where CSI takes different forms depending on the effect we want to study, as expressed by (23)–(25). In addition, the temporal parameter α and $\mathbf{R}_{jlm} \forall l$ are assumed known by BS j. Based on the approach [6], [10], $\tilde{y}_{\mathrm{r},jm}[n+1]$ can be described as the received signal of a single-input single-output (SISO) system with the effective channel to be $\mathbf{w}_{jm}^{\mathsf{H}}[n+1]\mathbf{g}_{jjm}[n+1]$, while any other term constitutes the uncorrelated additive Gaussian noise. Thus, the desired signal power is

$$S_{\mathrm{r},jm} = \left|\mathbf{w}_{jm}^{\mathsf{H}}[n+1]\mathbf{g}_{jjm}[n+1]\right|^2, \quad (27)$$

while the terms, concerning the interference and noise powers, are

$$\begin{aligned}I_{\mathrm{r},jm} &= \left|\mathbf{w}_{jm}^{\mathsf{H}}[n+1]\left(\mathbf{h}_{jjm}[n+1]-\mathbf{g}_{jjm}[n+1]\right)\right|^2 \\ &+ \frac{\sigma_j^2}{p_{\mathrm{r}}}\left|\mathbf{w}_{jm}^{\mathsf{H}}[n+1]\right|^2 + \sum_{(l,k)\neq(j,m)} \left|\mathbf{w}_{jm}^{\mathsf{H}}[n+1]\mathbf{h}_{jlk}[n+1]\right|^2. \quad (28)\end{aligned}$$

As a result, the uplink ergodic achievable rate of user m in cell j, based on a bound that considers the worst-case uncorrelated additive noise [36], is

$$R_{\mathrm{r},jm} = \mathbb{E}\left[\log_2(1+\gamma_{\mathrm{r},jm})\right], \quad (29)$$

where $\gamma_{\mathrm{r},jm} = \frac{S_{\mathrm{r},jm}}{I_{\mathrm{r},jm}}$ is the instantaneous uplink SINR, and the expectation is computed over all channel realizations.

The deterministic approximation of the uplink SINR is obtained such that $\gamma_{\mathrm{r},jm} - \bar{\gamma}_{\mathrm{r},jm} \xrightarrow[N\to\infty]{\text{a.s.}} 0$, where $\bar{\gamma}_{\mathrm{r},jm}$ is the deterministic equivalent SINR, and $\xrightarrow[N\to\infty]{\text{a.s.}}$ denotes almost sure convergence. The corresponding deterministic equivalent uplink sum-rate can be obtained by means of the dominated convergence [37] and the continuous mapping theorem [38] as

$$R_{\mathrm{r},jm} - \log_2(1+\bar{\gamma}_{\mathrm{r},jm}) \xrightarrow[N\to\infty]{\text{a.s.}} 0. \quad (30)$$

If delayed CSIT is considered, the estimated desired channel at time $n+1$ is $\mathbf{g}_{jjm}[n+1] = \alpha\hat{\mathbf{h}}_{jjm}[n]$ according to (24). The linear filter, taken into account, is the MMSE detector

$$\begin{aligned}\mathbf{w}_{jm}^{\mathrm{MMSE}}[n+1] &= \boldsymbol{\Sigma}_j^{(G)}[n+1]\mathbf{g}_{jjm}[n+1] \\ &= \alpha\hat{\boldsymbol{\Sigma}}_j^{(H)}[n]\hat{\mathbf{h}}_{jjm}[n], \quad (31)\end{aligned}$$

where we define the matrices

$$\boldsymbol{\Sigma}_j^{(G)}[n+1] = \left(\mathbf{G}_{jj}[n+1]\mathbf{G}_{jj}^{\mathsf{H}}[n+1] + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\right)^{-1} \quad (32)$$

$$\hat{\boldsymbol{\Sigma}}_j^{(H)}[n] = \left(\alpha^2\hat{\mathbf{H}}_{jj}[n]\hat{\mathbf{H}}_{jj}^{\mathsf{H}}[n] + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\right)^{-1}. \quad (33)$$

Hereafter, we set $\hat{\boldsymbol{\Sigma}}_j^{(H)}[n] = \hat{\boldsymbol{\Sigma}}_j[n]$ for simplicity. Regarding the parameters of $\hat{\boldsymbol{\Sigma}}_j[n]$, $\varphi_{\mathrm{r},j}$ is a positive scalar and $\mathbf{Z}_{\mathrm{r},j} \in \mathbb{C}^{N\times N}$ is a Hermitian nonnegative definite matrix, which can be optimized, and provide some new insights in case of presence of both delayed CSI and pilot contamination. Similar analysis has been presented in [39], where different regularization parameters for RZF precoding were derived under different scenarios with large number of antennas. Thus, it is an interesting topic of future work to explore and determine the corresponding parameters and shed light on their behavior.

If the BS has predicted CSI, we substitute $\mathbf{g}_{jjm}[n+1] = \bar{\mathbf{h}}_{jjm}[n+1]$, $\mathbf{w}_{jjm}^{\mathrm{MMSE}}[n+1] = \boldsymbol{\Sigma}_j^{(G)}[n+1]\mathbf{g}_{jjm}[n+1] = \bar{\boldsymbol{\Sigma}}_j[n+1]\bar{\mathbf{h}}_{jjm}[n+1]$, and $\mathbf{h}_{jjm}[n+1] - \bar{\mathbf{h}}_{jjm}[n+1] = \check{\mathbf{h}}_{jjm}[n+1]$ into (27) and (28), where $\bar{\boldsymbol{\Sigma}}_j[n+1] = \left(\bar{\mathbf{H}}_{jj}[n+1]\bar{\mathbf{H}}_{jj}^{\mathsf{H}}[n+1] + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\right)^{-1}$.

Theorems 1 and 2 present the deterministic equivalent uplink SINRs that consider delayed and predicted CSI, respectively, after applying the MMSE detector.

Theorem 1: The deterministic equivalent uplink SINR for user m in cell j with MMSE detector, accounting for delayed CSI, is given by (34) with $\hat{\delta}_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjm}\mathbf{T}_j$, $\hat{\delta}'_{\mathrm{r},jm} = \frac{1}{N}\operatorname{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Phi}_{jjm})\mathbf{T}_j^{\mathrm{r}_{d1}}$, $\hat{\delta}''_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjm}\mathbf{T}_j^{\mathrm{r}_{d2}}$, $\hat{\vartheta}_{jlk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j$, $\hat{\vartheta}'_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d1}}$, $\hat{\delta}'''_{jk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjk}\mathbf{T}_j^{\mathrm{r}_{d1}}$, and

$$\hat{\mu}_{jlkm} = \frac{\operatorname{tr}\mathbf{R}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d1}}}{N} - \frac{2\mathrm{Re}\{\hat{\vartheta}_{jlk}^*\hat{\vartheta}'_{jlkm}\}\left(1+\hat{\delta}_{jk}\right) - \hat{\vartheta}_{jlk}^2\hat{\delta}'''_{jk}}{\left(1+\hat{\delta}_{jk}\right)^2},$$

where

- $\mathbf{T}_j = \mathbf{T}(\phi_{\mathrm{r},j})$ and $\boldsymbol{\delta}_j = [\delta_{j1},\ldots,\delta_{jK}]^{\mathsf{T}} = \boldsymbol{\delta}(\phi_{\mathrm{r},j})$ are given by Theorem 5 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\,\forall k$,
- $\mathbf{T}_j^{\mathrm{r}_{d1}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{r}_{d1}} = \boldsymbol{\Phi}_{jjm}$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\,\forall k$,
- $\mathbf{T}_j^{\mathrm{r}_{d2}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{r}_{d1}} = \mathbf{I}_N$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\,\forall k$.

