



**UNIVERSIDADE ESTADUAL DE CAMPINAS
SISTEMA DE BIBLIOTECAS DA UNICAMP
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELLECTUAL DA UNICAMP**

Versão do arquivo anexado / Version of attached file:

Versão do Editor / Published Version

Mais informações no site da editora / Further information on publisher's website:

<https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1021-9>

DOI: 10.1186/s13638-018-1021-9

Direitos autorais / Publisher's copyright statement:

©2018 by Springer. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

RESEARCH

Open Access



Channel estimation for massive MIMO TDD systems assuming pilot contamination and flat fading

Felipe A. P. de Figueiredo^{1*} , Fabbryccio A. C. M. Cardoso², Ingrid Moerman¹ and Gustavo Fraidenraich³

Abstract

Channel estimation is crucial for massive multiple-input multiple-output (MIMO) systems to scale up multi-user (MU) MIMO, providing great improvement in spectral and energy efficiency. This paper presents a simple and practical channel estimator for multi-cell MU massive MIMO time division duplex (TDD) systems with pilot contamination in flat Rayleigh fading channels, i.e., the gains of the channels follow the Rayleigh distribution. We also assume uncorrelated antennas. The proposed estimator addresses performance under moderate to strong pilot contamination without previous knowledge of the cross-cell large-scale channel coefficients. This estimator performs asymptotically as well as the minimum mean square error (MMSE) estimator with respect to the number of antennas. An approximate analytical mean square error (MSE) expression is also derived for the proposed estimator.

Keywords: Massive MU-MIMO, Channel estimation, Flat fading, Pilot contamination, Maximum likelihood

1 Introduction

Massive multiple-input multiple-output (MIMO) antenna systems potentially allow base stations (BSs) to operate with huge improvements in spectral and radiated energy efficiency, using relatively low-complexity linear processing. The higher spectral efficiency is attained by serving several terminals in the same time-frequency resource through spatial multiplexing, and the increase in energy efficiency is mostly due to the array gain provided by the large set of antennas [1].

The expected massive MIMO improvements assume that accurate channel estimations are available at both the receiver and transmitter for detection and precoding, respectively. Additionally, the reuse of frequencies and pilot reference sequences in cellular communication systems causes interferences in channel estimation, degrading its performance. Since both the time-frequency resources allocated for pilot transmission and the channel coherence time are limited, the number of possible orthogonal pilot sequences is also limited, and as a

consequence, the pilot sequences have to be reused in neighbor cells of cellular systems. Therefore, channel estimates obtained in a given cell get contaminated by the pilots transmitted by the users in other cells [2]. This coherent interference is known in the literature as pilot contamination, i.e., the channel estimate at the base station in one cell becomes contaminated by the pilots of the users from other cells [3]. The contamination not only reduces the quality of the channel estimates, i.e., increases the MSE, but also makes the channel estimates statistically dependent, even though the true channels are statistically independent. Moreover, pilot contamination does not disappear with the addition of more antennas [4].

Massive MIMO systems operating in TDD 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 proportional to the number of terminals and improved estimation quality can be achieved due to the large number of antennas [5, 6]. Base stations estimate channels usually based on least squares (LS) [3] or minimum mean square error (MMSE) [7–9] methods. Besides, inter and intra-cell large-scale fading coefficients are assumed to be perfectly known when applying the MMSE method in the great majority of works [5, 9–13].

*Correspondence: felipe.pereira@ugent.be

¹Department of Information Technology, Ghent University, Technologiepark-Zwijnaarde, 15, 9052 Gent, Belgium

Full list of author information is available at the end of the article

In a real-world network deployment, although changing slowly, the large-scale fading coefficients must be estimated and updated from time to time. Additionally, the estimation error of the large-scale fading coefficients impacts significantly on the performance of uplink data decoding and downlink transmission (e.g., precoding and beamforming) [14–16]. Approaches on how to estimate the large-scale fading coefficients are presented in the following pieces of work [10, 14, 17].

The most commonly used analytical massive MIMO channel is the spatially i.i.d. frequency non-selective (flat) fading channel model. Flat fading channels are also known as *amplitude varying channels* and *narrowband channels* as the signal's bandwidth is narrow compared to channel's bandwidth [18]. In this narrowband channel model, the channel gain between any pair of transmit-receive antennas is modeled as a complex Gaussian random variable. This model relies on two assumptions: (i) the antenna elements in the transmitter and receiver being spatially well separated once the more widely spaced (in wavelengths) the antenna elements, the smaller the spatial channel correlation [19, 20], and (ii) the presence of a large number of temporally but narrowly separated multipaths (common in a *rich-scattering* environment), whose combined gain, by the central-limit theorem, can be approximated by a Gaussian random variable [20].

Flat fading channels present a channel response that exhibits flat gain and linear phase over a bandwidth (coherence bandwidth) that is greater than the signal's bandwidth. Therefore, all frequency components of the signal will experience the same magnitude of fading, resulting in a scalar channel response. The gain applied to the signal varies over time according to a fading distribution. In this work, we consider that the gain applied to the signal passing through this channel will vary randomly, according to a Rayleigh distribution. We additionally assume that the antenna spacing is sufficiently large so that the antennas are uncorrelated.

In this paper, we deal with the channel estimation and pilot contamination problems associated with uplink training in flat Rayleigh fading channels and understand its impact on the operation of multi-cell MU massive MIMO TDD cellular systems. We propose and evaluate an efficient and practical channel estimator that does not require previous knowledge of inter/intra-cell large-scale fading coefficients (i.e., interference) and noise power. Differently from [21], we employ the maximum likelihood (ML) method to find an estimator for the interference plus noise power term in the MMSE channel estimator. We show that this estimator is not only unbiased but also achieves the Crámer-Rao lower bound. We replace this estimator back into the MMSE estimator and prove that the performance of the new channel estimator asymptotically approaches that of the MMSE estimator. Simulation

results confirm that the performance of the proposed channel estimator approaches that of the ideal MMSE estimator asymptotically with the number M of antennas, i.e., $M \rightarrow \infty$. Additionally, in contrast with [21], we derive an approximate analytical MSE expression for the proposed channel estimator that is more mathematically tractable and not susceptible to numerical issues.

1.1 Related work

In this section, we survey previous work on channel estimation and pilot contamination mitigation.

A TDD cellular system employing BSs equipped with large numbers of antennas that communicate simultaneously with smaller numbers of cheap, single-antenna terminals through MU MIMO techniques is proposed in [3]. The author employs LS channel estimation in order to study and evaluate the problems caused by pilot contamination to such systems. He concludes that even when different sets of orthogonal pilots are used in different cells, it makes little difference to the resulting signal-to-interference ratio (SIR). This work is the first one to present the massive MIMO concept and identify its intrinsic issues, however, it fails to suggest ways to mitigate the pilot contamination problem.

The impact of pilot contamination on multi-cell systems is studied in [5]. The authors adopt MMSE channel estimation for the analysis of pilot contamination and the achievable rates in a massive MIMO system suffering from such problem. They propose a multi-cell MMSE-based precoding method that mitigates the pilot contamination problem by considering the set of training sequences assigned to the users in the solution of an optimization problem that minimizes the error seen by users in the serving cell and the interference seen by the users in all other cells. Simulation results show that the proposed approach has significant gains over certain single-cell precoding methods such as zero-forcing. In summary, the authors address the pilot contamination problem through a precoding technique and assume that the large-scale fading coefficients are known to all BSs.

