IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 5, 2017, accepted August 29, 2017, date of publication September 7, 2017, date of current version September 27, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2749602

Channel Estimation for Massive MIMO TDD Systems Assuming Pilot Contamination and Frequency Selective Fading

FELIPE A. P. DE FIGUEIREDO[1], FABBRYCCIO A. C. M. CARDOSO[2], INGRID MOERMAN[1], AND GUSTAVO FRAIDENRAICH[3]

[1]Department of Information Technology, Ghent University, 9052 Ghent, Belgium
[2]CPqD–Research and Development Center on Telecommunications, Campinas 13086-902, Brazil
[3]DECOM/FEEC, State University of Campinas, Campinas 13083-852, Brazil

Corresponding author: Felipe A. P. de Figueiredo (felipe.pereira@ugent.be)

ABSTRACT Channel estimation is crucial for massive multiple-input multiple-output (MIMO) systems to scale up multi-user MIMO, providing significant improvement in spectral and energy efficiency. In this paper, we present a simple and practical channel estimator for multipath multi-cell massive MIMO time division duplex systems with pilot contamination, which poses significant challenges to channel estimation. The proposed estimator addresses performance under moderate to strong pilot contamination without previous knowledge of the inter-cell large-scale fading coefficients and noise power. Additionally, we derive and assess an approximate analytical mean square error (MSE) expression for the proposed channel estimator. We show through simulations that the proposed estimator performs asymptotically as well as the minimum MSE estimator with respect to the number of antennas and multipath coefficients.

INDEX TERMS Massive MU-MIMO, channel estimation, pilot contamination, multipath, Zadoff-Chu sequences.

I. INTRODUCTION

Massive Multiple-Input Multiple-Output, (MMIMO) antenna systems potentially allow base stations (BS) to operate with huge improvements in spectral and radiated energy efficiency, using relatively low complexity linear processing techniques. The improvement in spectral efficiency is achieved by serving several terminals in the same time-frequency resource through spatial multiplexing, and the increase in energy efficiency is due to the array gain yielded by the large number of antennas available at the BS [1], [2].

The expected massive MIMO improvements assume that accurate channel estimations are available at both the receiver and transmitter for detection and precoding operations, respectively. Additionally, the reuse of frequencies and pilot sequences in cellular communications systems leads to degradation in the channel estimation performance. As mentioned in the literature, this inter-cellular interference is known as pilot contamination [2], [3].

Massive MIMO systems operating in time division duplex (TDD) mode assume channel reciprocity between uplink and downlink in order to minimize pilot overhead, transmitting pilot reference signals only in the uplink. In this scenario, pilot overhead cost is only proportional to the number of terminals and improved channel estimation quality can be achieved due to the large number of antennas available at the BS [3]–[5]. A BS estimates channel states usually based on least squares (LS) [6] or minimum mean square error (MMSE) [7], [8] linear estimators. In addition, inter and intra-cell large-scale fading coefficients are assumed to be perfectly known at the BS when applying the MMSE estimator in a great number of works [7]–[9].

A countless number of works in the literature only deals with the problem of channel estimation in flat fading channels [3], [5]–[18], which is applicable only for OFDM based systems and does not reflect the true nature of communications channels in practice [19].

In this work, we study the channel estimation problem and its influences on the performance of multi-cell massive MIMO TDD systems under the assumptions of time-varying frequency selective fading channels and frequency/pilot-sequence reuse. This is an extension of the work presented in [18] where only the assumption of flat fading

channels is considered. Our main contributions in this work are: (i) the proposal of an uplink training scheme using Zadoff-Chu (ZC) sequences to ease the channel estimation process; (ii) the use of the minimum-variance unbiased estimator (MVUE) method to estimate the interference (*i.e.*, inter-cell large-scale fading coefficients) plus noise power term and replace it into the ideal MMSE channel estimator, resulting in an efficient, more practical and simpler channel estimator; (iii) the derivation of an approximate analytical MSE expression for the proposed estimator; and (iv) a thorough analysis of the proposed channel estimator and its approximate MSE expression. Numerical results demonstrate that the proposed channel estimator performs asymptotically as well as the ideal MMSE channel estimator as the number of antennas and channel paths increases.

The paper is organized in the following way. Section II defines the problem at hand and briefly discusses two linear channel estimators, namely, LS and MMSE, which will be used for comparison analysis with the proposed channel estimator. Section III presents and analyses the proposed channel estimator and its approximate analytical MSE expression. The performance of the proposed estimator is assessed in Section IV. Section V then concludes the work.

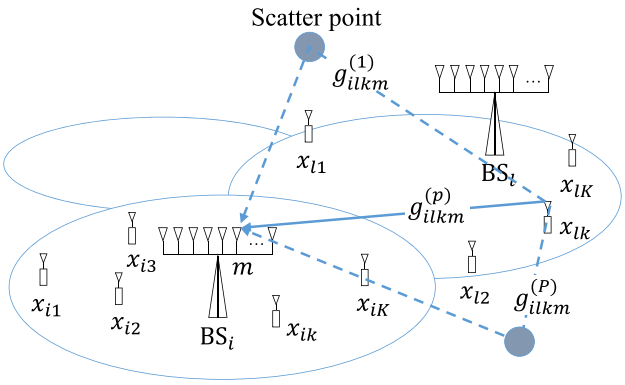

FIGURE 1. Problem structure.

II. PROBLEM DEFINITION

In this work we consider a multi-cell multi-user massive MIMO system with L cells, where a TDD protocol is employed. Each one of the cells has a BS at its center equipped with M co-located antennas and K single antenna users randomly located within the cell, where $M \gg K$. All the K users in a cell are served at the same time/frequency resource. Fig. 1 illustrates the problem we are dealing with, where the solid arrow corresponds to the possible direct path and the dashed lines correspond to paths where the signal from the kth user in the lth cell reaches the mth antenna at the ith base station through reflections caused by objects such as buildings and mountains.

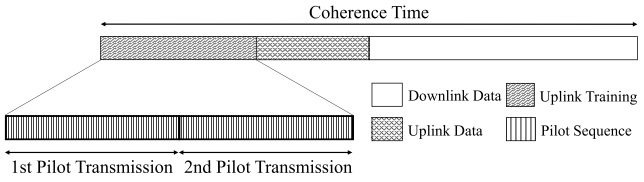

FIGURE 2. TDD transmission protocol.

Based on the assumption of channel reciprocity, we adopt the TDD protocol depicted in Fig. 2 and proposed in [20]. Due to the reciprocity principle, only uplink channels need to be estimated while downlink channels are equal to the transpose of the uplink channels. It is important to note that the length of the TDD frames are limited by the channel coherence time [20], [21]. According to the adopted TDD protocol, first, all users in all cells send their uplink training sequences synchronously. After that, the BSs use the training sequences to estimate the uplink channels. Next, the users send uplink data signals. Then, the BSs use the estimated channels to detect and decode uplink data and generate precoding matrices employed to transmit downlink data.