Proof: See Appendix B. ∎

Theorem 2: The deterministic equivalent uplink SINR for user m in cell j with MMSE detector, accounting for predicted CSI by means of a pth-order Wiener predictor, is given by (35) with $\bar{\delta}_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j$, $\bar{\delta}'_{\mathrm{r},jm} = \frac{1}{N}\operatorname{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$, $\bar{\delta}''_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p2}}$, $\bar{\vartheta}_{jlk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jlk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p2}}$, $\bar{\vartheta}'_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jlk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$, $\bar{\delta}'''_{jk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$, and

$$\bar{\mu}_{jlkm} = \frac{\operatorname{tr}\mathbf{R}_{jlk}\mathbf{T}_j^{\mathrm{r}_{p1}}}{N} - \frac{2\mathrm{Re}\{\alpha^4\bar{\vartheta}_{jlk}^*\bar{\vartheta}'_{jlkm}\}\left(1+\alpha^2\bar{\delta}_{jk}\right) - \alpha^6\bar{\vartheta}_{jlk}^2\bar{\delta}'''_{jk}}{\left(1+\alpha^2\bar{\delta}_{jk}\right)^2},$$

where

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

$$\bar{\gamma}_{\mathrm{r},jm}(\alpha) = \frac{\alpha^2 \hat{\delta}^2_{jm}}{\frac{1}{N}\hat{\delta}'_{\mathrm{r},jm} + \frac{\sigma^2_j}{p_\mathrm{r}}\frac{1}{N}\hat{\delta}''_{jm} + \sum\limits_{(l,k)\neq(j,m)} \frac{1}{N}\hat{\mu}_{jlkm} + \alpha^2 \sum\limits_{l\neq j} |\hat{\vartheta}_{jlm}|^2} \tag{34}$$

$$\bar{\gamma}_{\mathrm{r},jm}(\alpha) = \frac{\alpha^2 \bar{\delta}^2_{jm}}{\frac{1}{N}\bar{\delta}'_{\mathrm{r},jm} + \frac{\sigma^2_j}{p_\mathrm{r}}\frac{1}{N}\bar{\delta}''_{jm} + \sum\limits_{(l,k)\neq(j,m)} \frac{1}{N}\bar{\mu}_{jlkm} + \sum\limits_{l\neq j} \alpha^{2(p+2)} |\bar{\vartheta}_{jlm}|^2} \tag{35}$$

* $\mathbf{T}_j = \mathbf{T}(\phi_{\mathrm{r},j})$ and $\boldsymbol{\delta}_j = [\delta_{j1}, \ldots, \delta_{jK}]^\mathsf{T} = \boldsymbol{\delta}(\phi_{\mathrm{r},j})$ are given by Theorem 5 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\, \forall k$,
* $\mathbf{T}_j^{\mathrm{r}_{p1}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{r}_{p1}} = \boldsymbol{\Theta}_{jjm}(p,\alpha)$, $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\, \forall k$,
* $\mathbf{T}_j^{\mathrm{r}_{p2}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{r}_{p2}} = \mathbf{I}_N$, $\mathbf{R}_k = \boldsymbol{\Phi}_{jjk}\, \forall k$.

Proof: See Appendix C. ∎

B. Downlink Transmission

During the downlink phase, users do not have instantaneous CSI. For this reason, we use similar techniques to [40], in order to obtain an expression for the downlink sum-rate under the assumption that user m, found in cell j, has knowledge of only $\mathbb{E}[\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]]$. Specifically, the received signal at user m in cell j, scaled by $1/\sqrt{p_\mathrm{f}}$, is given after reforming (4) by

$$\begin{aligned} y_{\mathrm{f},jm}[n+1] =& \sqrt{\lambda_j}\mathbb{E}\left[\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]\right] x_{\mathrm{f},jm}[n+1] \\ &+ \sum_{(l,k)\neq(j,m)} \sqrt{\lambda_l}\mathbf{h}^\mathsf{H}_{ljm}[n+1]\mathbf{f}_{lk}[n+1]x_{\mathrm{f},lk}[n+1] + \frac{1}{p_\mathrm{f}} z_{\mathrm{f},jm}[n+1] \\ &+ \sqrt{\lambda_j}\left(\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]\right. \\ &\left. - \mathbb{E}\left[\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]\right]\right) x_{\mathrm{f},jm}[n+1]. \end{aligned}$$

If we assume uncorrelated additive noise, we may consider a SISO channel. Thus, the signal power at user m in cell j, divided by N^2, can be given by

$$S_{\mathrm{f},jm} = \frac{1}{N^2}\lambda_j \left|\mathbb{E}\left[\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]\right]\right|^2, \tag{36}$$

while the interference plus noise power at user m in cell j, scaled by $1/N^2$, is

$$\begin{aligned} I_{\mathrm{f},jm} =& \frac{1}{N^2}\lambda_j \mathrm{var}\left[\mathbf{h}^\mathsf{H}_{jjm}[n+1]\mathbf{f}_{jm}[n+1]\right] + \frac{1}{N^2}\frac{\sigma^2_{jm}}{p_\mathrm{f}} \\ &+ \sum_{(l,k)\neq(j,m)} \frac{1}{N^2}\lambda_l \mathbb{E}\left[\left|\mathbf{h}^\mathsf{H}_{ljm}[n+1]\mathbf{f}_{lk}[n+1]\right|^2\right]. \end{aligned} \tag{37}$$

Here, we focus on the use of RZF precoding $\mathbf{F}^{\mathrm{RZF}}_j[n+1]$ in cell j that considers delayed CSI. It is given by

$$\mathbf{F}^{\mathrm{RZF}}_j[n+1] = \hat{\boldsymbol{\Sigma}}^{(G)}_j[n+1]\hat{\mathbf{G}}_{jj}[n+1] \overset{(24)}{=} \alpha\hat{\boldsymbol{\Sigma}}_j[n]\hat{\mathbf{H}}_{jj}[n], \tag{38}$$

where we have defined $\hat{\boldsymbol{\Sigma}}^{(G)}_j[n+1]$ and $\hat{\boldsymbol{\Sigma}}_j[n]$ as in (32) and (33). The design parameters $\varphi_{\mathrm{f},j} > 0$ and $\mathbf{Z}_{\mathrm{f},j} \in \mathbb{C}^{N\times N}$, which is a Hermitian nonnegative definite matrix, correspond to the downlink case and can be optimized as in the uplink phase. In case of predicted CSI, the precoding vector is $\mathbf{f}^{(\mathrm{p})}_{jm} \overset{(25)}{=} \bar{\boldsymbol{\Sigma}}_j[n+1]\bar{\mathbf{h}}_{jjm}[n+1]$, where $\bar{\boldsymbol{\Sigma}}_j[n+1]$, can be defined similar to the uplink case, although the optimization parameters now refer to the downlink case.

The downlink deterministic equivalent SINRs at user m in cell j with RZF precoding, including delayed and predicted CSI considerations, are expressed by mean of Theorems 3 and 4, respectively.

Theorem 3: The downlink deterministic equivalent SINR at user m in cell j, applying a RZF precoder and accounting for delayed CSI, is given by (39) with $\hat{\delta}_{jm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{jjm}\mathbf{T}_j$, $\hat{\delta}'_{\mathrm{f},jm} = \frac{1}{N}\mathrm{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Phi}_{jjm})\mathbf{T}^{\mathrm{f}_{d2}}_j$, $\hat{\vartheta}_{ljm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}_l$, $\hat{\vartheta}'_{ljmk} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{ljm}\mathbf{T}^{\mathrm{f}_{d3}}_l$, $\hat{\delta}'''_{lm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Phi}_{llm}\mathbf{T}^{\mathrm{f}_{d3}}_l$, and

$$\hat{\lambda}^{\mathrm{d}}_j = \frac{1}{N^2}\frac{K}{\alpha^2\left(\frac{1}{N}\mathrm{tr}\,\mathbf{T}_j - \frac{1}{N}\mathrm{tr}\left(\frac{\mathbf{Z}^{\mathrm{dl}}_j}{N} + \varphi^{\mathrm{dl}}_j\mathbf{I}_N\right)\mathbf{T}^{f_{d1}}_j\right)}$$

$$\hat{\mu}_{ljmk} = \frac{\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}^{\mathrm{f}_{d3}}_l}{N} - \frac{2\mathrm{Re}\{\hat{\vartheta}^*_{ljm}\hat{\vartheta}'_{ljmk}\}\left(1+\hat{\delta}_{lm}\right) - \hat{\vartheta}^2_{ljm}\hat{\delta}'''_{lm}}{\left(1+\hat{\delta}_{lm}\right)^2},$$

where

* $\mathbf{T}_j = \mathbf{T}(\phi_{\mathrm{r},j})$ and $\boldsymbol{\delta}_j = [\delta_{j1}, \ldots, \delta_{jK}]^\mathsf{T} = \boldsymbol{\delta}(\phi_{\mathrm{r},j})$ are given by Theorem 5 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\, \forall k$,
* $\mathbf{T}_j^{\mathrm{f}_{d1}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{d1}} = \mathbf{I}_N$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\, \forall k$,
* $\mathbf{T}_j^{\mathrm{f}_{d2}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{d2}} = \boldsymbol{\Phi}_{jjm}$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\, \forall k$,
* $\mathbf{T}_j^{\mathrm{f}_{d3}} = \mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D} = \mathbf{I}_N$, $\mathbf{S} = \mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{d3}} = \boldsymbol{\Phi}_{llk}$, $\mathbf{R}_k = \alpha^2\boldsymbol{\Phi}_{jjk}\, \forall k$.