MMSE channel estimation is used in [7] to derive approximations of the achievable uplink and downlink rates with several linear precoders and detectors for realistic system dimensions, i.e., systems where the number of antennas is not extremely large compared to the number of users. Simulation results show that the approximations are asymptotically tight, but accurate for realistic systems. The authors do not propose any approach to mitigate the pilot contamination problem, however, they study and evaluate its impact on the achievable rates.

The impact of pilot contamination effect on the achievable uplink ergodic rate when using linear detection in multi-cell MU massive MIMO systems under a more realistic physical channel model is assessed in [8]. The

authors assume that the channel vectors for different users are correlated, or not asymptotically orthogonal due to the antennas not being sufficiently well separated and/or the propagation environment not offering rich enough scattering. Moreover, they assume that the BS performs MMSE channel estimation based on training sequences received on the uplink and a priori knowledge of the large-scale fading coefficients.

In [9], the polynomial expansion (PE) technique is applied to channel estimation of massive MIMO systems in order to approximate the MMSE estimator and thereby obtain a new set of low-complexity channel estimators. Conventional MMSE estimators present cubic complexity due to an inversion operation while the estimator proposed in [9] reduces this to square complexity by approximating the inverse by a L -degree matrix polynomial. The proposed estimator achieves near-optimal MSE with low polynomial degrees. However, statistical knowledge of channel and disturbance parameters at the receiver is assumed in this paper.

Outer multi-cellular precoding is employed in [10] to devise a method used to eliminate pilot contamination in massive MIMO systems. Each BS performs two levels of precoding, firstly it estimates and shares only the large-scale fading coefficients with a central entity (network controller) which computes the precoding matrices and sends them back to the BSs, i.e., outer precoding. Next, each BS performs local precoding using estimates of the fast-fading vectors, i.e., inner precoding. The proposed approach is shown to completely mitigate the pilot contamination problem, making it possible to construct interference and noise free multi-cell massive MIMO systems with frequency reuse one and infinite downlink and uplink signal-to-interference-plus-noise ratios (SINRs). The proposed method employs MMSE channel estimation, however, the effectiveness of this method lies in the estimation accuracy of the shared large-scale fading coefficients from each BS. The authors also propose a method to estimate the large-scale fading coefficients. As this approach needs to share the large-scale coefficients with the network controller for outer precoding computation, it presents a higher computational complexity than non-cooperative approaches.

The authors in [11], adopt a massive MIMO system model that is based on spatially correlated channels. They devise a covariance aided channel estimation method which exploits the covariance information of both desired and interfering user channels. The Bayesian method is used to derive two different channel estimators (it is also shown that the Bayesian estimators coincide with the MMSE estimators), one for all channels from users in all cells to the target cell and the other one for the channels from users within the target cell. Results

show that in the ideal case, where the desired and the interference covariance matrices span distinct subspaces, the pilot contamination effect tends to vanish in the large antenna array case. As a consequence, users with mutually non-overlapping angle of arrival (AoA) hardly contaminate each other. Based on the results, the authors propose a coordinated pilot assignment strategy which assigns carefully selected groups of users to identical pilot sequences.

A semi-blind iterative space-alternating generalized expectation maximization (SAGE) based channel estimation algorithm for massive MIMO systems with pilot contamination is proposed in [13]. The proposed method does not assume a priori knowledge on the large-scale fading coefficients of the interfering cells, employing an estimate obtained from the received signal. The method updates the pilot based MMSE channel estimates iteratively with the help of the SAGE algorithm, which improves the initial estimate with the help of pilot symbols and soft information of the transmitted data. However, as it refines the channel estimates over some iterations starting from an initial MMSE channel estimation, it presents a computational complexity that is higher than the one presented by pure blind and linear estimators.

After surveying the literature on channel estimation and pilot contamination mentioned above, it is clear that, for clarity, in the great majority of studies the authors always assume complete knowledge on large-scale fading coefficients, i.e., path-loss and shadow fading, of the interfering cells, which is not the case in practical deployments of MU Massive MIMO systems. Furthermore, several studies propose solutions that present additional computational complexity in order to mitigate the pilot contamination problem.

The main contribution of our work is the proposal and assessment of a simple and practical channel estimator used to mitigate the pilot contamination problem. The proposed estimator does not assume a priori knowledge of the large-scale fading coefficients of the interfering cells. Moreover, it does not require the heavy overhead created by their estimation once it obtains them from the received signal.

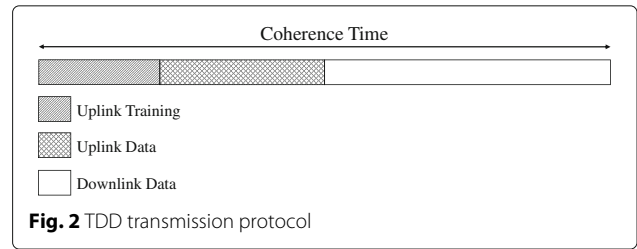
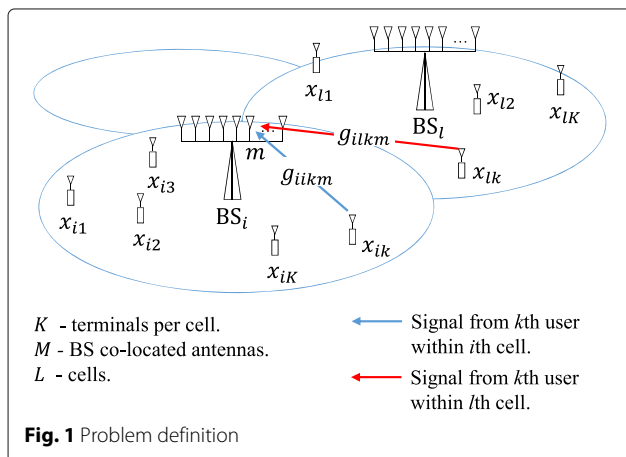
1.2 Organization

The remainder of this work is divided into four parts: First, we present the problem structure, signal model adopted for this study and briefly discuss two well-known channel estimators, namely, LS and MMSE linear estimators. Then, we introduce the proposed channel estimator for flat Rayleigh fading channels. Later, some numerical results are presented in order to support the effectiveness of the proposed estimator against the well-known linear estimators. Finally, we present our conclusions.

2 Problem structure

Let us assume as illustrated in Fig. 1 a multi-cell system with L cells, where each cell has a BS at its center with M co-located antenna elements and K randomly located single antenna users. Let us also assume Rayleigh fading channels being independent across users and antennas. Let g_{ilk} represent the complex gain of the channel from the k th user in the l th cell to the m th BS antenna in the i th cell. We can write $g_{ilk} = \sqrt{\beta_{ilk}}h_{ilk}$ where $\sqrt{\beta_{ilk}}$ is the large-scale coefficient encompassing both path loss and log-normal shadowing. We assume the same large-scale coefficient value for all BS co-located antennas, and h_{ilk} is the small-scale coefficient with a circularly symmetric complex normal distribution $\mathcal{CN}(0, 1)$. We assume that the large-scale fading coefficients do not depend on the frequency as well as on the antenna index m of a given BS because typically, the distance between a user and a BS is significantly larger than the distance between the BS antennas [10]. Therefore, between a BS and a user, there is only one large-scale fading coefficient. Moreover, these coefficients only change when a user considerably change its geographical location. The wireless channels are considered static during the channel coherence time (i.e., channel estimates are effective only in this time interval) and independent across users and antennas.