We also assume, in this work, wireless radio channels that take into account both large and small scale fading effects. Large-scale fading represents shadowing and attenuation or path-loss over a large area, while small-scale fading refers to the abrupt changes in signal amplitude and phase as a result of very small changes (in the order of $\lambda/2$) in the spatial separation between the user and the BS [19]. We consider that multipath propagation is the main effect behind small scale fading. Multipath causes multiple copies of a signal to travel over different paths with different propagation delays. These copies are received at the receiver with different phase angles and strengths [22].

The wireless channels are considered static during the channel coherence time and independent across all the M antennas and K users. We begin our derivation by defining $g_{ilkm}^{(p)}$ as the complex gain of the pth path belonging to the channel connecting the kth user in the lth cell to the mth antenna of the BS in the ith cell. Additionally, the complex gain $g_{ilkm}^{(p)}$ can be expressed as a complex fast-fading factor (small-scale fading coefficient) times an amplitude factor that accounts for path-loss and log-normal shadow fading (large-scale fading coefficient),

$$g_{ilkm}^{(p)} = h_{ilkm}^{(p)} \sqrt{\beta_{ilk}}, \tag{1}$$

where $i = 1, \cdots L$; $l = 1, \cdots L$; $k = 1, \cdots, K$; $m = 1, \cdots, M$; $p = 0, \cdots, P-1$ and the fast fading coefficients, $h_{ilkm}^{(p)}$, are assumed to follow a circularly-symmetric complex normal distribution with $\mathcal{CN}(0, 1)$. The amplitude factor, $\sqrt{\beta_{ilk}}$, is assumed constant with respect to the index of the base station antenna since the path-loss and shadow fading change slowly over space [1], [22], [23]. Additionally, we assume that the set of large-scale fading coefficients, represented by $\{\beta_{ilk}\}$, is deterministic during the channel estimation. This assumption is in accordance with other works in the literature [1], [9], [14], [18].

The multipath channels connecting the kth user in the lth cell to the mth antenna at the BS in the ith cell are modeled as $1 \times P$ vectors, $\mathbf{g}_{ilkm} = \left[g_{ilkm}^{(0)}, g_{ilkm}^{(1)}, ..., g_{ilkm}^{(P-1)}\right]$, where P denotes the number of paths of the channel. Additionally, we assume that the channels have at most P paths. The $M \times KP$ channel matrix including all the $\mathbf{g}_{ilkm}$ vectors is represented by

$$\mathbf{G}_{il} = \begin{bmatrix} \mathbf{g}_{il11} & \mathbf{g}_{il21} & \cdots & \mathbf{g}_{ilK1} \\ \mathbf{g}_{il12} & \mathbf{g}_{il22} & \cdots & \mathbf{g}_{ilK2} \\ \mathbf{g}_{il13} & \mathbf{g}_{il23} & \cdots & \mathbf{g}_{ilK3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{g}_{il1M} & \mathbf{g}_{il2M} & \cdots & \mathbf{g}_{ilKM} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{il1} & \mathbf{G}_{il2} & \cdots & \mathbf{G}_{ilK} \end{bmatrix}. \quad (2)$$

This matrix contains all the channels for all K users in the lth cell to all the M antennas at the BS in the ith cell. The columns in $\mathbf{G}_{il}$, defined by $\mathbf{G}_{ilk}$, represent the channels from the kth user in the lth cell to all the M antennas at the BS in the ith cell. The BS at the ith cell requires knowledge of the channels of the users within its coverage area, denoted by $\mathbf{G}_{iik} : \forall k$, in order to perform detection and precoding operations.

A. UPLINK TRAINING

In order for the BS at the ith cell to estimate all the M multipath channels from a given user to itself, each user is required to transmit an uplink training sequence. We suppose the existence of pilot sequence reuse, which is a common scenario in massive MIMO. Additionally, we assume a pilot reuse factor of one and the existence of time-synchronization among all the L cells. In this scenario, users from all the L cells transmit the same pilot sequences at the same time and frequency resources. All these assumptions constitute the worst possible case for a massive MIMO system [1]. The kth user has their pilot sequence represented by $\mathbf{s}_k = [s_k(0), s_k(1), \cdots, s_k(N-1)]^T$, where N denotes the length of the pilot sequences.

As we consider time-varying frequency selective fading channels, each one of the K uplink training sequences has to be convolved with all the M channels connecting a user to a specific BS. Therefore, with the purpose of expressing these convolutions in matrix form, we define the $N \times P$ matrix, $\mathbf{S}_k$, with circularly-shifted versions of the pilot sequence $\mathbf{s}_k$ as shown below

$$\mathbf{S}_k = \begin{bmatrix} s_k(0) & s_k(N-1) & \cdots & s_k(N-P+1) \\ s_k(1) & s_k(0) & \cdots & s_k(N-P+2) \\ s_k(2) & s_k(1) & \cdots & s_k(N-P+3) \\ \vdots & \vdots & \ddots & \vdots \\ s_k(N-1) & s_k(N-2) & \cdots & s_k(N-P+N) \end{bmatrix}. \quad (3)$$

This is a circulant matrix with orthogonality property $\mathbf{S}_k^H \mathbf{S}_k = N\mathbf{I}_p$. In order to obtain circular matrices as shown in (3), users have to transmit their pilot sequences, $\mathbf{s}_k$, twice, rendering them uplink training sequences with length $2N$, as showed in Fig. 2. This is mainly due to the orthogonality property exhibited by (3). The pilot sequences are formed through the application of cyclic shifts to ZC root sequences with prime length N. It is worth mentioning that cyclically shifted versions of a ZC sequence are orthogonal to each other [24]. ZC sequences are used in this work due to their properties of constant amplitude and zero autocorrelation, however, any other type of sequence could have been adopted as long as it exhibits the properties mentioned earlier. In order for all the users in a given cell to have valid pilots, the inequality $N > KP$ must hold true. For the case when $N \geq KPL$ pilot contamination vanishes as each one of the users in the L cells will have sequences that are orthogonal to all other pilot sequences in all other cells, however, this case could be impractical as the pilot sequence could get very long depending on the number of cells, L. It is also important to mention that the number of pilot symbols, N, must be the closest prime number to $K \times P$. In case $N \leq KP$, it is not possible to form orthogonal pilot sequences out of the same ZC root sequence for all users. Additionally, the set of pilot sequences of the K users can be described by a $N \times KP$ matrix $\mathbf{S}$ of the form $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_K]$.

The received uplink training sequences at the BS in the ith cell can be represented as a $M \times N$ matrix defined as

$$\mathbf{Y}_i = \sqrt{q} \sum_{l=1}^{L} \mathbf{G}_{il} \mathbf{S}^H + \mathbf{N}_i, \quad (4)$$

where q is regarded as the transmit signal to noise ratio (Tx-SNR) and $\mathbf{N}_i$ represents a $M \times N$ noise matrix with independent and identically distributed elements following a circularly-symmetric complex normal distribution, $\mathcal{CN}(0, 1)$.