Theorem 4: The downlink deterministic equivalent SINR at user m in cell j with RZF precoding and predicted CSI by using a pth-order Wiener filter is given by (40) with $\bar{\delta}_{jm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j$, $\bar{\delta}'_{\mathrm{f},jm} = \mathrm{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}^{\mathrm{f}_{p2}}_j$, $\bar{\vartheta}_{ljm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Theta}_{ljm}(p,\alpha)\mathbf{T}_l$, $\bar{\vartheta}'_{ljmk} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Theta}_{ljm}(p,\alpha)\mathbf{T}^{\mathrm{f}_{p3}}_l$, $\bar{\delta}'''_{lm} = \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Theta}_{llm}(p,\alpha)\mathbf{T}^{\mathrm{f}_{p3}}_l$, and

$$\bar{\lambda}^{\mathrm{p}}_j = \frac{1}{N^2}\frac{K}{\left(\frac{1}{N}\mathrm{tr}\,\mathbf{T}_j - \mathrm{tr}\left(\frac{\mathbf{Z}^{\mathrm{dl}}_j}{N} + \varphi^{\mathrm{dl}}_j\mathbf{I}_N\right)\bar{\mathbf{T}}^{p_1}_j\right)}$$

$$\bar{\mu}_{ljmk} = \frac{\mathrm{tr}\,\mathbf{R}_{ljm}\mathbf{T}^{\mathrm{f}_{p3}}_l}{N} - \frac{2\mathrm{Re}\{\alpha^4\bar{\vartheta}^*_{ljm}\bar{\vartheta}'_{ljmk}\}\left(1+\alpha^2\bar{\delta}_{lm}\right) - \alpha^6\bar{\vartheta}^2_{ljm}\bar{\delta}'''_{lm}}{\left(1+\alpha^2\bar{\delta}_{lm}\right)^2},$$

where

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

$$\bar{\gamma}_{\mathrm{f},jm}(p,\alpha)=\frac{\hat{\lambda}_j^{\mathrm{d}}\alpha^4\hat{\delta}_{jm}^2}{\alpha^2\hat{\lambda}_j^{\mathrm{d}}\frac{1}{N}\hat{\delta}'_{\mathrm{r},jm}+\frac{\sigma_{jm}^2(1+\hat{\delta}_{jm})^2}{p_{\mathrm{f}}}+\sum\limits_{(l,k)\neq(j,m)}\alpha^2\hat{\lambda}_l^{\mathrm{d}}\left(\frac{1+\hat{\delta}_{jm}}{1+\hat{\delta}_{lk}}\right)^2\frac{1}{N}\hat{\mu}_{ljmk}+\sum\limits_{l\neq j}\alpha^4\hat{\lambda}_l^{\mathrm{d}}\left(\frac{1+\hat{\delta}_{jm}}{1+\hat{\delta}_{lm}}\right)^2|\hat{\vartheta}_{ljm}|^2} \tag{39}$$

$$\bar{\gamma}_{\mathrm{f},jm}(p,\alpha)=\frac{\alpha^4\bar{\lambda}_j^{\mathrm{p}}\bar{\delta}_{jm}^2}{\alpha^2\bar{\lambda}_j^{\mathrm{p}}\frac{1}{N}\bar{\delta}'_{\mathrm{f},jm}+\frac{\sigma_{jm}^2(1+\alpha^2\bar{\delta}_{jm})}{p_{\mathrm{f}}}+\sum\limits_{(l,k)\neq(j,m)}\alpha^2\bar{\lambda}_l^{\mathrm{p}}\left(\frac{1+\alpha^2\bar{\delta}_{jm}}{1+\alpha^2\bar{\delta}_{lk}}\right)^2\frac{1}{N}\bar{\mu}_{ljmk}+\sum\limits_{l\neq j}\alpha^{2(p+2)}\bar{\lambda}_l^{\mathrm{p}}\left(\frac{1+\alpha^2\bar{\delta}_{jm}}{1+\alpha^2\bar{\delta}_{lm}}\right)^2|\bar{\vartheta}_{ljm}|^2} \tag{40}$$

- $\mathbf{T}_j=\mathbf{T}(\phi_{\mathrm{r},j})$ and $\boldsymbol{\delta}_j=[\delta_{j1},\ldots,\delta_{jK}]^{\mathsf{T}}=\boldsymbol{\delta}(\phi_{\mathrm{r},j})$ are given by Theorem 5 for $\mathbf{D}=\mathbf{I}_N$, $\mathbf{S}=\mathbf{Z}_j/N$, $\mathbf{R}_k=\boldsymbol{\Phi}_{jjk}\ \forall k$,
- $\mathbf{T}_j^{\mathrm{f}_{p1}}=\mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D}=\mathbf{I}_N$, $\mathbf{S}=\mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{p1}}=\mathbf{I}_N$, $\mathbf{R}_k=\boldsymbol{\Phi}_{jjk}\ \forall k$,
- $\mathbf{T}_j^{\mathrm{f}_{p2}}=\mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D}=\mathbf{I}_N$, $\mathbf{S}=\mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{p2}}=\boldsymbol{\Theta}_{jjm}(p,\alpha)$, $\mathbf{R}_k=\boldsymbol{\Phi}_{jjk}\ \forall k$,
- $\mathbf{T}_j^{\mathrm{f}_{p3}}=\mathbf{T}'(\phi_{\mathrm{r},j})$ is given by Theorem 6 for $\mathbf{D}=\mathbf{I}_N$, $\mathbf{S}=\mathbf{Z}_j/N$, $\mathbf{K}^{\mathrm{f}_{p3}}=\boldsymbol{\Theta}_{llk}(p,\alpha)$, $\mathbf{R}_k=\boldsymbol{\Phi}_{jjk}\ \forall k$.

V. Discussion and Numerical Results

This section starts with comments concerning the proposed results. Interestingly, regarding the uplink, each term of the denominator of (34) and (35) discriminates among different effects due to pilot contamination. In particular, the terms from left to right indicate the channel estimation error, the post-processed local noise at BS j, the post-processed inter-cell and intra-cell interference from users sharing different pilot, and the post-processed inter-cell interference from users sharing the same pilot with user m. Especially, in the case of knowledge of current CSIT, i.e., $\alpha=1$ and $p=0$, our results coincide with [10, eq. 25]. From the physical point of view, the SINR with delayed CSIT cannot be greater than the SINR with current CSIT. This can also be shown intuitively, since increase of α brings an increase of the desired estimated signal power, while it results to a decrease of the estimation error term. Moreover, the move of user m affects not only the desired channel at BS j, but also the interference channels from users in other cells sharing the same pilot as well. Basically, the higher the velocity of the user (decreasing α), the greater the degradation of the channels, as expressed by the reduction of the uplink SINR.

As far as the downlink is concerned, expressions (39) and (40) reduce to Theorem 6 of [10], if $\alpha=1$ and $p=0$. The pilot contamination affects the downlink sum-rate in the same way as in the uplink, while delayed CSIT affects apart of the desired channel and the interference channel from BS l belonging to different cells ($l\neq j$), but also the term corresponding to the channel estimation error.

In addition, in this section, numerical results verify our analysis and demonstrate the performance of MMSE decoder and RZF precoder under the effects of delayed CSI and pilot contamination, as well as they reveal the benefits of the necessary channel prediction in massive MIMO cellular networks. Nevertheless, comparisons with MRC and MRT are provided that give an insight on the outperformance of MMSE and RZF. It is worthwhile to mention that the simulations testify that the proposed approximations are tight for practical system dimensions. Space limitations allow us to present selected representative results. However, we avoid any particular omission of presenting important results due to the similarities between the downlink and the uplink.

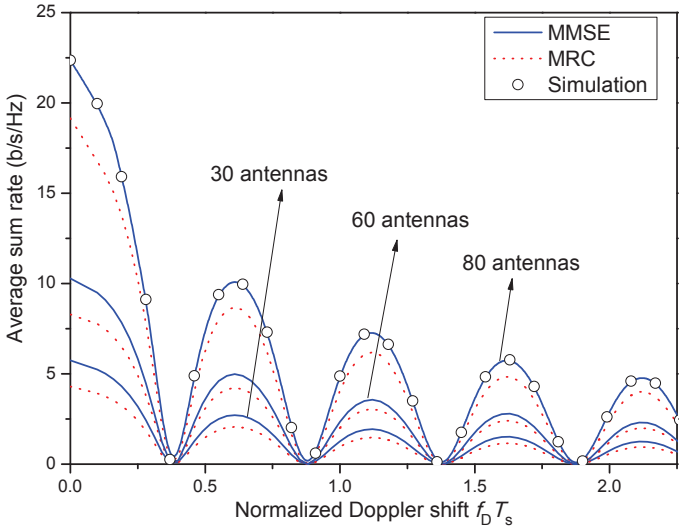

Fig. 2. Simulated and deterministic equivalent uplink sum-rates with MMSE and MRC decoders versus the normalized Doppler shift $f_{\mathrm{D}}T_s$ for different number of BS antennas N.