The $M \times 1$ channel vector from the k th user in the l th cell to the M antennas at the i th BS is defined by $\mathbf{g}_{ilk} = [g_{ilk1}, g_{ilk2}, \dots, g_{ilkM}]^T$. The overall $M \times K$ channel matrix \mathbf{G}_{il} is obtained by column concatenating vectors \mathbf{g}_{ilk} for all cell users, that is, $\mathbf{G}_{il} = [\mathbf{g}_{il1}, \mathbf{g}_{il2} \dots \mathbf{g}_{ilK}]$. For detection and precoding, BS i needs to know the channels of the users in cell i , namely $\{\mathbf{g}_{ilk}, \forall k\}$. The same way as in the literature, we treat $\{\beta_{ilk}\}$ as being deterministic during the channel estimation [1, 6, 8, 13]. As described, the overall channel matrix \mathbf{G}_{il} can also be defined directly by the channel coefficients,



$$\mathbf{G}_{il} = \begin{bmatrix} g_{il11} & g_{il21} & \dots & g_{ilK1} \\ g_{il12} & g_{il22} & \dots & g_{ilK2} \\ g_{il13} & g_{il23} & \dots & g_{ilK3} \\ \vdots & \vdots & \ddots & \vdots \\ g_{il1M} & g_{il2M} & \dots & g_{ilKM} \end{bmatrix}. \quad (1)$$

Based on the assumption of channel reciprocity, we adopt the TDD protocol depicted in Fig. 2 and proposed in [22]. Due to the reciprocity principle, only the uplink channels need to be estimated while the downlink channels are equal to the transpose of the uplink channels. It is important to note that the length of the TDD frames is limited by the channel coherence time [22, 23]. According to the 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 uplink data and generate precoding matrices used to transmit downlink data.

2.1 Uplink training

Each user transmits an uplink training sequence so that the user serving BS can estimate the channels per antenna and subsequently detect the transmitted user data. We assume that users in different cells transmit data at the same time-frequency resource (a typical scenario in massive MIMO) and that the pilot reuse factor is one, the worst possible use case scenario [3]. As all BSs reuse the same set of pilots and transmit at the same time-frequency resource, the pilot contamination problem arises, consequently, all the other BSs will also receive the pilots sent by users being served by other BSs, limiting the quality of the channel estimation [24].

The pilot signals of K users are represented by a $N \times K$ matrix \mathbf{S} of the form $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_K]$, where N is the length of the pilot sequences. Each pilot sequence is of the form $\mathbf{s}_k = [s_k^0, s_k^1, \dots, s_k^{N-1}]^T$. The pilot matrix, \mathbf{S} , exhibits orthogonal property $\mathbf{S}^H \mathbf{S} = N \mathbf{I}_K$.

The pilots are created by applying cyclic shifts to Zadoff-Chu (ZC) root sequences with length N , where N is a prime number. These sequences exhibit some useful properties: (i) cyclically shifted versions of themselves are orthogonal to each other, (ii) constant amplitude, (iii) zero

auto-correlation, (iv) flat frequency domain response, and (v) cross-correlation between two ZC sequences is low [25]. Some of the reasons why they are adopted in communication systems like long-term evolution (LTE) are (i) channel estimation at receiver is made simpler due to their small variation in frequency, (ii) inter-cell interference is reduced as they present low cross-correlation, (iii) high peak to average power ratio (PAPR) is reduced due to their small variation in time. ZC sequences are used in this work due to the properties mentioned above [25]; however, any other sequences could be used as long as they exhibit the required orthogonal property. Additionally, we assume that $N \geq K$ in order to avoid underdetermined systems.

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

$$\mathbf{Y}_i = \sqrt{q} \sum_{l=1}^L \mathbf{G}_{il} \mathbf{S}^H + \mathbf{N}_i, \quad (2)$$

where q is the uplink power or transmit signal to noise ratio (TX SNR) and \mathbf{N}_i is a $M \times N$ noise matrix with independent and identically distributed elements following $\mathcal{CN}(0, 1)$.

Equation (2) can also be written as showed below, which clearly highlights the coherent inter-cell interference caused by users employing the same pilot sequences in other BSs.

$$\mathbf{Y}_i = \underbrace{\sqrt{q} \mathbf{G}_{ii} \mathbf{S}^H}_{\text{Desired pilot signals}} + \underbrace{\sqrt{q} \sum_{l=1, l \neq i}^L \mathbf{G}_{il} \mathbf{S}^H}_{\text{Undesired pilot signals}} + \underbrace{\mathbf{N}_i}_{\text{Noise}}. \quad (3)$$

2.2 LS channel estimator

For estimation of the channel \mathbf{g}_{ilk} at BS i , a sufficient statistic [26–28] is given by

$$\begin{aligned} \mathbf{z}_{ik} &= \frac{1}{\sqrt{qN}} \mathbf{Y}_i \mathbf{s}_k = \sum_{l=1}^L \mathbf{g}_{ilk} + \frac{\mathbf{N}_i \mathbf{s}_k}{\sqrt{qN}} \\ &= \underbrace{\mathbf{g}_{ilk}}_{\text{Desired channel}} + \underbrace{\sum_{l=1, l \neq i}^L \mathbf{g}_{ilk}}_{\text{Inter-cell interference}} + \underbrace{\frac{\mathbf{N}_i \mathbf{s}_k}{\sqrt{qN}}}_{\text{Noise}}. \end{aligned} \quad (4)$$

where \mathbf{z}_{ik} is a column vector with a $\mathcal{CN}(\mathbf{0}_M, \zeta_{ik} \mathbf{I}_M)$ distribution and

$$\zeta_{ik} = \sum_{l=1}^L \beta_{ilk} + \frac{1}{qN}. \quad (5)$$

Additionally, the term corresponding to noise in (4) has a $\mathcal{CN}(\mathbf{0}_M, \frac{1}{qN} \mathbf{I}_M)$ distribution.

Therefore, the least square estimator is given by [26]

$$\hat{\mathbf{g}}_{ilk}^{\text{LS}} = \mathbf{z}_{ik}. \quad (6)$$

The MSE per antenna of the LS estimator is given by

$$\eta_{ik}^{\text{LS}} = \frac{1}{M} \mathbb{E} \left[\left\| \hat{\mathbf{g}}_{ilk}^{\text{LS}} - \mathbf{g}_{ilk} \right\|^2 \right] = \zeta_{ik} - \beta_{ilk}. \quad (7)$$

As known, the LS estimator has larger MSE than the MMSE estimator; however, it does not need prior knowledge of the large-scale fading coefficients, $\{\beta_{ilk}\}$.