B. LS CHANNEL ESTIMATOR

As shown in the last section, $\mathbf{S}_k$ represents the kth pilot matrix of the overall matrix, $\mathbf{S}$. Therefore, a sufficient statistic for estimating channel $\mathbf{G}_{ilk}$ at the BS in the ith cell can be defined by

$$\mathbf{Z}_{ik} = \frac{1}{\sqrt{qN}} \mathbf{Y}_i \mathbf{S}_k = \sum_{l=1}^{L} \mathbf{G}_{ilk} + \frac{\mathbf{N}_i \mathbf{S}_k}{\sqrt{qN}}, \quad (5)$$

where $\mathbf{Z}_{ik}$ is a $M \times P$ matrix. Each one of $\mathbf{Z}_{ik}$'s columns follows a $\mathcal{CN}(\mathbf{0}_M, \zeta_{ik}\mathbf{I}_M)$ distribution where

$$\zeta_{ik} = \sum_{l=1}^{L} \beta_{ilk} + \frac{1}{qN}. \quad (6)$$

As can be noticed, ζ_{ik} is the term representing the combined interference, *i.e.*, the summation of all large-scale fading coefficients, plus noise power from the kth users with the same pilot sequence to the BS at the ith cell.

Additionally, the term corresponding to noise in (5) has each one of its columns following a $\mathcal{CN}(\mathbf{0}_M, \frac{1}{qN}\mathbf{I}_M)$ distribution.

The least squares channel estimator is given by [25]

$$\hat{\mathbf{G}}_{iik}^{\mathrm{LS}} = \mathbf{Z}_{ik}. \quad (7)$$

The LS channel estimator has its MSE per antenna and per channel path defined by

$$\eta_{ik}^{\mathrm{LS}} = \frac{1}{MP}\,\mathrm{Tr}\left\{\mathbb{E}\left[\left(\hat{\mathbf{G}}_{iik}^{\mathrm{LS}} - \mathbf{G}_{iik}\right)\left(\hat{\mathbf{G}}_{iik}^{\mathrm{LS}} - \mathbf{G}_{iik}\right)^{H}\right]\right\} = \zeta_{ik} - \beta_{iik}. \tag{8}$$

When compared to the MMSE channel estimator, the LS estimator presents higher MSE, however, on the other hand, it does not require previous knowledge of the set of large-scale fading coefficients, $\beta_{ilk}, \forall l$. Although presenting higher MSE than the MMSE estimator, the low complexity and simplicity of the LS estimator are important aspects that can justify its adoption in real-time embedded systems. As can be noticed, the MSE, η_{ik}^{LS}, decreases with increasing q and decreasing β_{ilk}, $\forall i \neq l$ (*i.e.*, smaller interference level).

Remark 1: As a result of pilot contamination, as $q \to \infty$, $\eta_{ik}^{\mathrm{LS}} \to \sum_{l=1, l\neq i}^{L} \beta_{ilk}$.

C. MMSE CHANNEL ESTIMATOR

Statistical Bayesian MMSE channel estimation supposes previous knowledge of interference, *i.e.*, large-scale fading coefficients and noise statistics. In general, MMSE channel estimation provides simpler analytical results for lower-bound error performance, and a large number of works in massive MIMO literature employs it in order to gain knowledge of the channels [3], [8]. Those works suppose perfect knowledge of the interference and noise statistics, which might not be reasonable in practice [19]. Therefore, based on the assumption of perfect knowledge of the set of large-scale fading coefficients, $\{\beta_{ilk}\}$, and noise statistics, the ideal MMSE estimator is defined by [25]

$$\hat{\mathbf{G}}_{iik}^{\mathrm{MMSE}} = \frac{\beta_{iik}}{\zeta_{ik}}\mathbf{Z}_{ik}, \tag{9}$$

where each one of its columns follows the $\mathcal{CN}(\mathbf{0}_M, \frac{\beta_{iik}^2}{\zeta_{ik}}\mathbf{I}_M)$ distribution.

The ideal MMSE channel estimator has its MSE per antenna and per channel path defined by [25]

$$\eta_{ik}^{\mathrm{MMSE}} = \frac{1}{MP}\,\mathrm{Tr}\left\{\mathbb{E}\left[\left(\hat{\mathbf{G}}_{iik}^{\mathrm{MMSE}} - \mathbf{G}_{iik}\right)\left(\hat{\mathbf{G}}_{iik}^{\mathrm{MMSE}} - \mathbf{G}_{iik}\right)^{H}\right]\right\} = \beta_{iik}\left(1 - \frac{\beta_{iik}}{\zeta_{ik}}\right). \tag{10}$$

As can be noticed, the MSE, $\eta_{ik}^{\mathrm{MMSE}}$, decreases with increasing q, decreasing β_{iik}, or decreasing β_{ilk}, $\forall i \neq l$ (*i.e.*, smaller interference level).

Remark 2: As a result of pilot contamination, as $q \to \infty$, $\eta_{ik}^{\mathrm{MMSE}} \to \beta_{iik}\left(1 - \frac{\beta_{iik}}{\sum_{l=1}^{L}\beta_{ilk}}\right)$.

III. PROPOSED CHANNEL ESTIMATOR

The process of obtaining the inter-cell large-scale fading coefficients is unjustified given the unreasonable overburden necessary to estimate them [26]. For example, assuming the deployment of L cells where each cell has K users, one given BS would have to estimate $K(L-1)$ inter-cell large-scale fading coefficients. Through a careful analysis of the ideal MMSE channel estimator, it is possible to notice that instead of the individual large-scale fading coefficients, β_{ilk}, as supposed in many previous papers [1], [3], [6]–[9], [14], it is only needed to estimate the term ζ_{ik}, which represents the sum of all large-scale fading coefficients plus noise statistic. This straightforward observation produces a simpler and efficient answer to the problem of acquiring the inter-cell large-scale coefficients. Therefore, the approach we put forward in this work consists of the estimation of ζ_{ik} and its subsequent use in the ideal MMSE channel estimator (9), resulting in the channel estimator derived next.

A possible minimum variance unbiased estimator (MVUE) of ζ_{ik}, given the observed signal $\mathbf{Z}_{ik}$ is defined by [25]

$$\hat{\zeta_{ik}} = \frac{\|\mathbf{Z}_{ik}\|_F^2}{MP}, \tag{11}$$

where $\|\mathbf{Z}_{ik}\|_F^2 = \mathrm{Tr}\left(\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}\right)$ is the squared Frobenius norm of $\mathbf{Z}_{ik}$. It is important to note that the Trace operator, Tr(.), eliminates, setting to 0, all the inner products between estimated channels of the kth user to the M antennas at the BS in the ith cell while summing all the elements on the main diagonal. Other possible MVUE's would be the ones that take into account one or more (less than P) elements in the main diagonal of $\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}$.