Our focus is on the simulation of uplink and downlink sum-rates in the center cell of a multi-cell hexagonal network with $L=5$ cells sharing the same frequency band. We assume that each cell has normalized radius to one, while $K=4$ users are uniformly distributed around each BS on a circle of radius 0.7. The physical model includes a distance-based path loss model with path loss parameter $\beta=3.5$ and without antenna correlation. In this case, $\tilde{\mathbf{R}}_{jlk}=d_{jlk}^{-\frac{\beta}{2}}\mathbf{I}_N$, where d_{jlk} denotes the distance between BS j and user k in cell j. Moreover, the transmit, receive, and training SNRs are $p_{\mathrm{f}}=8$dB, $p_{\mathrm{r}}=8$dB, and $p_{\mathrm{p}}=5$dB, respectively. The design parameters for MMSE, as well as RZF are $\phi_{\mathrm{r},j}=\frac{1}{p_{\mathrm{r}}}$ and $\mathbf{Z}_{\mathrm{r},j}=0$ as well as $\phi_{\mathrm{f},j}=\frac{1}{p_{\mathrm{f}}}$ and $\mathbf{Z}_{\mathrm{f},j}=0$, respectively. Monte-Carlo simulations enable us to obtain (15) and (22), and provide the simulation curves for the sum-rate with delayed and predicted CSIT, respectively. More specifically, the ratios of (27) by (28) and (36) by (37) provide the uplink and downlink SINRs by which we the simulated curves result. The analytical curves for MMSE and RZF as well as MRC and MRT are computed via (34), (35) and (39), (40) as well as [24, eq. 53, 75] and [24, eq. 100, 120], respectively. The simulated and analytical results match exactly in all cases, even for moderate values of N.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

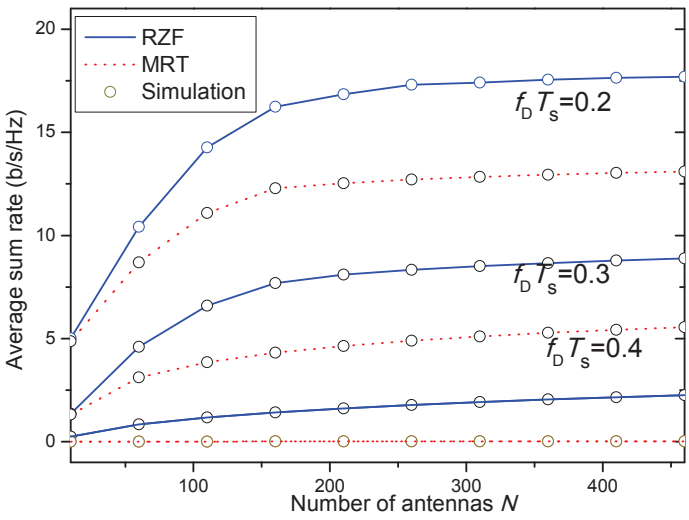

Fig. 3. Simulated and deterministic equivalent downlink sum-rates of the users with RZF and MRT precoders versus the number of antennas N for different normalized Doppler shifts $f_D T_s$.

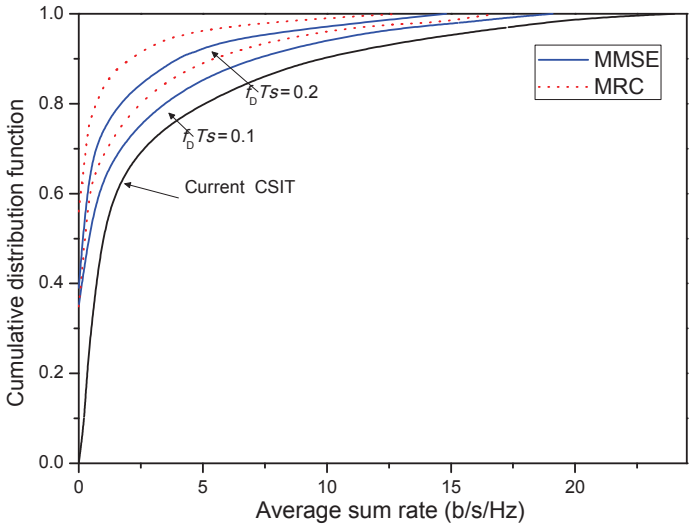

Fig. 4. The cumulative distribution function of the deterministic equivalent uplink sum-rate of the users with MMSE and MRC decoding for different normalized Doppler shifts $f_D T_s$, and $N = 80$ BS antennas.

Fig. 2 shows the uplink sum-rate of the users in the center cell as a function of the normalized Doppler shifts $f_D T_s$ for $N = 30, 60, 90$. Delayed CSIT decreases the downlink sum-rate to zero with some ripples following according to the behavior of the Bessel function $J_0(\cdot)$. Specifically, at $f_D T_s \approx 0.4$ we observe the first zero point and then with following ripples, the magnitude increases and later decreases to zero again and again, tending finally to zero. Note that by increasing the number of BS antennas N, a higher sum-rate is obtained, but the shape of the curves stays the same by keeping the zero points at constant specific values of $f_D T_s$. Apparently, the dependence of MMSE by delayed CSIT is identical to MRC, although MMSE achieves higher sum-rate under the same conditions of delayed CSIT, i.e., specific Doppler shift.

The effect of delayed CSIT is also shown in Fig. 3 for different Doppler shifts. In particular, the downlink sum-rate of the users is depicted as a function of the number of the BS antennas N. First, note that when $f_D T_s \approx 0.4$, the achievable sum-rate is almost zero with negligible increment as the number of antennas N increases. Nonetheless, the downlink sum-rate appears a rise with N, as expected. An in the case of MMSE, RZF provides higher sum-rate than MRT for the same delayed CSIT.

Fig. 4 shows the cumulative distribution function of uplink deterministic equivalent sum-rate of the users for different Doppler shifts together with the case of current CSIT, when $N = 80$. As expected, lower Doppler shift allows the transmission with a higher sum-rate at a given probability. Specifically, with probability 1, MMSE achieves 18.21 b/s/Hz with $f_D T_s = 0.1$ and 14.98 b/s/Hz with $f_D T_s = 0.2$ against MRT, which achieves 17.34 b/s/Hz and 13.46 b/s/Hz, respectively.

The efficiency of the channel predictor is investigated in Fig 5, which shows the uplink sum-rate of the users versus $f_D T_s$ for varying filter order p, as well as the sum-rate with current CSIT, when $N = 90$. Interestingly, the comparison between MMSE and MRC in the cases of no channel prediction ($p = 0$), $p = 5$ being of practical interest, and $p = 20,\ 30$

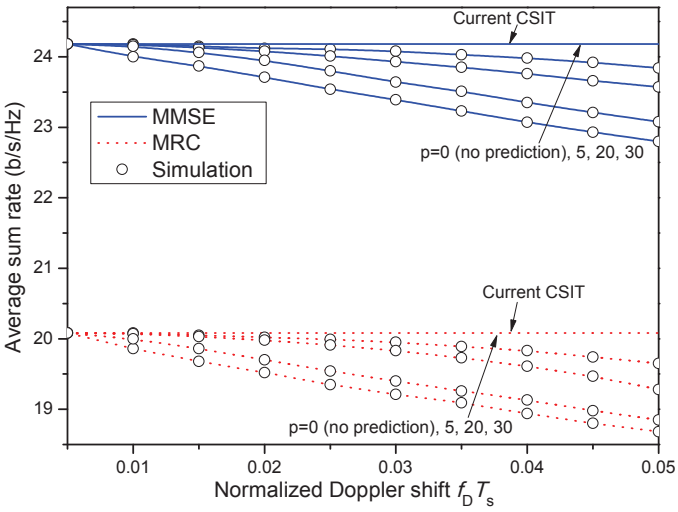

Fig. 5. Simulated and deterministic equivalent uplink sum-rates of the users with MMSE and MRC decoders versus the normalized Doppler shifts $f_D T_s$ for different order of the channel predictor, and $N = 90$ BS antennas.