Remark 1 Due to pilot contamination, as $q \rightarrow \infty$, $\eta_{ik}^{\text{LS}} \rightarrow \sum_{l=1, l \neq i}^L \beta_{ilk}$.

2.3 MMSE channel estimator

A great number of massive MIMO works adopt the MMSE estimation method to obtain channel knowledge [5, 8]. Those works assume that all large-scale fading coefficients, i.e., $\{\beta_{ilk}, i \geq 1, l \leq L, 1 \leq k \leq K\}$, are perfectly known. In practice, this assumption might not be reasonable. In case we consider the coefficients $\{\beta_{ilk}\}$ perfectly known at the BS, the ideal MMSE estimator is given by [26]

$$\hat{\mathbf{g}}_{ilk}^{\text{MMSE}} = \frac{\beta_{ilk}}{\zeta_{ik}} \mathbf{z}_{ik}, \quad (8)$$

where $\hat{\mathbf{g}}_{ilk}^{\text{MMSE}} \sim \mathcal{CN}(\mathbf{0}_M, \frac{\beta_{ilk}^2}{\zeta_{ik}} \mathbf{I}_M)$ and the MSE of MMSE estimator is given by

$$\eta_{ik}^{\text{MMSE}} = \frac{1}{M} \mathbb{E} \left[\left\| \hat{\mathbf{g}}_{ilk}^{\text{MMSE}} - \mathbf{g}_{ilk} \right\|^2 \right] = \beta_{ilk} \left(1 - \frac{\beta_{ilk}}{\zeta_{ik}} \right). \quad (9)$$

Remark 2 Due to pilot contamination, as $q \rightarrow \infty$, $\eta_{ik}^{\text{mmse}} \rightarrow \beta_{ilk} \left(1 - \frac{\beta_{ilk}}{\sum_{l=1}^L \beta_{ilk}} \right)$.

3 Proposed channel estimator

In this work, we employ the ML method to estimate the parameter ζ_{ik} [26]. Applying the ML method to $f(\mathbf{z}_{ik}; \zeta_{ik}) \sim \mathcal{CN}(\mathbf{0}_M, \zeta_{ik} \mathbf{I}_M)$, we find the following estimator for ζ_{ik} given the observation \mathbf{z}_{ik}

$$\hat{\zeta}_{ik} = \frac{\|\mathbf{z}_{ik}\|^2}{M}. \quad (10)$$

This estimator has $\mathbb{E}[\hat{\zeta}_{ik}] = \zeta_{ik}$, which shows that the ML estimator is unbiased, and $\text{var}\{\hat{\zeta}_{ik}\} = \zeta_{ik}^2/M$. In order to assess the efficiency of the estimator we derive the Cramér-Rao bound as [26]

$$\text{var}(\hat{\zeta}_{ik}) \geq \frac{\zeta_{ik}^2}{M}. \quad (11)$$

Therefore, the ML estimator derived for ζ_{ik} is the minimum variance unbiased estimator (MVUE), i.e., it is an unbiased estimator that has lower variance than any other unbiased estimator for all possible values of the parameter [26].

This simple and effective estimator is derived based on the observation that the MMSE estimator does not need to know the individual large-scale fading coefficients, $\{\beta_{iik}\}$, as assumed in the existing literature, but just ζ_{ik} suffices. The proposed estimator for ζ_{ik} makes the acquisition of inter-cell large-scale fading coefficients unnecessary. The task of gaining knowledge of those coefficients may be unjustifiable in practice due to the excessive, e.g., in case there are L cells serving K users in each one of them, each BS needs to acquire $(L - 1)K$ inter-cell large-scale coefficients.

Swapping ζ_{ik} with $\hat{\zeta}_{ik}$ in (8) produces the proposed channel estimator, which is defined by

$$\hat{\mathbf{g}}_{iik}^{\text{prop}} = M\beta_{iik} \frac{\mathbf{z}_{ik}}{\|\mathbf{z}_{ik}\|^2}. \quad (12)$$

This estimator approaches the ideal MMSE estimator asymptotically with respect to M . The estimator has $\mathbb{E}[\hat{\mathbf{g}}_{iik}^{\text{prop}}] = \mathbf{0}_M$ and variance given by

$$\text{Var}[\hat{\mathbf{g}}_{iik}^{\text{prop}}] = \mathbb{E}[\hat{\mathbf{g}}_{iik}^{\text{prop}} (\hat{\mathbf{g}}_{iik}^{\text{prop}})^H] = \left(\frac{M^2}{M-1} \frac{\beta_{iik}^2}{\zeta_{ik}} \right) \mathbf{I}_M. \quad (13)$$

As can be seen by analyzing equation (13), as $M \rightarrow \infty$, $\text{Var}[\hat{\mathbf{g}}_{iik}^{\text{prop}}] \rightarrow \frac{\beta_{iik}^2}{\zeta_{ik}}$. An approximation to the MSE per antenna of this estimator is given by

$$\eta_{ik}^{\text{prop}} = \frac{1}{M} \mathbb{E}[\|\hat{\mathbf{g}}_{iik}^{\text{prop}} - \mathbf{g}_{iik}\|^2] \approx \beta_{iik} \left[1 - \frac{(M-2)\beta_{iik}}{(M-1)\zeta_{ik}} \right]. \quad (14)$$

The approximate MSE in (14) for the proposed estimator decreases with increasing transmitting power q , increasing M or decreasing β_{iik} , which means smaller interference level from other cells, i.e., smaller pilot contamination.

Remark 3 Due to pilot contamination, as $q \rightarrow \infty$ and $M \rightarrow \infty$, $\eta_{ik}^{\text{prop}} \rightarrow \beta_{iik} \left(1 - \frac{\beta_{iik}}{\sum_{l=1}^L \beta_{iik}} \right)$.

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

For the sake of clarity, we reproduce below the closed-form MSE equation (9) presented in [21].

$$\eta_{ik}^{\text{prop(closed-form)}} = \frac{M}{M-1} \frac{\beta_{iik}^2}{\zeta_{ik}} + \beta_{iik} - 2\beta_{iik}\theta_{ik} \quad (15)$$

where

$$\theta_{ik} = \int_0^1 \int_{-1}^1 \frac{k_{ik}^2(1-t) + k_{ik}w\sqrt{t(1-t)}}{k_{ik}^2(1-t) + 2k_{ik}w\sqrt{t(1-t)} + t} f_T(t)f_W(w)dw dt \quad (16)$$

with $k_{ik} = \sqrt{\frac{\beta_{iik}}{\zeta_{ik} - \beta_{iik}}}$, and $f_T(t)$ and $f_W(w)$ are given by

$$f_T(t) = \frac{\Gamma(2M)}{(\Gamma(M))^2} (t(1-t))^{M-1}, 0 < t < 1 \quad (17)$$

$$f_W(w) = \frac{M}{\pi} B\left(\frac{1}{2}, M\right) (1-w^2)^{M-\frac{1}{2}}, |w| < 1. \quad (18)$$

The difference between the closed-form, given by Eq. (9) in [21], and the approximated MSE expressions are defined by

$$\eta_{ik}^{\text{prop(closed-form)}} - \eta_{ik}^{\text{prop(approx.)}} = 2\beta_{iik} \left\{ \frac{\beta_{iik}}{\zeta_{ik}} - \theta_{ik} \right\}, \quad (19)$$

where θ_{ik} is defined in [21].