Equation (11) arises from the fact that $\mathbb{E}\left\{\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}\right\} = M\zeta_{ik}\mathbf{I}_P$, that the unbiased estimator has $\mathbb{E}\{\hat{\zeta_{ik}}\} = \zeta_{ik}$ and also from the observation that as P increases, the term $\|\mathbf{Z}_{ik}\|_F^2/P$ tends to $M\zeta_{ik}$, therefore resulting in a better estimator in the sense that it approaches ζ_{ik} faster than the other possible MVUE's once it is the average over all P inner products in the main diagonal of $\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}$.

It is possible to write $\zeta_{ik} = \hat{\zeta_{ik}} + e_{ik}$ where e_{ik} is the estimation error. The estimator $\hat{\zeta_{ik}}$ has Gamma distribution with shape (k) equal to MP, scale (θ) equal to ζ_{ik}/MP, $\mathbb{E}\{\hat{\zeta_{ik}}\} = \zeta_{ik}$ and $\mathrm{var}\{\hat{\zeta_{ik}}\} = \zeta_{ik}^2/MP$. Under the MVUE framework, the parameter to be estimated is considered to be deterministic [25], and thus e_{ik} has zero mean and variance given by ζ_{ik}^2/MP, which is the same variance of $\hat{\zeta_{ik}}$.

Finally, swapping ζ_{ik} with $\hat{\zeta_{ik}}$ in (9) yields the proposed channel estimator, which is given by

$$\hat{\mathbf{G}}_{iik}^{\mathrm{prop}} = MP\frac{\beta_{iik}}{\|\mathbf{Z}_{ik}\|_F^2}\mathbf{Z}_{ik}. \tag{12}$$

As will be shown next, the proposed channel estimator asymptotically approaches the ideal MMSE channel estimator with regard to both M and P. For the case when $P = 1$ it is easy to notice that equation (12) reduces to equation (8) presented in [18]. An approximation to the MSE per antenna and per channel path of the proposed channel estimator is

defined by

$$\eta_{ik}^{\text{prop}} = \frac{1}{MP}\,\text{Tr}\left\{\mathbb{E}\left[\left(\hat{\mathbf{G}}_{iik}^{\text{prop}} - \mathbf{G}_{iik}\right)\left(\hat{\mathbf{G}}_{iik}^{\text{prop}} - \mathbf{G}_{iik}\right)^{H}\right]\right\} \approx \beta_{iik}\left[1 - \frac{(MP-2)\beta_{iik}}{(MP-1)\zeta_{ik}}\right]. \quad (13)$$

The approximate MSE for the proposed estimator, η_{ik}^{prop}, decreases with increasing q, increasing MP, decreasing β_{iik}, or decreasing β_{ilk}, which means smaller interference level from other cells, *i.e.*, smaller pilot contamination.

Remark 3: As a result of pilot contamination, as $q \to \infty$ and $MP \to \infty$, $\eta_{ik}^{\text{prop}} \to \beta_{iik}\left(1 - \frac{\beta_{iik}}{\sum_{l=1}^{L}\beta_{ilk}}\right)$.

Remark 3 clearly shows that the MSE of the proposed channel estimator tends to the MSE of the ideal MMSE channel estimator when both q and $MP \to \infty$. The proof for the approximation of the MSE is given in Appendix A.

Remark 4: The mean squared Euclidean distance per antenna and per channel between $\hat{\mathbf{G}}_{iik}^{\text{prop}}$ and $\hat{\mathbf{G}}_{iik}^{\text{MMSE}}$ is given by

$$\frac{1}{MP}\,\text{Tr}\left\{\mathbb{E}\left[\left(\hat{\mathbf{G}}_{iik}^{\text{prop}} - \hat{\mathbf{G}}_{iik}^{\text{MMSE}}\right)\left(\hat{\mathbf{G}}_{iik}^{\text{prop}} - \hat{\mathbf{G}}_{iik}^{\text{MMSE}}\right)^{H}\right]\right\} = \frac{1}{MP-1}.\frac{\beta_{iik}^{2}}{\zeta_{ik}}. \quad (14)$$

The proof of Remark 4 is presented in Appendix B. After analyzing equations (6) and (14) it is easily noticeable that the mean squared Euclidean distance in (14) decreases as q decreases, MP increases, β_{iik} decreases, and $\beta_{ilk}, i \neq l$ increases.

TABLE 1. Simulation parameters.

Parameter	Description	Value
L	Total number of cells	7
K	Number of users per cell	10
P	Number of multipath coefficients per channel	20
N	Pilot length in symbols	223
-	Frequency and pilot reuse factors	1

IV. SIMULATION RESULTS AND ANALYSIS

A thorough analysis of the performance of the proposed channel estimator is carried out in this section. We use the MMSE and LS channel estimators for some of the comparisons with the proposed channel estimator. A usual multi-cell deployment as depicted in Fig. 1 with the parameters described in Table 1 is adopted in this work.

We consider two different cases for setting the large-scale fading coefficients, $\{\beta_{ilk}\}$, one considering constant values and other considering random values. In the case where the coefficients are considered constant, β_{iik} is set to 1 and $\beta_{ilk}, \forall\, l \neq i$ is set to a. In the case where the coefficient's values are randomly selected, the users belonging to each one of the L cells are uniformly distributed within the area

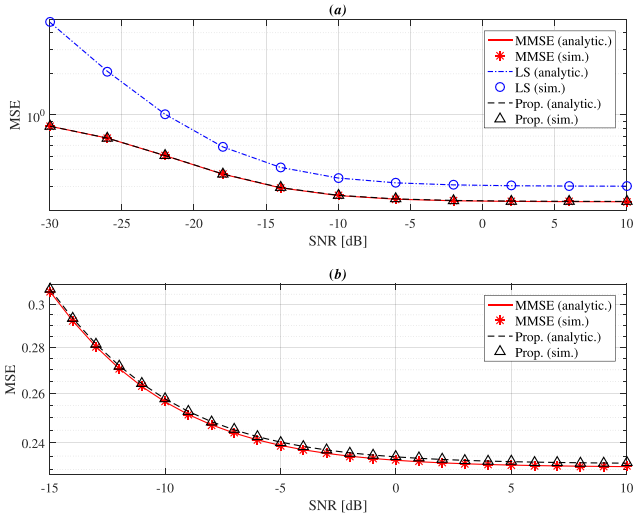

FIGURE 3. MSE of channel Estimation versus uplink pilot power. (a) Comparison of proposed estimator with LS and MMSE estimators. (b) Difference between ideal MMSE and proposed channel estimators.

comprised by two circles with radius $d_0 = 100$ m and $d_1 = 1000$ m respectively. The set of large-scale fading coefficients $\{\beta_{ilk}\}$ is independently generated by $\beta_{ilk} = \psi / \left(\frac{d_{ilk}}{d_0}\right)^{v}$, where $v = 3.8, 10\log_{10}(\psi) \sim \mathcal{N}(0, \sigma^2_{\text{shadow,dB}})$ with $\sigma_{\text{shadow,dB}} = 8$, and d_{ilk} is the distance of the kth user in lth cell to BS in the ith cell.