for the sake of demonstration, enables insightful observations. Especially, channel prediction with higher value of p, i.e., more past observations, is required, in order to predict the degraded channel with higher precision. In addition, the inadequacy of the Wiener filter for a practical value of its order ($p = 5$) is shown. As a result, the need for efficient prediction leads to the quest of more sophisticated predictors, but this is outside the scope of this work. Moreover, in larger Doppler shifts and after applying both detectors, the ability of the filter to compensate for the loss because of the delayed CSIT weakens, but notably, MMSE behaves better by approaching closer than MRC the nominal value representing no delayed CSIT, i.e., current CSIT, as the Doppler shift increases.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

VI. Conclusion

In this paper, we analyzed in detail the uplink and downlink channels in a cellular large MIMO setting. In particular, we provided the corresponding deterministic equivalents of achievable rates with MMSE detector and RZF precoder prone to practical impairments, i.e., path loss, antenna correlation, pilot contamination, and delayed CSIT. However, our main purpose was to focus on the dependence by the delayed CSIT and how it affects the performance of the system in terms of sum-rate after applying a MMSE decoder (uplink) and a RZF precoder (downlink). We demonstrated that the sum-rate becomes negligible for specific Doppler shifts. Nevertheless, simulated results were presented, in order to verify our analysis. They showed that the deterministic analytical results are reliable even for practical values of BS antennas N and users K. Having tight approximations makes any study of the system by lengthy Monte Carlo simulations unnecessary. In particular, we noticed similar behavior of these techniques but with higher achievable rates with comparison to their simpler MRC detector and MRT precoder regarding the effect of delayed CSIT. Moreover, we applied a linear FIR Wiener predictor and derived the deterministic equivalent sum-rates for both MMSE and RZF. Especially, the predictor for MMSE achieves to regain most of the loss due to delayed CSIT, while in the case of MRC the results were not satisfactory. Overall, it was shown how effectively MMSE and RZF outperform against MRC and MRT without sacrificing any computational complexity due to the deterministic essence of the proposed results.

Appendix A
Useful Lemmas

Lemma 1 (Matrix inversion lemma (I) [41, Eq. 2.2]): Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be Hermitian invertible. Then, for any vector $\mathbf{x} \in \mathbb{C}^N$ and any scalar $\tau \in \mathbb{C}^N$ such that $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible,

$$\mathbf{x}^{\mathsf{H}}(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 2 (Matrix inversion lemma (II) [10, Lemma 2]): Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be Hermitian invertible. Then, for any vector $\mathbf{x} \in \mathbb{C}^N$ and any scalar $\tau \in \mathbb{C}^N$ such that $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible,

$$(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \mathbf{A} - \frac{\mathbf{A}^{-1}\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 3 (Rank-1 perturbation lemma [42, Lemma 2.1]): Let $z \in< 0$, $\mathbf{A} \in \mathbb{C}^{N\times N}$, $\mathbf{B} \in \mathbb{C}^{N\times N}$ with $\mathbf{B}$ Hermitian nonnegative definite, and $\mathbf{x} \in \mathbb{C}^N$. Then,

$$|\operatorname{tr}\left((\mathbf{B} - z\mathbf{I}_N)^{-1} - (\mathbf{B} + \mathbf{x}\mathbf{x}^{\mathsf{H}} - z\mathbf{I}_N)^{-1}\mathbf{A}\right| \leq \frac{\|\mathbf{A}\|}{|z|}.$$

Lemma 4: Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ with uniformly bounded spectral norm (with respect to N). Consider $\mathbf{x}$ and $\mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_x)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_y)$, are mutually independent and independent of $\mathbf{A}$. Then, we have

$$\frac{1}{N}\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x} - \frac{1}{N}\operatorname{tr}\mathbf{A}\mathbf{\Phi}_x \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{41}$$

$$\frac{1}{N}\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{y} \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{42}$$

$$\mathbb{E}\left[\left|\left(\frac{1}{N}\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x}\right)^2 - \left(\frac{1}{N}\operatorname{tr}\mathbf{A}\mathbf{\Phi}_x\right)^2\right|\right] \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{43}$$

$$\frac{1}{N^2}|\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{y}|^2 - \frac{1}{N^2}\operatorname{tr}\mathbf{A}\mathbf{\Phi}_x\mathbf{A}^{\mathsf{H}}\mathbf{\Phi}_y \xrightarrow[N\to\infty]{\text{a.s.}0} . \tag{44}$$

Theorem 5 ([14, Theorem 1]): Let $\mathbf{D} \in \mathbb{C}^{N\times N}$ and $\mathbf{S} \in \mathbb{C}^{N\times N}$ be Hermitian nonnegative definite matrices and let $\mathbf{H} \in \mathbb{C}^{N\times K}$ be a random matrix with columns $\mathbf{v}_k \sim \mathcal{CN}\left(0, \frac{1}{N}\mathbf{R}_k\right)$. Assume that $\mathbf{D}$ and the matrices $\mathbf{R}_k$, $k = 1, \ldots, K$, have uniformly bounded spectral norms (with respect to N). Then, for any $\rho > 0$,

$$\frac{1}{N}\operatorname{tr}\mathbf{D}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{D}\mathbf{T}(\rho) \xrightarrow[N_{\mathrm{t}}\to\infty]{\text{a.s.}} 0,$$

where $\mathbf{T}(\rho) \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}(\rho) = \left(\frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k}{1 + \delta_k(\rho)} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1},$$

and the elements of $\boldsymbol{\delta}(\rho) = [\delta_1(\rho)\ldots\delta_K(\rho)]^{\mathsf{T}}$ are defined as $\delta_k(\rho) = \lim_{t\to\infty}\delta_k^{(t)}(\rho)$, where for $t = 1, 2, \ldots$

$$\delta_k^{(t)}(\rho) = \frac{1}{N}\operatorname{tr}\mathbf{R}_k\left(\frac{1}{N}\sum_{j=1}^{K}\frac{\mathbf{R}_j}{1 + \delta_j^{(t-1)}(\rho)} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1} \tag{45}$$

with initial values $\delta_k^{(0)}(\rho) = \frac{1}{\rho}$ for all k.

Theorem 6 ([10, Theorem 2]): Let $\mathbf{\Theta} \in \mathbb{C}^{N\times N}$ be a Hermitian nonnegative definite matrix with uniformly bounded spectral norm (with respect to N). Under the same conditions as in Theorem 5,

$$\frac{1}{N}\operatorname{tr}\mathbf{D}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1}\mathbf{K}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{S} + \rho\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{D}\mathbf{T}'(\rho) \xrightarrow{\text{a.s.}} 0,$$

where $\mathbf{T}'(\rho) \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}'(\rho) = \mathbf{T}(\rho)\mathbf{K}\mathbf{T}(\rho) + \mathbf{T}(\rho)\frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k\delta_k'(\rho)}{\left(1 + \delta_k(\rho)\right)^2}\mathbf{T}(\rho)$$

with $\mathbf{T}(\rho)$ and $\boldsymbol{\delta}_k(\rho)$ as defined in Theorem 5 and $\boldsymbol{\delta}'(\rho) = [\delta_1'(\rho)\ldots\delta_K'(\rho)]^{\mathsf{T}}$ given by

$$\boldsymbol{\delta}'(\rho) = \left(\mathbf{I}_K - \mathbf{J}(\rho)\right)^{-1}\mathbf{v}(\rho). \tag{46}$$

$\mathbf{J}(\rho) \in \mathbb{C}^{K\times K}$ and $\mathbf{v}(\rho) \in \mathbb{C}^K$ are defined as

$$[\mathbf{J}(\rho)]_{kl} = \frac{\frac{1}{N}\operatorname{tr}\mathbf{R}_k\mathbf{T}(\rho)\mathbf{R}_l\mathbf{T}(\rho)}{N\left(1 + \delta_k(\rho)\right)^2},$$

$$[\mathbf{v}(\rho)]_k = \frac{1}{N}\operatorname{tr}\mathbf{R}_k\mathbf{T}(\rho)\mathbf{K}\mathbf{T}(\rho). \tag{47}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

APPENDIX B
PROOF OF THEOREM 1

After appropriate substitutions and scaling by $\frac{1}{N^2}$, the desired signal power becomes

$$S_{\mathrm{r},jm} = \frac{1}{N^2}\alpha^4\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\hat{\mathbf{h}}_{jjm}[n]\right|^2 = \frac{1}{N^2}\alpha^4\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2, \quad (48)$$

where we have used Lemma 1, and $\hat{\boldsymbol{\Sigma}}_{jm}[n]$ is defined as

$$\hat{\boldsymbol{\Sigma}}_{jm}[n] = \left(\alpha^2\hat{\mathbf{H}}_{jj}[n]\hat{\mathbf{H}}_{jj}^{\mathsf{H}}[n] - \hat{\mathbf{h}}_{jjm}[n]\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n] + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\right)^{-1}.$$