Remark 4 As both q and $M \rightarrow \infty$, $\theta_{ik} \rightarrow \frac{\beta_{iik}}{\zeta_{ik}}$ and then, $\eta_{ik}^{\text{prop(closed-form)}} - \eta_{ik}^{\text{prop(approx.)}} \rightarrow 0$.

We find Remark 4 by using Remark 3 and equating the closed-form and approximated MSE expressions. This remark shows that the difference between the closed-form and the approximated MSE expressions decreases, tending to 0, as both uplink power, q , and number of receiving antennas, M , increase.

Remark 5 The average normalized squared Euclidean distance between $\hat{\mathbf{g}}_{iik}^{\text{prop}}$ and $\hat{\mathbf{g}}_{iik}^{\text{MMSE}}$ is given by

$$\frac{1}{M} \mathbb{E}[\|\hat{\mathbf{g}}_{iik}^{\text{prop}} - \hat{\mathbf{g}}_{iik}^{\text{MMSE}}\|^2] = \frac{1}{M-1} \frac{\beta_{iik}^2}{\zeta_{ik}}. \quad (20)$$

The proof of (20) is given in Appendix B. From (5) and (20), it is easily noticeable that the average distance decreases with increasing M , decreasing q , increasing β_{iik} , $i \neq l$, and decreasing β_{iik} .

4 Numerical results and discussion

In this section, we compare the performance of the proposed channel estimator with that of the MMSE and LS estimators. We adopt a typical multi-cell structure as the one shown in Fig. 1 with $L = 7$ cells (one central cell surrounded by 6 other cells), $K = 10$ users in each cell, frequency reuse factor of 1 and $N = K$ pilot symbols. We consider two different types of setups for $\{\beta_{iik}\}$, one with fixed values and other with random values. For the fixed case, we set $\beta_{iik} = 1$ and $\beta_{ilk} = a, \forall l \neq i$, where a represents the cross-cell interference level. The value selected for a in the fixed case is 0.05, and it is chosen so that there is moderate cross-cell interference level from users being

served by other BSs, i.e., not being served by the central cell. For the random case, users in each cell are uniformly distributed within a ring with radii $d_0 = 100$ m and $d_1 = 1000$ m respectively. The large-scale fading coefficients $\{\beta_{ilk}\}$ are independently generated by $\beta_{ilk} = \psi / \left(\frac{d_{ilk}}{d_0}\right)^\nu$, where $\nu = 3.8$, $10 \log_{10}(\psi) \sim \mathcal{N}(0, \sigma_{\text{shadow, dB}}^2)$ with $\sigma_{\text{shadow, dB}} = 8$, and d_{ilk} is the distance of the k th user in the l th cell to the i th BS. Both, the path loss exponent, ν , and the standard deviation of the log-normal shadowing, $\sigma_{\text{shadow, dB}}$, are common values for outdoor shadowed urban cellular radio environments [18, 29].

The results in Fig. 3 show MSE versus SNR (uplink pilot power q) performances for $a = 0.05$ and $M = 70$. As can be seen, analytical, approximated, and simulation MSEs match for all estimators. With the increase of SNR, MSEs of all the estimation methods decrease. There are MSE floors for all the three estimators due to pilot contamination (see Remarks 1, 2, and 3). At low SNR, the MSE of the proposed estimator is very close to that of the ideal MMSE estimator. On the other hand, as can be noticed, with the increase of the SNR, the gap between the ideal MMSE estimator and the proposed one increases (see Remark 5).

In Fig. 4, we compare MSE versus the number of BS antennas M under the setting of $a = 0.05$ and TX SNR $q = 10$ dB. With the increase of M , the MSE of the proposed estimator approaches that of the ideal MMSE, while the MSE of LS estimator does not change. Due to numerical issues, the closed-form MSE expression presented in [21] does not produce values for $M > 85$. During our simulations, comparing the closed-form expression given by equation (15) and the approximated MSE expression given by (14), we noticed that the $\Gamma(2M)$ function in the numerator of equation (16) grows without bound, reaching values that are greater than the largest possible

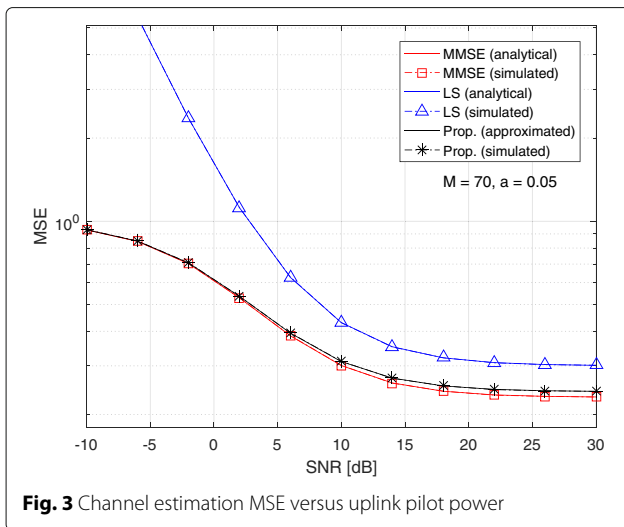


Fig. 3 Channel estimation MSE versus uplink pilot power

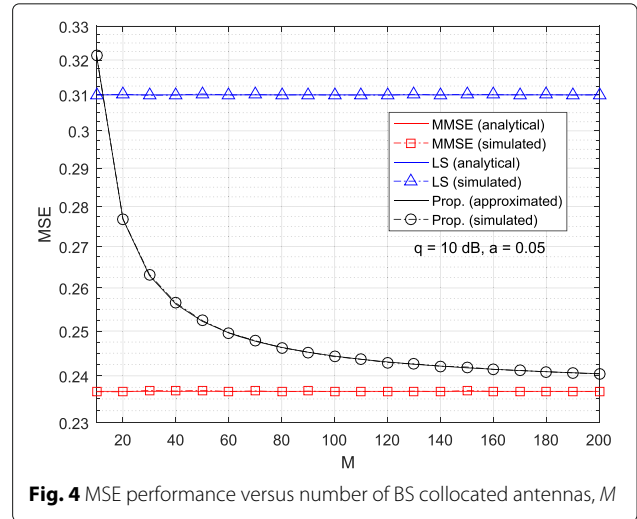


Fig. 4 MSE performance versus number of BS collocated antennas, M

finite floating-point number represented by the IEEE double precision format, i.e., $1.7977e+308$ [30], for values of M greater than 85. A double precision variable goes to $+\text{Inf}$ after the largest possible number [30]. On the other hand, as can be seen in Fig. 4, the approximate analytical MSE expression (14) does not present the same problem and, therefore, can be used to evaluate the MSE for any number of antennas, M without any numerical issue.