Fig. 3 (a) depicts the channel estimation MSE versus the transmit signal to noise ratio, q, for $M = 30$ and $a = 0.05$. As is evident from the figure, both simulation and analytical MSE results match for all the compared estimators. The MSE of all channel estimators decreases with the increase of the transmit signal to noise ratio. It can also be easily noticed that all the three channel estimators present MSE floors due to pilot contamination. This confirmation is in accordance with Remarks 1, 2 and 3. The MSE of the proposed channel estimator gets quite close to the MSE of the ideal MMSE channel estimator at low transmit signal to noise ratio values while the difference between the LS and both the ideal MMSE and proposed channel estimators increases. Fig. 3 (b) shows that the difference between the ideal MMSE channel estimator and the proposed one increases, tending to a constant value as the transmit signal to noise ratio increases. This confirmation is in accordance with Remark 4.

The impact caused by the number of averaged elements on the performance of the proposed channel estimator is depicted in Fig. 4. In this simulation, the number of elements on the main diagonal of $\mathbf{Z}_{ik}^{H}\mathbf{Z}_{ik}$ that are averaged to calculate the proposed channel estimator is varied. The figure clearly indicates that the performance of the proposed channel estimator is directly affected by the number of elements taken into account for the calculation of the estimator. Consequently, the greater the number of averaged elements, the smaller the MSE difference between the proposed and the ideal MMSE channel estimators. Additionally, the number of elements of $\mathbf{Z}_{ik}^{H}\mathbf{Z}_{ik}$ considered in the average can also

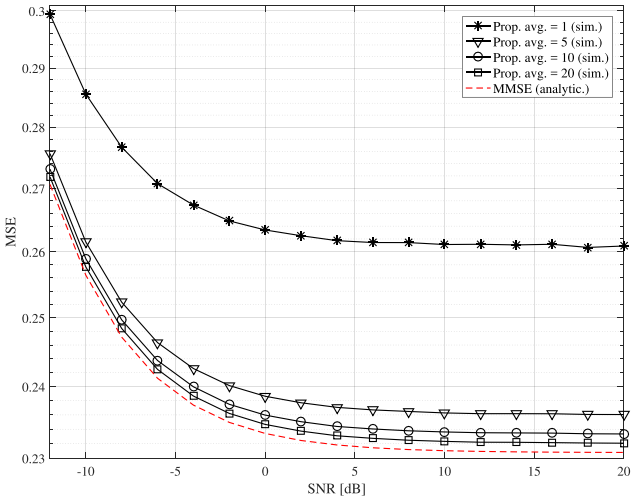

FIGURE 4. Performance of the proposed estimator when varying the number of averaged elements.

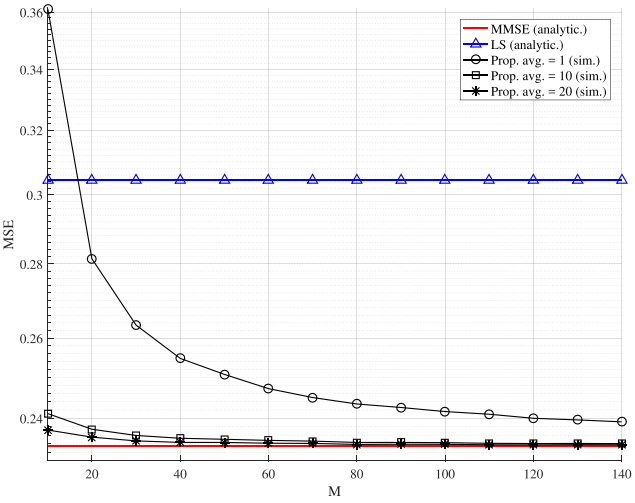

FIGURE 5. MSE of channel estimation versus number of antennas, M.

be thought of as the number of paths, P, a time-varying frequency selective fading channel has. The figure also indicates that the performance of the proposed channel estimator improves with P. Therefore, this result demonstrates that the proposed estimator not only exploits spatial diversity, but also takes advantage of the time diversity provided by the delay spread of the channels, since its performance improves with P.

Fig. 5 shows the impact of varying the number of antennas, M, installed at one BS on the channel estimation MSE. For this result, we considered $a = 0.05$ and Tx-SNR, $q = 0$ dB. The MSE of the proposed channel estimator asymptotically approaches that of the ideal MMSE channel estimator as the number of antennas, M, increases. On the other hand, the channel estimation MSE of the LS channel estimator stays fixed at a constant value for all values of M. In this figure, we also show the impact of the number of averaged elements of $\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}$ on the performance of the proposed channel estimator. From the results, it is observable

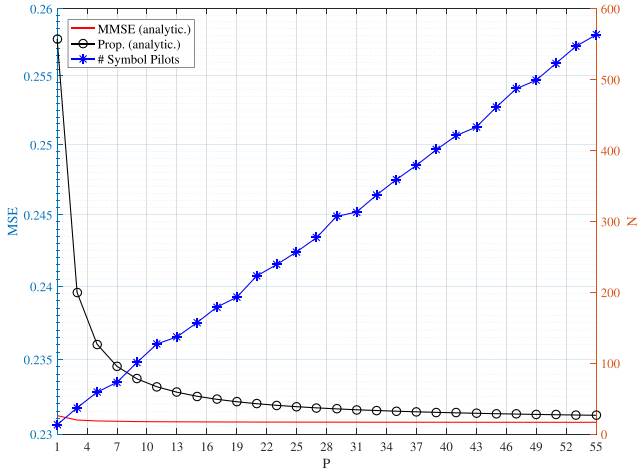

FIGURE 6. MSE of channel estimation versus number of channel paths versus pilot length.

that the proposed channel estimator has its MSE performance improved as the number of elements from the main diagonal of $\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}$ taken into account in the average increases, *i.e.*, as time diversity, represented here by P, increases.

Fig. 6 depicts the comparison of channel estimation MSE versus the number of channel paths, P, versus the number of pilot symbols, N, for both the ideal MMSE and proposed channel estimators. For this result, we considered the number of BS antennas, $M = 30$ and Tx-SNR, $q = 20$ dB. The figure shows that as the number of channel paths, P, increases, the MSE of the proposed channel estimator asymptotically approaches that of the ideal MMSE channel estimator. It is also shown that as P increases, the length of the pilot sequence, N, has to be also increased in order to provide users with orthogonal sequences. Without orthogonal sequences, the BS is not able to estimate the channels from its users to itself properly.

Fig. 7 shows the impact of different levels of inter-cell interference, here denoted by a, on the performance of channel estimation MSE for the three estimators considered in this work. For this result, we considered two different number of BS antennas, $M = 30$ and 100 and Tx-SNR, $q = 10$ dB. When compared with the proposed channel estimator, the LS estimator presents a slightly better MSE result for low levels of inter-cell interference and a small number of BS antennas, M, *e.g.*, $M = 30$. However, as M increases, this difference disappears. It can be clearly noticed in the plot with $M = 100$. It is clear from both plots that the proposed channel estimator substantially surpasses the LS estimator and approaches the performance of the ideal MMSE estimator as the inter-cell interference level increases. This observation is in accordance with Remark 4.