If we denote the deterministic equivalent signal power $\bar{S}_{\mathrm{r},bu}$, and apply Lemma 4 and Theorem 5, we have

$$\bar{S}_{\mathrm{r},jm} - \frac{1}{N^2}\alpha^4\left(\frac{\hat{\delta}_{jm}}{1+\hat{\delta}_{jm}}\right)^2 \xrightarrow[N_{\mathrm{t}}\to\infty]{\text{a.s.}} 0, \quad (49)$$

where $\hat{\delta}_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjm}\mathbf{T}_j$. Similarly, by taking into account $\mathbf{h}_{jjm}[n+1] - \mathbf{g}_{jjm}[n+1] = \tilde{\mathbf{e}}_{jjm}[n+1]$, (28) yields to

$$I_{\mathrm{r},jm} = \frac{\alpha^2}{N^2}\Bigg(\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\tilde{\mathbf{e}}_{jjm}[n+1]\right|^2 + \frac{{\sigma_j}^2}{p_{\mathrm{r}}}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\right|^2 + \sum_{(l,k)\neq(j,m)}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\mathbf{h}_{jlk}[n+1]\right|^2\Bigg). \quad (50)$$

Each term of (50) will be derived separately by using Lemmas 1 and 4 as well as Theorems 5 and 6. Specifically, the first term is written as[2]

$$\frac{\alpha^2}{N^2}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\tilde{\mathbf{e}}_{jjm}[n+1]\right|^2 = \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\tilde{\mathbf{e}}_{jjm}[n+1]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \asymp \frac{\alpha^2}{N^2}\frac{\frac{1}{N}\hat{\delta}'_{\mathrm{r},jm}}{\left(1+\hat{\delta}_{jm}\right)^2}, \quad (51)$$

where $\hat{\delta}'_{\mathrm{r},jm} = \frac{1}{N}\operatorname{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Phi}_{jjm})\mathbf{T}_j^{\mathrm{r}_{d1}}$ and $\mathbf{K}^{\mathrm{r}_{d1}} = \boldsymbol{\Phi}_{jjm}$, while, hereafter, the superscripts of $\mathbf{T}$ and $\mathbf{K}$ correspond to different instances of these variables. The next term is obtained as

$$\frac{\alpha^2}{N^2}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\right|^2 = \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 = \frac{\alpha^2}{N^2}\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}^2[n]\hat{\mathbf{h}}_{jjm}[n]}{\left(1+\frac{1}{N}\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]\right)^2} \asymp \frac{\alpha^2}{N^2}\frac{\frac{1}{N}\hat{\delta}''_{jm}}{\left(1+\hat{\delta}_{jm}\right)^2}, \quad (52)$$

where $\hat{\delta}''_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjm}\mathbf{T}_j^{\mathrm{r}_{d2}}$ and $\mathbf{K}^{\mathrm{r}_{d2}} = \mathbf{I}_N$. Consider now the last term. In particular, if $k \neq m$, we have

$$\frac{\alpha^2}{N^2}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\mathbf{h}_{jlk}[n+1]\right|^2 = \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlk}[n+1]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \asymp \frac{\alpha^2}{N^2}\frac{\mathbf{h}_{jlk}^{\mathsf{H}}[n+1]\hat{\boldsymbol{\Sigma}}_{jm}[n]\boldsymbol{\Phi}_{jjm}\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlk}[n+1]}{\left(1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]\right)^2}. \quad (53)$$

Since $\hat{\boldsymbol{\Sigma}}_{jm}[n]$ is not independent of $\mathbf{h}_{jjk}[n+1]$, the use of Lemma 2 gives

$$\hat{\boldsymbol{\Sigma}}_{jm}[n] = \hat{\boldsymbol{\Sigma}}_{jmk}[n] - \frac{\hat{\boldsymbol{\Sigma}}_{jmk}[n]\hat{\mathbf{h}}_{jjk}[n]\hat{\mathbf{h}}_{jjk}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jmk}[n]}{1+\hat{\mathbf{h}}_{jjk}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jmk}[n]\hat{\mathbf{h}}_{jjk}[n]}, \quad (54)$$

which introduces a new matrix $\hat{\boldsymbol{\Sigma}}_{jmk}[n]$ to (53) defined as

$$\hat{\boldsymbol{\Sigma}}_{jmk}[n] = \Big(\alpha^2\hat{\mathbf{H}}_{jj}[n]\hat{\mathbf{H}}_{jj}^{\mathsf{H}}[n] - \hat{\mathbf{h}}_{jjm}[n]\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n] - \hat{\mathbf{h}}_{jjk}[n]\hat{\mathbf{h}}_{jjk}^{\mathsf{H}}[n] + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\Big)^{-1}. \quad (55)$$

By substituting (54) into (53) and applying Theorems 5 and 6, we obtain

$$\frac{\alpha^2}{N^2}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\mathbf{h}_{jlk}[n+1]\right|^2 = \frac{\hat{\mu}_{jlkm}}{N\left(1+\hat{\delta}_{jm}\right)^2}, \quad (56)$$

where

$$\hat{\mu}_{jlkm} = \frac{\operatorname{tr}\mathbf{R}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d1}}}{N} - \frac{2\mathrm{Re}\{\hat{\vartheta}_{jlk}^*\hat{\vartheta}'_{jlkm}\}\left(1+\hat{\delta}_{jk}\right) - \hat{\vartheta}_{jlk}^2\hat{\delta}'''_{jk}}{\left(1+\hat{\delta}_{jk}\right)^2}$$

with $\hat{\vartheta}_{jlk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j$, $\hat{\vartheta}'_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d1}}$, and $\hat{\delta}'''_{jk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jjk}\mathbf{T}_j^{\mathrm{r}_{d1}}$. Note that we have used $\hat{\mathbf{h}}_{jjk}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jmk}[n]\hat{\mathbf{h}}_{jlk}[n] \asymp \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}T_j$ by means of Lemma 3 and Theorem 5. If $k = m$, we need to define

$$\hat{\mathbf{z}}_{\mathrm{p},jlm}[n] = \sum_{l'\neq l}\mathbf{h}_{jl'm}[n] + \frac{1}{\sqrt{p_{\mathrm{p}}\tau}}\tilde{\mathbf{z}}_{\mathrm{p},j}[n]. \quad (57)$$

As a result, $\hat{\mathbf{z}}_{\mathrm{p},jlm}[n]$ and $\mathbf{h}_{jlm}[n+1]$ are independent, while $\hat{\mathbf{z}}_{\mathrm{p},jlm}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{jm}^{-1} - \mathbf{R}_{jlm})$. Thus, the last term of (50) becomes

$$\frac{\alpha^2}{N^2}\left|\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_j[n]\mathbf{h}_{jlm}[n+1]\right|^2 = \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n+1]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \overset{(12)}{=} \frac{\alpha^2}{N^2}\left|\frac{\alpha\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 + \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{e}_{jlm}[n+1]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \overset{(8)}{=} \frac{\alpha^2}{N^2}\left|\frac{\alpha\mathbf{h}_{jlm}^{\mathsf{H}}[n]\mathbf{Q}_{jm}\mathbf{R}_{jjm}\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n]}{1+\hat{\mathbf{h}}_{jjm}^{\mathsf{H}}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 +$$

[2]Let a_n and b_n two infinite sequences. $a_n \asymp b_n$ denotes the equivalence relation $a_n - b_n \xrightarrow[N\to\infty]{\text{a.s.}} 0$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

$$\frac{\alpha^2}{N^2}\left|\frac{\alpha\hat{\mathbf{z}}^{\mathsf{H}}_{\mathrm{p},jlm}[n]\mathbf{Q}_{jm}\mathbf{R}_{jjm}\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n]}{1+\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2+ \frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{e}_{jlm}[n+1]}{1+\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2. \quad (58)$$