In Fig. 5, we compare MSE performance with respect to various levels of cross-cell interference, a , with $q = 10$ dB and two different number of antennas, $M = 30$ and $M = 90$. We can see that when a increases (the effect of pilot contamination increases), the estimation performance degrades. At a low cross-cell interference level, LS presents a slightly better MSE when compared to the proposed estimator. This difference disappears as M increases, as can be noticed in the plot with $M = 90$. As

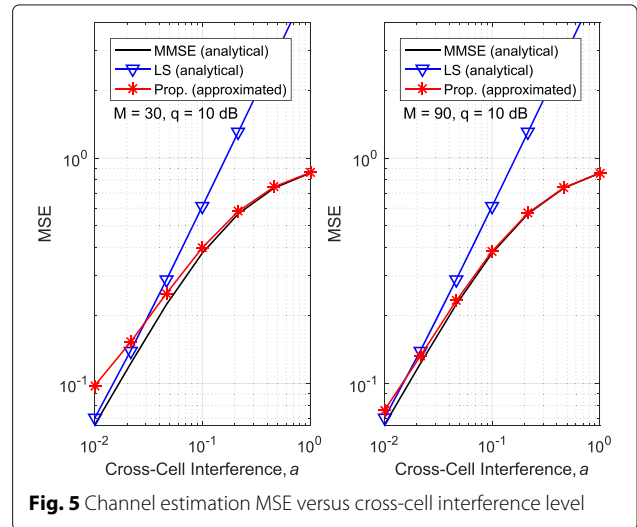


Fig. 5 Channel estimation MSE versus cross-cell interference level

the interference level increases, the proposed method outperforms the LS estimator substantially and approaches the ideal MMSE performance (see Remark 5).

In Fig. 6, we evaluate the MSE performance under random large-scale fading coefficients $\{\beta_{ilk}\}$ with $M = 30$. The results are obtained by averaging MSEs over 10000 realizations of $\{\beta_{ilk}\}$. As can be observed, simulation MSE matches with the analytical MSE. Additionally, the sensitivity of the proposed estimator against inaccuracy of β_{ilk} by using an estimate $\hat{\beta}_{ilk} = \beta_{ilk} (1 + \mathcal{N}(0, \sigma^2))$ is investigated. The performance degradation for $\sigma^2 = 0.1$ is noticeable at high SNR but for $\sigma^2 = 0.01$, it is insignificant. The proposed estimator still outperforms the LS estimator significantly.

In Fig. 7, we compare the distance between the proposed and MMSE channel estimators for different number of antennas, M , with $a = 0.05$. As the Remark 5 states, the distance is small at low SNR, increasing with SNR until a ceiling is reached. As can be also noticed, the ceiling value decreases with the number of antennas, M .

In Fig. 8, we compare the absolute distance between the approximated MSE expression presented in (13) and the analytical (closed form) MSE expression presented in [21] for various SNR and M values with $a = 0.05$. The distance between the MSE expressions is small at low SNR, increasing with SNR until a ceiling value is reached. As can be noticed, the ceiling value decreases with the number of antennas, M . For $M = 50$, the ceiling distance is smaller than $1e - 4$, showing that the approximated MSE expression can replace the one presented in [21].

5 Conclusions

In this work, we have introduced a simple and practical channel estimator for massive MIMO TDD systems

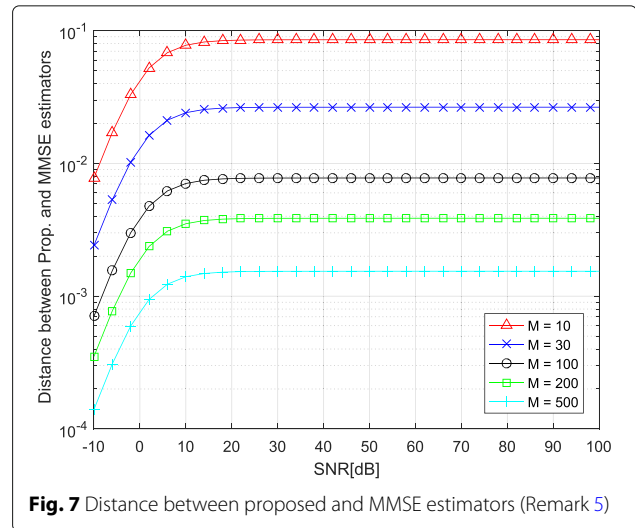


Fig. 7 Distance between proposed and MMSE estimators (Remark 5)

with pilot contamination in a flat channel environment. The proposed estimator replaces the combined interference plus noise power term in the ideal MMSE estimator with a maximum likelihood estimator for that term. Moreover, the proposed estimator presents MSE results that are very close to that of the ideal MMSE estimator without requiring previous knowledge of noise and interference statistics. Additionally, we have derived an approximate analytical MSE expression for the proposed estimator which can be useful in system design and performance evaluation. We have also shown that the MSE expression presented here asymptotically approaches that of the MMSE estimator. Finally, the simpler approximate analytical MSE expression presented here can be used instead of the more complex and susceptible to numerical issues one presented in [21].

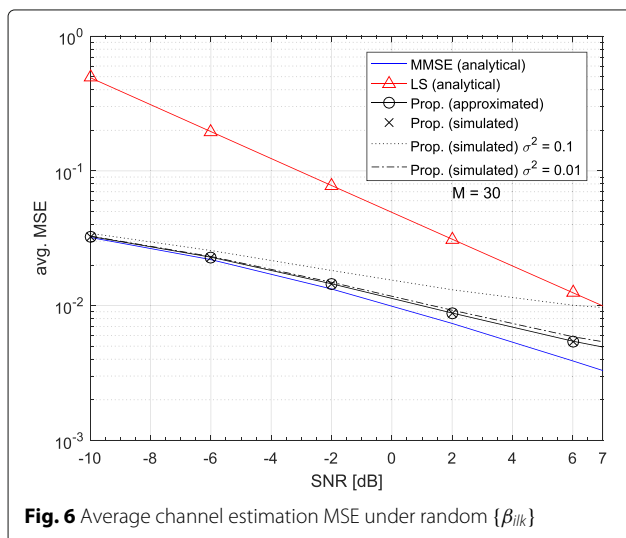


Fig. 6 Average channel estimation MSE under random $\{\beta_{ilk}\}$

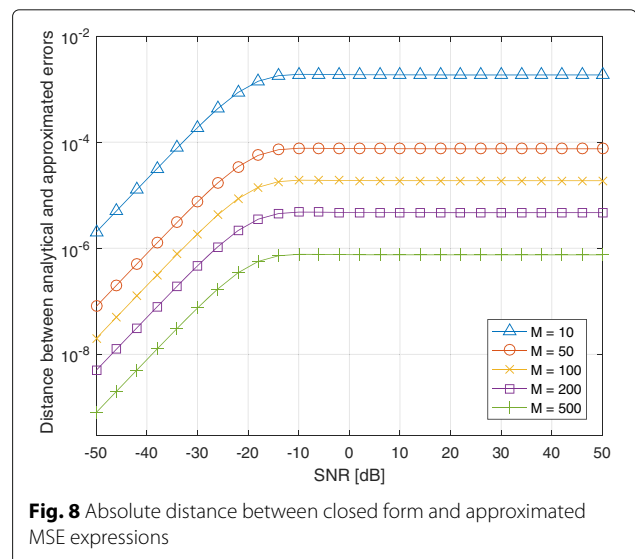


Fig. 8 Absolute distance between closed form and approximated MSE expressions

Appendix A

For the proof of the approximate MSE of the proposed 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*

$$\mathbb{E}\left\{\frac{X}{Y}\right\} \approx \frac{\mu_X}{\mu_Y} - \frac{\sigma_{XY}}{\mu_Y^2} + \frac{\mu_X}{\mu_Y^3} \sigma_Y^2. \quad (21)$$