Fig. 8 compares the performance of channel estimation MSE for the case when the set of large-scale fading coefficients, $\{\beta_{ilk}\}$, is randomly selected. For this result, we considered the number of BS antennas, $M = 30$. The MSE results shown in the figure are the result of averaging the MSE values over 10000 realizations of the set $\{\beta_{ilk}\}$. As can

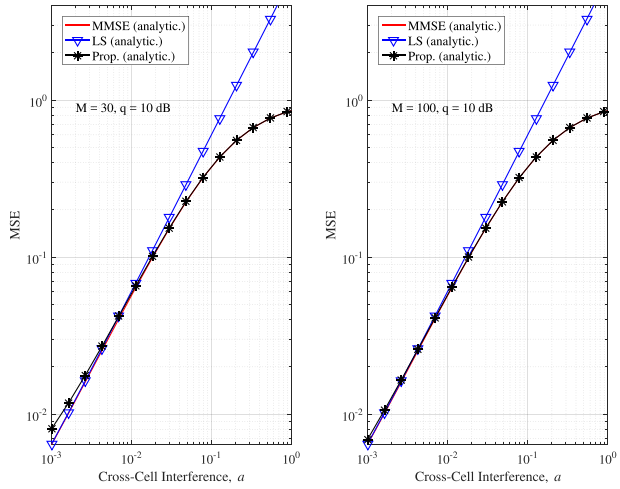

FIGURE 7. MSE of channel estimation versus inter-cell interference, a.

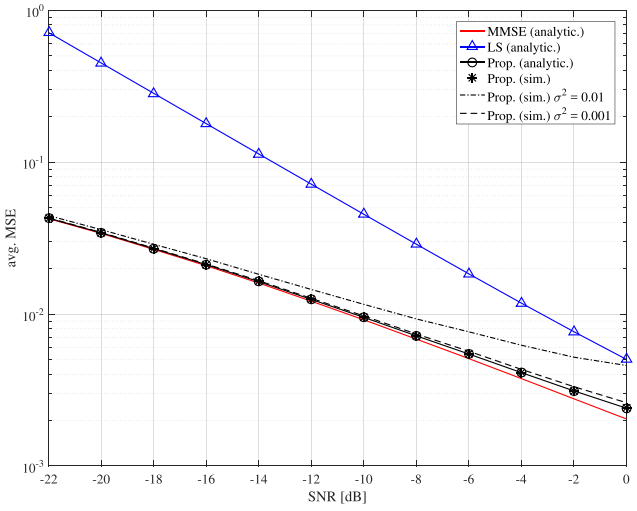

FIGURE 8. Averaged MSE of channel estimation for random$\{\beta_{ilk}\}$.

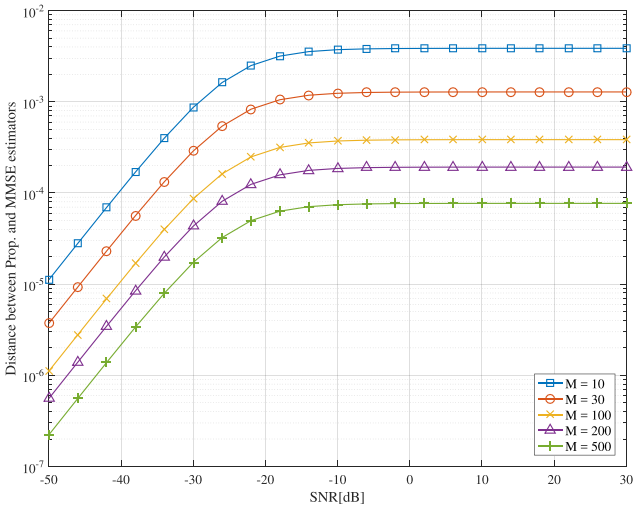

FIGURE 9. Distance between proposed and MMSE estimators (Remark 4).

be noticed, the approximated analytical MSE expression matches the simulated MSE results. Furthermore, we investigate the sensitivity of the proposed channel estimator against inaccuracy in the estimation of the intra-cell large-scale fading coefficients, β_{iik}, when we use the estimate $\beta_{iik} = \beta_{iik}(1 + \mathcal{N}(0, \sigma^2))$. The degradation in the performance of the proposed estimator becomes quite clear at high Tx-SNR values when $\sigma^2 = 0.01$, however, when $\sigma^2 = 0.001$ it becomes negligible. The performance of the proposed channel estimator still surpasses that of the LS estimator significantly.

In Fig. 9, we compare the distance between the proposed and MMSE channel estimators for different number of antennas, M, with $a = 0.05$. As Remark 4 states, the distance between them is small at low SNR, increasing with SNR until a ceiling value is reached. As can be also noticed, the ceiling value decreases as the number of antennas, M, increases. Also aligned with the interpretation of (14), the figure shows that increasing the pilot power does not improve the performance of the proposed estimator.

V. CONCLUSIONS

This work has proposed the use of Zadoff-Chu pilot sequences for the uplink training phase of massive MU-MIMO TDD cellular systems, in such a way that we could derive a simple and practical channel estimator for the respective frequency-selective massive MIMO fading channels, assuming pilot contamination. The estimator proposed in this work is defined by replacing the combined interference plus noise power term, ζ_{ik}, in the ideal MMSE estimator with a MVUE estimator for that term. The resulting channel estimator has proved to be simpler and more practical than the ideal MMSE channel estimator. Additionally, we derive and evaluate an approximate analytical MSE expression for the proposed channel estimator. Through analytical analysis and extensive simulations, the resulting channel estimator has proved to be simpler and more practical than the ideal MMSE estimator. Its MSE results asymptotically approaches that of the ideal MMSE estimator without any previous knowledge of the inter-cell large-scale coefficients and noise power term, ζ_{ik}. Finally, when $P = 1$, the simpler approximate analytical MSE expression presented here can be used instead of the more complex one presented in [18].

APPENDIX A

For the proof of the approximate analytical MSE expression of the proposed channel estimator we need to present a few Lemmas.

Lemma 1: If $X_m \sim \mathcal{CN}(0, \sigma^2)$ $\forall m$ are independent, then $\sum_{m=1}^{M} |X_m|^2 \sim \Gamma(M, \sigma^2)$.

Lemma 2: If $X \sim \Gamma(k, \theta)$ and $\frac{1}{X} \sim \Gamma^{-1}(k, \theta)$, *i.e.*, the Inverse-Gamma distribution, then $\mathbb{E}\left\{\frac{1}{X}\right\} = \frac{1}{\theta(k-1)}$.