As shown above, application of Lemma 4 as well as Theorems 5 and 6 to the first and second term of (58) gives

$$\frac{\alpha^4}{N^2}\left|\frac{\mathbf{h}^{\mathsf{H}}_{jlm}[n]\mathbf{Q}_{jm}\mathbf{R}_{jjm}\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n]}{1+\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \asymp \frac{\alpha^4}{N^2}\left|\frac{\hat{\vartheta}_{jlm}}{1+\hat{\delta}_{jm}}\right|^2$$

$$\frac{\alpha^4}{N^2}\left|\frac{\hat{\mathbf{z}}^{\mathsf{H}}_{\mathrm{p},jlm}[n]\mathbf{Q}_{jm}\mathbf{R}_{jjm}\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{h}_{jlm}[n]}{1+\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \asymp \frac{\alpha^4}{N^2}\frac{\frac{1}{N}\hat{\vartheta}''_{jlkm}}{\left(1+\hat{\delta}_{jm}\right)^2} - \frac{\alpha^4}{N^2}\frac{\frac{1}{N}\hat{\vartheta}'''_{jlkm}}{\left(1+\hat{\delta}_{jm}\right)^2},$$

where $\hat{\vartheta}''_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d3}}$, $\hat{\vartheta}'''_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Phi}_{jlk}\mathbf{T}_j^{\mathrm{r}_{d4}}$, $\mathbf{K}^{\mathrm{r}_{d3}} = \mathbf{R}_{jjm}$ and $\mathbf{K}^{\mathrm{r}_{d4}} = \boldsymbol{\Phi}_{jlm}$. Similarly, for the last term of (58) we have

$$\frac{\alpha^2}{N^2}\left|\frac{\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\mathbf{e}_{jlm}[n+1]}{1+\hat{\mathbf{h}}^{\mathsf{H}}_{jjm}[n]\hat{\boldsymbol{\Sigma}}_{jm}[n]\hat{\mathbf{h}}_{jjm}[n]}\right|^2 \asymp \frac{\alpha^2}{N^2}\frac{\frac{1}{N^2}\operatorname{tr}(1-\alpha^2)\mathbf{R}_{jlm}\mathbf{T}_j^{\mathrm{r}_{d1}}}{\left(1+\hat{\delta}_{jm}\right)^2}.$$

If we neglect the terms that vanish asymptotically in (58) and make the necessary substitutions into (50), we obtain the corresponding deterministic equivalent. As a result, $\bar{\gamma}_{\mathrm{r},jm}(\alpha)$ can be derived and this concludes the proof.

Appendix C
Proof of Theorem 2

The signal power, divided by N^2, yields by using Lemma 1

$$S_{\mathrm{r},jm} = \frac{1}{N^2}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\bar{\mathbf{h}}_{jjm}[n+1]\right|^2 = \frac{1}{N^2}\left|\frac{\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}{1+\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2, \quad (59)$$

where we have defined $\bar{\boldsymbol{\Sigma}}_{jm}[n+1]$ as

$$\bar{\boldsymbol{\Sigma}}_{jm}[n+1] = \left(\bar{\mathbf{H}}_{jj}[n+1]\bar{\mathbf{H}}^{\mathsf{H}}_{jj}[n+1] - \bar{\mathbf{h}}_{jjm}\bar{\mathbf{h}}^{\mathsf{H}}_{jjm} + \mathbf{Z}_{\mathrm{r},j} + N\varphi_{\mathrm{r},j}\mathbf{I}_N\right)^{-1}.$$

When predicted CSI is known, the deterministic equivalent signal power $\bar{S}_{\mathrm{r},jm}$ becomes

$$\bar{S}_{\mathrm{r},jm} \asymp \frac{1}{N^2}\left|\frac{\alpha^2\bar{\delta}_{jm}}{1+\alpha^2\bar{\delta}_{jm}}\right|^2, \quad (60)$$

where $\bar{\delta}_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j$. As far as (28) is concerned, we have after dividing by N^2

$$I_{\mathrm{r},jm} = \frac{1}{N^2}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\breve{\mathbf{h}}_{jjm}[n+1]\right|^2 + \frac{1}{N^2}\frac{\sigma_j^2}{p_{\mathrm{r}}}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\right|^2 + \sum_{(l,k)\neq(j,m)}\frac{1}{N^2}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\mathbf{h}_{jlk}[n+1]\right|^2. \quad (61)$$

We start by applying Lemmas 1 and 4 as well as Theorems 5 and 6 to each term of (61). For the first term, we have

$$\frac{1}{N^2}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\breve{\mathbf{h}}_{jjm}[n+1]\right|^2 = \frac{1}{N^2}\left|\frac{\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\breve{\mathbf{h}}_{jjm}[n+1]}{1+\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2 \asymp \frac{1}{N^2}\frac{\frac{\alpha^2}{N}\bar{\delta}'_{\mathrm{r},jm}}{\left(1+\alpha^2\bar{\delta}_{jm}\right)^2}, \quad (62)$$

where $\bar{\delta}'_{\mathrm{r},jm} = \frac{1}{N}\operatorname{tr}(\mathbf{R}_{jjm} - \alpha^2\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$ and $\mathbf{K}^{\mathrm{r}_{p1}} = \boldsymbol{\Theta}_{jjm}(p,\alpha)$. Regarding the second term, we obtain

$$\frac{1}{N^2}\frac{\sigma_j^2}{p_{\mathrm{r}}}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\right|^2 = \frac{1}{N^2}\frac{\sigma_j^2}{p_{\mathrm{r}}}\frac{\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}^2_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}{\left(1+\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]\right)^2} \asymp \frac{1}{N^2}\frac{\sigma_j^2}{p_{\mathrm{r}}}\frac{\frac{\alpha^2}{N}\bar{\delta}''_{jm}}{\left(1+\alpha^2\bar{\delta}_{jm}\right)^2}, \quad (63)$$

where $\bar{\delta}''_{jm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p2}}$ and $\mathbf{K}^{\mathrm{r}_{p2}} = \mathbf{I}_N$. Finally, we focus on the last term, which in case $k \neq m$, since $\bar{\mathbf{h}}_{jjm}[n+1]$ and $\mathbf{h}_{jlk}[n+1]$ are mutually independent and after similar steps with (56), becomes

$$\frac{1}{N^2}\left|\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_j[n+1]\mathbf{h}_{jlk}[n+1]\right|^2 = \frac{1}{N^2}\left|\frac{\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\mathbf{h}_{jlk}[n+1]}{1+\bar{\mathbf{h}}^{\mathsf{H}}_{jjm}[n+1]\bar{\boldsymbol{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2 \asymp \frac{1}{N^2}\frac{\frac{\alpha^2}{N}\bar{\mu}_{jlkm}}{\left(1+\frac{\alpha^2}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjm}(p,\alpha)\mathbf{T}_j\right)^2}, \quad (64)$$

where

$$\bar{\mu}_{jlkm} = \frac{\operatorname{tr}\mathbf{R}_{jlk}\mathbf{T}_j^{\mathrm{r}_{p1}}}{N} - \frac{2\mathrm{Re}\{\alpha^4\bar{\vartheta}_{jlk}\bar{\vartheta}'_{jlkm}\}\left(1+\alpha^2\bar{\delta}_{jk}\right)\alpha^6\bar{\vartheta}^2_{jlk}\bar{\delta}'''_{jk}}{\left(1+\alpha^2\bar{\delta}_{jk}\right)^2}$$

with $\bar{\vartheta}_{jlk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jlk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p2}}$, $\bar{\vartheta}'_{jlkm} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jlk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$, and $\bar{\delta}'''_{jk} = \frac{1}{N}\operatorname{tr}\boldsymbol{\Theta}_{jjk}(p,\alpha)\mathbf{T}_j^{\mathrm{r}_{p1}}$. However, if $k = m$, $\bar{\mathbf{h}}_{jjm}[n+1]$ and $\frac{1}{N}\mathbf{h}_{jlm}[n+1]$ are not mutually independent, since from (7) and (16) we have

$$\bar{\mathbf{h}}_{jjm}[n+1] = \sum_{q=0}^{p}\mathbf{V}_{jjm,q}\left(\sum_{l'=1}^{L}\mathbf{h}_{jl'm}[n-q] + \frac{1}{\sqrt{p_{\mathrm{p}}\tau}}\tilde{\mathbf{z}}_{\mathrm{p},j}[n-q]\right). \quad (65)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

Consequently, the last term in (61) is written as

$$\frac{1}{N^2}\left|\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_j[n+1]\mathbf{h}_{jlm}[n+1]\right|^2 = \frac{1}{N^2}\left|\frac{\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_{jm}[n+1]\mathbf{h}_{jlm}[n+1]}{1+\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2 = \frac{1}{N^2}\left|\frac{\bar{\mathbf{y}}_{\mathrm{p},jm}[n]\mathbf{V}_{jjm}^{\mathsf{H}}\bar{\mathbf{\Sigma}}_{jm}[n+1]\mathbf{h}_{jlm}[n+1]}{1+\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2 = \frac{1}{N^2}\left|\frac{\left(\sum_{q=0}^{p}\tilde{\mathbf{y}}_{jm}^{\mathsf{H}}[n-q]\mathbf{V}_{jjm,q}^{\mathsf{H}}\right)\bar{\mathbf{\Sigma}}_{jm}[n+1]\mathbf{h}_{jlm}[n+1]}{1+\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_{jm}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2. \quad (66)$$