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

$$\begin{aligned} g(x, y) &= g(a, b) + g_x(a, b)(x - a) + g_y(a, b)(y - b) \\ &\quad + \frac{1}{2!} (g_{xx}(a, b)(x - a)^2 + 2g_{xy}(a, b)(x - a)(y - b) \\ &\quad + g_{yy}(a, b)(y - b)^2), \end{aligned} \quad (22)$$

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 (22), the second order Taylor expansion of $g(X, Y) = X/Y$ around the mean point (μ_X, μ_Y) , the following is obtained

$$\begin{aligned} \frac{X}{Y} &\approx \frac{\mu_x}{\mu_y} - \frac{\mu_x}{\mu_y^2} (Y - \mu_y) + \frac{1}{\mu_y} (X - \mu_x) \\ &\quad + \frac{1}{2!} \left(\frac{2\mu_x}{\mu_y^3} (Y - \mu_y)^2 - \frac{2}{\mu_y^2} (Y - \mu_y)(X - \mu_x) \right). \end{aligned} \quad (23)$$

Finally, applying the expectation operator, $\mathbb{E}\{\cdot\}$, to (23) concludes the proof. \square

Proof of the approximate MSE, η_{ik}^{prop}

For the proof of the approximate MSE, we expand it as

$$\begin{aligned} \eta_{ik}^{\text{prop}} &= \frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{prop}}\|^2] + \frac{1}{M} \mathbb{E} [\|\mathbf{g}_{ik}\|^2] \\ &\quad - \frac{2}{M} \mathbb{E} \left[\Re \left[(\hat{\mathbf{g}}_{ik}^{\text{prop}})^H \mathbf{g}_{ik} \right] \right], \end{aligned} \quad (24)$$

and find these three expectations.

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

$$\begin{aligned} \frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{prop}}\|^2] &= M\beta_{ik}^2 \mathbb{E} \left\{ \frac{\|\mathbf{z}_{ik}\|^2}{[\|\mathbf{z}_{ik}\|^2]^2} \right\} \\ &= M\beta_{ik}^2 \mathbb{E} \left\{ \frac{1}{\|\mathbf{z}_{ik}\|^2} \right\}. \end{aligned} \quad (25)$$

From Lemma 1, we know that $\|\mathbf{z}_{ik}\|^2 \sim \Gamma(MP, \zeta_{ik})$. Then, applying Lemma 2 to (25), we figure out that $\mathbb{E} \left\{ 1/\|\mathbf{z}_{ik}\|^2 \right\} = 1/\zeta_{ik}(M-1)$ and consequently, the first expectation term is defined as

$$\frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{prop}}\|^2] = \frac{M\beta_{ik}^2}{\zeta_{ik}(M-1)}. \quad (26)$$

The second expectation term is defined as

$$\frac{1}{M} \mathbb{E} [\|\mathbf{g}_{ik}\|^2] = \frac{1}{M} \sum_{m=1}^M \mathbb{E} [\|\mathbf{g}_{ikm}\|^2] = \beta_{ik}. \quad (27)$$

Finally, in order to find the expected value of the third term, first, we use (4) and (12) to rewrite it as

$$\begin{aligned} -2\beta_{ik} \mathbb{E} \left\{ \Re \left[\frac{\mathbf{z}_{ik}^H \mathbf{g}_{ik}}{\|\mathbf{z}_{ik}\|^2} \right] \right\} &= -2\beta_{ik} \left\{ \mathbb{E} \left[\Re \left[\frac{\sum_{l=1}^L \mathbf{g}_{ik}^H \mathbf{g}_{ik}}{\|\mathbf{z}_{ik}\|^2} \right] \right] \right. \\ &\quad \left. + \mathbb{E} \left[\Re \left[\frac{\mathbf{w}_{ik}^H \mathbf{g}_{ik}}{\|\mathbf{z}_{ik}\|^2} \right] \right] \right\} \end{aligned} \quad (28)$$

where $\mathbf{w}_{ik} = \mathbf{N}_i \mathbf{s}_k / \sqrt{q}N \sim \mathcal{CN}(\mathbf{0}_M, \frac{1}{qN} \mathbf{I}_M)$.

In order to avoid the numerical issues mentioned earlier in this work and find a simpler and more tractable equation for the MSE of the proposed channel estimator, we find approximations to the two ratios of random variables in (28). 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 separately to each one of the terms in the second and third lines of (28), we are able to find an approximation to the third expectation, which is defined as

$$\begin{aligned} -\frac{2}{M} \mathbb{E} \left[\Re \left[(\hat{\mathbf{g}}_{ik}^{\text{prop}})^H \mathbf{g}_{ik} \right] \right] &\approx -2\beta_{ik} \left\{ \frac{\beta_{ik}}{\zeta_{ik}} \left(1 - \frac{\sum_{l=1}^L \beta_{ik}}{M\zeta_{ik}} + \frac{1}{M} \right) \right. \\ &\quad \left. + \left[\frac{-\beta_{ik}}{M\zeta_{ik}^2 qN} \right] \right\} = -\frac{2\beta_{ik}^2}{\zeta_{ik}}. \end{aligned} \quad (29)$$

After finding the three expectations, (26), (27), and (29), by substituting them back in the expansion of η_{ik}^{prop} , we complete the proof.

Appendix B

Here, we present proof for (20). First, we expand the normalized Euclidean distance between $\hat{\mathbf{g}}_{ik}^{\text{prop}}$ and $\hat{\mathbf{g}}_{ik}^{\text{MMSE}}$ as

$$\begin{aligned} \frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{prop}}\|^2] &+ \frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{MMSE}}\|^2] \\ &- \frac{2}{M} \mathbb{E} \left[\Re \left[(\hat{\mathbf{g}}_{ik}^{\text{prop}})^H \hat{\mathbf{g}}_{ik}^{\text{MMSE}} \right] \right]. \end{aligned} \quad (30)$$

Then, we compute these three different expectations. The first one is given by (26), $\frac{1}{M} \mathbb{E} [\|\hat{\mathbf{g}}_{ik}^{\text{prop}}\|^2] = M\beta_{ik}^2/\zeta_{ik}(M-1)$. Next, by recalling that $\hat{\mathbf{g}}_{ik}^{\text{MMSE}} \sim$

$\mathcal{CN}\left(\mathbf{0}_M, \frac{\beta_{iik}^2}{\zeta_{ik}} \mathbf{I}_M\right)$, we have that $\frac{1}{M} \mathbb{E} \left[\|\hat{\mathbf{g}}_{iik}^{\text{MMSE}}\|^2 \right] = \beta_{iik}^2 / \zeta_{ik}$. For the last expectation term, using (8) and (12), we can write it as

$$-\frac{2}{M} \mathbb{E} \left[\Re \left[\left(\hat{\mathbf{g}}_{iik}^{\text{prop}} \right)^H \hat{\mathbf{g}}_{iik}^{\text{MMSE}} \right] \right] = -\frac{2\beta_{iik}^2}{\zeta_{ik}} \mathbb{E} \left\{ \Re \left[\frac{\|\mathbf{z}_{ik}\|^2}{\|\mathbf{z}_{ik}\|^2} \right] \right\} = -\frac{2\beta_{iik}^2}{\zeta_{ik}}. \tag{31}$$

Finally, by substituting these results back into the expansion, we arrive at (20).