Lemma 3: Let μ_X and μ_Y be the expectations of X and Y, σ_Y^2 be the variance of Y, and σ_{XY} be their covariance. Then the expectation, $\mathbb{E}\{X/Y\}$, can be approximated by

$$\mathbb{E}\left\{\frac{X}{Y}\right\} \approx \frac{\mu_X}{\mu_Y} - \frac{\sigma_{XY}}{\mu_Y^2} + \frac{\mu_X}{\mu_Y^3}\sigma_Y^2. \tag{15}$$

Proof: For a function that depends on two variables, x and y, the second order Taylor expansion series about the point (a, b) is given by

$$\begin{aligned} g(x,y) = {} & g(a,b) + g_x(a,b)(x-a) + g_y(a,b)(y-b) \\ & + \frac{1}{2!}(g_{xx}(a,b)(x-a)^2 + 2g_{xy}(a,b)(x-a)(y-b) \\ & + g_{yy}(a,b)(y-b)^2), \end{aligned} \tag{16}$$

where the subscripts denote the respective partial derivatives. The partial derivatives are defined by $g_y = -X/Y^2$, $g_{yy} = 2X/Y^3$, $g_x = 1/Y$, $g_{xx} = 0$, and $g_{xy} = -1/Y^2$. Applying the derivatives into (16), the second order Taylor expansion of $g(X, Y) = X/Y$ around the mean point (μ_X, μ_Y), the following is obtained

$$\begin{aligned} \frac{X}{Y} \approx {} & \frac{\mu_x}{\mu_y} - \frac{\mu_x}{\mu_y^2}(Y-\mu_y) + \frac{1}{\mu_y}(X-\mu_x) \\ & + \frac{1}{2!}\left(\frac{2\mu_x}{\mu_y^3}(Y-\mu_y)^2 - \frac{2}{\mu_y^2}(Y-\mu_y)(X-\mu_x)\right). \end{aligned} \tag{17}$$

Finally, applying the expectation operator, $\mathbb{E}\{.\}$, to (17) concludes the proof. ∎

A. PROOF OF THE APPROXIMATE MSE, η_{ik}^{prop}

With the purpose of helping the next proofs, the LS estimator can be expressed in the following way

$$\mathbf{Z}_{ik} = \begin{bmatrix} \overset{(0)}{z}_{ik1} & \overset{(1)}{z}_{ik1} & \cdots & \overset{(P-1)}{z}_{ik1} \\ \overset{(0)}{z}_{ik2} & \overset{(1)}{z}_{ik2} & \cdots & \overset{(P-1)}{z}_{ik2} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{(0)}{z}_{ikM} & \overset{(1)}{z}_{ikM} & \cdots & \overset{(P-1)}{z}_{ikM} \end{bmatrix}, \tag{18}$$

where $\overset{(p)}{z}_{ikm} = \sum_{l=1}^{L} \overset{(p)}{g}_{ilkm} + \overset{(p)}{w}_{ikm} \sim \mathcal{CN}(0, \zeta_{ik})\ \forall p$.

It is important to observe that the Trace of $\mathbf{Z}_{ik}^H\mathbf{Z}_{ik}$ results in a positive real-valued scalar random variable, $\|\mathbf{Z}_{ik}\|_F^2 = \sum_{m=1}^{M}\sum_{p=0}^{P-1} |\overset{(p)}{z}_{ikm}|^2$.

For the proof of the approximate MSE, we first expand η_{ik}^{prop} as

$$\begin{aligned} \eta_{ik}^{\text{prop}} = {} & \frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{prop}})^H\right]\right\} \\ & + \frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\mathbf{G}_{iik}(\mathbf{G}_{iik})^H\right]\right\} \\ & - \frac{2}{MP}\mathbb{E}\left\{\text{Tr}\left[\Re\left\{\hat{\mathbf{G}}_{iik}^{\text{prop}}(\mathbf{G}_{iik})^H\right\}\right]\right\}, \end{aligned} \tag{19}$$

and find these three expected values.

From (12), the first expectation can be written as

$$\begin{aligned} \frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{prop}})^H\right]\right\} &= MP\beta_{iik}^2\mathbb{E}\left\{\frac{\text{Tr}\left[\mathbf{Z}_{ik}\mathbf{Z}_{ik}^H\right]}{[\|\mathbf{Z}_{ik}\|_F^2]^2}\right\} \\ &= MP\beta_{iik}^2\mathbb{E}\left\{\frac{1}{\|\mathbf{Z}_{ik}\|_F^2}\right\}. \end{aligned} \tag{20}$$

From Lemma 1 we know that $\|\mathbf{Z}_{ik}\|_F^2 \sim \Gamma(MP, \zeta_{ik})$. Then, applying Lemma 2 to (20) we find that $\mathbb{E}\{1/\|\mathbf{Z}_{ik}\|_F^2\} = 1/\zeta_{ik}(MP-1)$ and consequently, the first expectation term is defined as

$$\frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{prop}})^H\right]\right\} = \frac{MP\beta_{iik}^2}{\zeta_{ik}(MP-1)}. \tag{21}$$

The second expectation term is defined as

$$\frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\mathbf{G}_{iik}(\mathbf{G}_{iik})^H\right]\right\} = \sum_{p=0}^{P-1}\sum_{m=1}^{M}\mathbb{E}\{|g_{iikm}^{(p)}|^2\} = \beta_{iik}. \tag{22}$$

Finally, in order to find the expected value of the third term, first we use (5) and (12) to rewrite it as

$$\begin{aligned} & -2\beta_{iik}\mathbb{E}\left\{\frac{\Re\left[\text{Tr}\left(\mathbf{Z}_{ik}\mathbf{G}_{iik}^H\right)\right]}{\|\mathbf{Z}_{ik}\|_F^2}\right\} \\ & = -2\beta_{iik}\left\{\mathbb{E}\left[\frac{\sum_{m=1}^{M}\sum_{p=0}^{P-1}\sum_{l=1}^{L}\overset{(p)}{g}_{ilkm}\overset{\star(p)}{g}_{iikm}}{\sum_{m=1}^{M}\sum_{p=0}^{P-1}|\overset{(p)}{z}_{ikm}|^2}\right]\right. \\ & \quad \left. + \mathbb{E}\left[\frac{\sum_{m=1}^{M}\sum_{p=0}^{P-1}\overset{(p)}{w}_{ikm}\overset{\star(p)}{g}_{iikm}}{\sum_{m=1}^{M}\sum_{p=0}^{P-1}|\overset{(p)}{z}_{ikm}|^2}\right]\right\}. \end{aligned} \tag{23}$$

The two ratios of random variables in (23) do not have a well-defined expectation, *i.e.* there is no closed formula for them. However, it is possible to approximate the moments of a function $g(X, Y)$ using Taylor series expansions, provided g is sufficiently differentiable and that the moments of X and Y are finite. Therefore, applying Lemma 3 to both terms in (23) we find an approximation to the third expectation

$$-\frac{2}{MP}\mathbb{E}\left\{\text{Tr}\left[\Re\left\{\hat{\mathbf{G}}_{iik}^{\text{prop}}(\mathbf{G}_{iik})^H\right\}\right]\right\} \approx -\frac{2\beta_{iik}^2}{\zeta_{ik}}. \tag{24}$$

We complete the proof by substituting (21), (22) and (24) back in the expansion of η_{ik}^{prop} defined by (19).