By considering (12), we obtain

$$\mathbf{h}_{jlm}[n+1] = \alpha^{p+1}\mathbf{h}_{jlm}[n-p] + \mathbf{v}_{jlm}[n+1], \quad (67)$$

where $\mathbf{v}_{jlm}[n+1] \sim \mathcal{CN}(\mathbf{0}, (1-\alpha^{2(p+1)})\mathbf{R}_{jlm})$. In addition, if we combine (7) and (67), we have for $q = 0, 1, \ldots, p$

$$\tilde{\mathbf{y}}_{jm}^{\mathsf{H}}[n-q] = \alpha^{p-q}\mathbf{h}_{jlm}^{\mathsf{H}}[n-p] + \sum_{t=0}^{p-q}\alpha^t\mathbf{e}_{jlm}[n-q-t] + \hat{\mathbf{z}}_{\mathrm{p},jlm}[n-q]. \quad (68)$$

Now, we can proceed with (66) by applying Lemma 4, i.e.,

$$\frac{1}{N^2}\left|\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_j^{\mathsf{H}}[n+1]\mathbf{h}_{jlm}[n+1]\right|^2 \overset{(67)}{\underset{(68)}{\asymp}} \frac{\alpha^{2(p+1)}}{N^2}\left|\frac{\mathbf{h}_{jlm}^{\mathsf{H}}[n-p]\left(\sum_{t=0}^{p}\alpha^{p-t}\mathbf{V}_{jjm,t}^{\mathsf{H}}\right)\bar{\mathbf{\Sigma}}_{jm}^{\mathsf{H}}[n+1]\mathbf{h}_{jlm}[n-p]}{1+\bar{\mathbf{h}}_{jjm}^{\mathsf{H}}[n+1]\bar{\mathbf{\Sigma}}_{jm}^{\mathsf{H}}[n+1]\bar{\mathbf{h}}_{jjm}[n+1]}\right|^2 \asymp \frac{\alpha^{2(p+2)}}{N^2}\left|\frac{\bar{\vartheta}_{jlm}}{1+\alpha^2\bar{\delta}_{jm}}\right|^2, \quad (69)$$

where the asymptotically negligible terms have been removed. After straightforward substitutions, we obtain $\bar{\gamma}_{\mathrm{r},jm}(\alpha, p)$.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Dr. Jakob Hoydis for his help and support in making this work possible.

REFERENCES

[1] A. Papazafeiropoulos and T. Ratnarajah, "Uplink performance of massive MIMO subject to delayed CSIT and anticipated channel prediction" in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, May 2014, pp. 3186–3189.

[2] ——, "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.

[3] M. Cudak, "IEEE 802.16m System Requirements," IEEE 802.16m Task Group, doc. IEEE 802.16m-07/002r10, Jan. 2010.

[4] "3GPP technical specification group radio access network, Further advancements for E-UTRA: LTE-Advanced feasibility studies in RAN WG4, 3GPP TR 36.815, Tech. Rep.," Mar. 2010.

[5] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Aug. 2013.

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

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

[8] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Sel. Areas in Com.*, vol. 31, no. 1, pp. 192–201, Feb. 2013.

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

[10] J. Hoydis, S. T. Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[11] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in Proc. of *IEEE Int. Conf. Acoustics Speech Signal Processing*, pp. 3464–3467, May 2011.

[12] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Channel estimation and linear precoding in multiuser multiple-antenna TDD systems." *IEEE Tran. on Veh. Tech.*, vol. 60, no. 5, pp. 2102–2116, Jun. 2011.

[13] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Sel. Areas in Com.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[14] S. Wagner, R. Couillet, M. Debbah, and D. Slock "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. on Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[15] R. Couillet and M. Debbah, "Signal processing in large systems: A new paradigm,' *IEEE Sign. Process. Mag.*, vol. 30, no. 1, pp. 24–39, Jan. 2013.

[16] J. Zhang, C.-K. Wen, S. Jin, X. Q. Gao, and K.-K. Wong, "Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4801–4813, Oct. 2013.

[17] Samsung, "Rel–12 and onward," in Workshop on *TD-LTE Enhancements and Evolution for Rel-12 and Beyond*, Apr. 2012.

[18] J. Hoydis, C. Hoek, T. Wild, and S. Brink, "Channel measurements for large antenna arrays," *IEEE Inter. Symp. on Wir. Com. Systems (ISWCS)*, pp. 811–815, Aug. 2012.

[19] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, "Pilot contamination reduction in multi-user TDD systems," in Proc. of *IEEE Int. Conf. Com.*, May 2010.

[20] B. Gopalakrishnan and N. Jindal, "An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas," in *Int. Work. Sign. Proc. Adv. in Wir. Commun.*, pp. 381–385, Jun. 2011.

[21] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[22] X. Zhang, M. Matthaiou, M. Coldrey, and E. Björnson, "Energy efficiency optimization in hardware-constrained large-scale MIMO systems," in *Proc. IEEE Intern. Sym. Wireless Commun. Systems (ISWCS)*, Barcelona, Spain, Aug. 2014.

[23] A. Pitarokoilis, S. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," accepted to appear in *IEEE Trans. Wireless Commun.*, Sep. 2014.

[24] K. T. Truong and R. W. Heath, Jr., "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, available on arXiv: 1305.6151.

[25] A. Papazafeiropoulos, "Impact of user mobility on optimal linear receivers in cellular networks CSIT," accepted in *IEEE International Conference on Communications (ICC), 2015*, London, June 2015.

[26] A. Papazafeiropoulos, H. Q. Ngo, M. Matthaiou, and T. Ratnarajah "Uplink performance of conventional and massive MIMO cellular systems with delayed CSIT," *in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Washington, D.C., September 2014, pp. 574–579.

[27] 3GPP TR 36.814, "Evolved universal terrestrial radio access (E-UTRA): Further advancements for E-UTRA physical layer aspects," available at *http://www.3gpp.org/ftp/Specs/archive/36series/36.814/36814-900.zip*.

[28] S. Verdu, "*Multiuser Detection.*" Cambridge University Press, 1998.

[29] K. E. Baddour and N. C. Beaulieu, "Autoregressive modelling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, July 2005.

[30] W. C. Jakes, *Microwave Mobile Communications. New York: Wiley*, 1974.

[31] M. Vu and A. Paulraj, "On the capacity of MIMO wireless channels with dynamic CSIT," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1269–1283, Sep. 2007.

[32] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE Journal of Sel. Top. in Sig. Proc.*, vol.8, no.5, pp. 787–801, Oct. 2014.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2443040, IEEE Transactions on Wireless Communications

[33] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE Journal of Sel. Top. in Sig. Proc.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.
[34] J. So, D. Kim, Y. Lee, and Y. Sung, "Pilot signal design for massive MIMO systems: A received signal-to-noise-ratio-based aApproach," *IEEE Sig. Proc. Letters*, vol.22, no.5, pp. 549–553, May 2015.
[35] S. Haykin, *Adaptive Filtering Theory*, 2nd ed. Englewood Cliffs, NY: Prentice-Hall, 1991.
[36] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[37] P. Billingsley, "*Probability and measure*," 3rd ed. John Wiley & Sons, Inc., 1995.
[38] A. W. van der Vaart, "*Asymptotic statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*," Cambridge University Press, New York, 2000.
[39] R. Muharar, R. Zakhour, and J. Evans, "Optimal power allocation and user loading for multiuser MISO channels with regularized channel inversion," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5030–5041, Dec. 2013.
[40] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE T. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.
[41] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.
[42] Z. D. Bai and J. W. Silverstein, "On the signal-to-interference ratio of CDMA systems in wireless communications," *The Annals of Applied Probability*, vol. 17, no. 1, pp. 81–101, 2007.

Anastasios Papazafeiropoulos is currently a Marie Curie Fellow at Imperial College London, 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. 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.

Tharmalingam. Ratnarajah is currently with the Institute for Digital Communications, University of Edinburgh, Edinburgh, U.K., as a Professor in signal processing and communications. His research interests include signal processing and information theoretic aspects of 5G 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 250 publications in these areas and holds four U.S. patents. He is currently the coordinator of the FP7 projects HARP (3.2M€) in the area of highly distributed MIMO and ADEL (3.7M€) in the area of licensed shared access. Previously, he was the coordinator of FP7 Future and Emerging Technologies project CROWN (2.3M€) in the area of cognitive radio networks and HIATUS (2.7M€) in the area of interference alignment. Dr Ratnarajah is a member of the American Mathematical Society and Information Theory Society and Fellow of Higher Education Academy (FHEA), U.K.