Abbreviations

AoA: Angle of arrival; BS: Base station; IEEE: Institute of Electrical and Electronics Engineers; LS: Least squares; LTE: Long-term evolution; MIMO: Multiple-input multiple-output; ML: Maximum likelihood; MMSE: Minimum mean square error; MSE: Mean square error; MVUE: Minimum variance unbiased estimator; MU-MIMO: Multi user multiple-input multiple-output; PAPR: High peak to average power ratio; SAGE: Space-alternating generalized expectation; SINR: Signal-to-interference plus noise ratio; SIR: Signal-to-interference ratio; SNR: Signal-to-noise ratio; TDD: Time division duplex; ZC: Zadoff-Chu

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

FAPdF is the main author and is responsible for the conception, simulation, interpretation of data and wrote the paper. FACMC revised equations, helped writing the introduction, and critically revised the paper. IM critically revised the paper and proofread it. GF supervised the research and approved the version to be published. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Information Technology, Ghent University, Technologiepark-Zwijnaarde, 15, 9052 Gent, Belgium. ²Research and Development Center on Telecommunications, (CPqD), Doutor Ricardo Benetton Martins Street, 1000, 13086-902 Campinas, Brazil. ³Department of Communications, School of Electrical and Computer Engineering, University of Campinas, Albert Einstein Avenue, 400, 13083-852 Campinas, Brazil.

Received: 26 June 2017 Accepted: 29 December 2017

Published online: 15 January 2018

References

1. EG Larsson, O Edfors, F Tufvesson, TL Marzetta, Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**(2) (2014)
2. O Elijah, CY Leow, TA Rahman, S Nunoo, SZ Iliya, A comprehensive survey of pilot contamination in massive MIMO—5G system. *IEEE Commun. Surv. Tutor.* **18**(2), 905–923 (2016)
3. TL Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wirel. Commun.* **9**(11), 3590–3600 (2010)
4. E Björnson, J Hoydis, L Sanguinetti, Massive MIMO networks: spectral, energy, and hardware efficiency. *Found. Trends Signal Process.* **11**(3–4), 154–655 (2017). <https://doi.org/10.1561/2000000093>
5. J Jose, A Ashikhmin, TL Marzetta, S Vishwanath, Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans. Wirel. Commun.* **10**(8), 2640–2651 (2011)

6. E Björnson, EG Larsson, TL Marzetta, Massive MIMO: ten myths and one critical question. *IEEE Commun. Mag.* **54**(2) (2016)
7. J Hoydis, S ten Brink, M Debbah, Massive MIMO in the UL/DL of cellular networks: how many antennas do we need? *IEEE J.Sel. Areas Commun.* **31**(2), 160–171 (2013)
8. HQ Ngo, EG Larsson, TL Marzetta, The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel. *IEEE Trans. Commun.* **61**(6), 2350–2361 (2013)
9. N Shariati, E Björnson, M Bengtsson, M Debb, Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics. *IEEE J. Sel. Top. Signal Process.* **8**(5), 815–830 (2014)
10. A Ashikhmi, TL Marzetta, L Li, Interference reduction in multi-cell massive MIMO systems I: large-scale fading precoding and decoding. *IEEE Trans. Inf. Theory* (2014). To appear. [Online]. Available: arxiv.org/abs/1411.4182
11. H Yin, D Gesbert, M Filippou, Y Liu, A coordinated approach to channel estimation in large-scale multiple antenna systems. *IEEE J. Sel. Areas Commun.* **31**(2), 264–273 (2013)
12. TL Marzetta, EG Larsson, H Yang, HQ Ngo, *Fundamentals of massive MIMO*. (Cambridge University Press, 2016)
13. K Mawatwal, D Sen, R Roy, A semi-blind channel estimation algorithm for massive MIMO systems. *IEEE Wirel. Commun. Lett.* **6**(1), 70–73 (2017)
14. K Li, X Song, OM Ahmad, MNS Swamy, An improved multicell MMSE channel estimation in a massive MIMO system. *Int. J. Antennas Propag.* **2014** (2014)
15. A Yang, Z He, C Xing, Z Fei, J Kuang, The role of large-scale fading in uplink massive MIMO systems. *IEEE Trans. Veh. Technol.* **65**(1), 477–483 (2016)
16. H Quoc, Ngo, Larsson, EG, TL Marzetta, Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. Commun.* **61**(4) (2013)
17. K-F Chen, Y-C Liu, YT Su, in *IEEE Globecom Workshops*. On composite channel estimation in wireless massive MIMO systems, (2013), pp. 135–139
18. TS Rappaport, *Wireless Communications Principles and Practice*, 2nd ed. (Pearson Education Inc., 2002)
19. E Björnson, E Jorswieck, B Ottersten, Impact of spatial correlation and precoding design in OSTBC MIMO systems. *IEEE Trans. Wirel. Commun.* **9**, 3578–3589 (2010)
20. A Chockalingam, BS Rajan, *Large MIMO Systems*. (Cambridge Univ. Press, Cambridge, 2013)
21. A Khansefid, H Minn, On channel estimation for massive MIMO with pilot contamination. *IEEE Commun. Lett.* **19**(9), 1660–1663 (2015)
22. TL Marzetta, in *Proc. 40th Asilomar Conf. Signals, Syst., Comput. (ACSSC)*. How much training is required for multiuser MIMO? (Pacific Grove, 2006), pp. 359–363
23. G Caire, N Jindal, M Kobayashi, N Ravindran, Multiuser MIMO achievable rates with downlink training and channel state feedback. *IEEE Trans. Inf. Theory.* **56**(6), 2845–2866 (2010)
24. E Björnson, EG Larsson, M Debbah, Massive MIMO for maximal spectral efficiency: how many users and pilots should be allocated? *IEEE Trans. Wirel. Commun.* **15**(2), 1293–1308 (2016)
25. DC Chu, Polyphase codes with good periodic correlation properties. *IEEE Trans. Inf. Theory.* **19**, 531–532 (1972)
26. SM Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. (Pearson Education, New York, 1993)
27. C Radhakrishna Rao, *Linear Statistical Inference and Its Applications*, 2nd. ed. (Wiley, New York, 1973)
28. RA Fisher, On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. A.* **222**, 309–368 (1922)
29. SR Saunders, A Aragon-Zavala, *Antennas and propagation for wireless communication systems*, 2nd edition. (Wiley, Chichester, 2007)
30. IEEE 754-2008, *IEEE 754-2008: Standard for Floating-Point Arithmetic*. (Institute of Electrical and Electronics Engineers, August, 2008)