APPENDIX B

In this Appendix, we present proof for (14). First, we expand the Euclidean distance between $\hat{\mathbf{G}}_{iik}^{\text{prop}}$ and $\hat{\mathbf{G}}_{iik}^{\text{MMSE}}$ as

$$\begin{aligned} & \frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{prop}})^H\right]\right\} \\ & + \frac{1}{MP}\mathbb{E}\left\{\text{Tr}\left[\hat{\mathbf{G}}_{iik}^{\text{MMSE}}(\hat{\mathbf{G}}_{iik}^{\text{MMSE}})^H\right]\right\} \\ & - \frac{2}{MP}\mathbb{E}\left\{\text{Tr}\left[\Re\left\{\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{MMSE}})^H\right\}\right]\right\}. \end{aligned} \tag{25}$$

Then we compute these three different expectations. The first one is given by (21), $\frac{1}{MP}\mathbb{E}\{\text{Tr}[\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{prop}})^H]\} = MP\beta_{iik}^2/\zeta_{ik}(MP-1)$. Next, by recalling that $\hat{\mathbf{G}}_{iik}^{\text{MMSE}} \sim \mathcal{CN}(\mathbf{0}_M, \frac{\beta_{iik}^2}{\zeta_{ik}}\mathbf{I}_M)$ we have that $\frac{1}{MP}\mathbb{E}\{\text{Tr}[\hat{\mathbf{G}}_{iik}^{\text{MMSE}}(\hat{\mathbf{G}}_{iik}^{\text{MMSE}})^H]\} = \beta_{iik}^2/\zeta_{ik}$. For the last expectation term, using (9) and (12), we can write it as

$$\begin{aligned} & -\frac{2}{MP}\mathbb{E}\left\{\text{Tr}\left[\Re\left\{\hat{\mathbf{G}}_{iik}^{\text{prop}}(\hat{\mathbf{G}}_{iik}^{\text{MMSE}})^H\right\}\right]\right\} \\ & \qquad = -\frac{2\beta_{iik}^2}{\zeta_{ik}}\mathbb{E}\left\{\frac{\text{Tr}\ (\mathbf{Z}_{ik}\mathbf{Z}_{ik}^H)}{\text{Tr}\ (\mathbf{Z}_{ik}^H\mathbf{Z}_{ik})}\right\} = -\frac{2\beta_{iik}^2}{\zeta_{ik}}. \end{aligned} \tag{26}$$

Finally, by substituting these results back into the expansion in (25), we arrive at (14).

REFERENCES

[1] 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.

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

[3] 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.

[4] E. Björnson, E. G. Larsson, and T. L Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[5] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, Nov. 2016.

[6] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[7] J. Hoydis, S. ten 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.

[8] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[9] A. Ashikhmi, T. L. Marzetta, and L. Li. (2014). "Interference reduction in multi-cell massive MIMO systems I: Large-scale fading precoding and decoding." [Online]. Available: https://arxiv.org/abs/1411.4182

[10] K. Mawatwal, D. Sen, and R. Roy, "A semi-blind channel estimation algorithm for massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 70–73, Feb. 2017.

[11] R. T. Kobayashi and T. Abro, "Theoretical error for asynchronous multi-user large-scale MIMO channel estimation," *IET Commun.*, vol. 11, no. 1, pp. 17–24, Dec. 2016.

[12] J. Fang, X. Li, H. Li, and F. Gao, "Low-rank covariance-assisted downlink training and channel estimation for FDD massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1935–1947, Mar. 2017.

[13] M. Soltanalian, M. M. Naghsh, N. Shariati, P. Stoica, and B. Hassibi, "Training signal design for correlated massive MIMO channel estimation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1135–1143, Feb. 2017.

[14] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.

[15] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 787–801, Oct. 2014.

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

[17] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[18] A. Khansefid and H. Minn, "On channel estimation for massive MIMO with pilot contamination," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1660–1663, Sep. 2015.

[19] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Singapore: Pearson Ed., 2002.

[20] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. 40th Asilomar Conf. Signals, Syst., Comput. (ACSSC)*, Pacific Grove, CA, USA, Oct. 2006, pp. 359–363.

[21] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[22] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. I. Characterization," *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 136–146, Sep. 1997.

[23] W. H. Tranter, K. S. Shanmugan, T. S. Rappaport, and K. L. Kosbar, *Principles of Communication Systems Simulation with Wireless Applications*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2004.

[24] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 4, pp. 531–532, Jul. 1972.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1. New York, NY, USA: Pearson Ed., 1993.

[26] K.-F. Chen, Y.-C. Liu, and Y. T. Su, "On composite channel estimation in wireless massive MIMO systems," in *Proc. IEEE Globecom Workshops*, Sep. 2013, pp. 135–139.

FELIPE A. P. DE FIGUEIREDO received the B.S. and M.S. degrees in telecommunications from the Instituto Nacional de Telecomunicações, Minas Gerais, Brazil, in 2004 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Internet Technology and Data Science Laboratory, Ghent University, Gent, Belgium. He has been with the research and development of telecommunications systems for over ten years. His research interests include digital signal processing, digital communications, mobile communications, MIMO, multicarrier modulations, and FPGA development.

FABBRYCCIO A. C. M. CARDOSO received the bachelor's degree in electrical engineering from the University of Brasilia, Brazil, in 1995, and the M.Sc. and Ph.D. degrees from the State University of Campinas, Brazil, in 1998 and 2004, respectively. He is currently a Senior Researcher with the Wireless Communications Technologies Division, CPqD Foundation, Brazil. He has over seven years of experience leading research and development projects focused on physical layer design for 4G (LTE and LTE-A) and defense-related SDR systems. His research interests include 5G wireless communication technologies for radio access networks.

INGRID MOERMAN received the degree in electrical engineering and the Ph.D. degree from Ghent University in 1987 and 1992, respectively. She became a part-time Professor at Ghent University in 2000. She is currently a Staff Member with the IDLab, a core research group of imec with research activities embedded in Ghent University and the University of Antwerp. She is coordinating the research activities on mobile and wireless networking, and she is leading a research team of about 30 members at IDLab–Ghent University. Her main research interests include: Internet of Things, low power wide area networks, high-density wireless access networks, collaborative and cooperative networks, intelligent cognitive radio networks, real-time software defined radio, flexible hardware/software architectures for radio/network control and management, and experimentally-supported research.

GUSTAVO FRAIDENRAICH received the degree in electrical engineering from the Federal University of Pernambuco, Brazil, in 1997, and the M.Sc. and Ph.D. degrees from the State University of Campinas (UNICAMP), Brazil, in 2002 and 2006, respectively. From 2006 to 2008, he was a Post-Doctoral Fellow with the Star Laboratory Group, Stanford University, USA. He is currently an Assistant Professor with UNICAMP. His research interests include multiple antenna systems, cooperative systems, radar systems, and wireless communications in general.