On Channel Estimation for Massive MIMO with Pilot Contamination and Multipath Fading Channels

Felipe A. P. de Figueiredo*, Fabiano S. Mathilde*, Fabricio P. Santos*,
Fabbryccio A. C. M. Cardoso* and Gustavo Fraidenraich†
*CPqD – Research and Development Center on Telecommunications, Brazil.
†DECOM/FEEC – State University of Campinas (UNICAMP), Brazil.
Email: *[felipep, fabianom, fpsantos, fcardoso]@cpqd.com.br, †gf@decom.fee.unicamp.br

Abstract— **This paper introduces a simple and practical channel estimator for multipath multi-cell massive MIMO TDD systems with pilot contamination. Least squares (LS) channel estimators present low performance under moderate to strong pilot contamination, while minimum mean square error (MMSE) estimators perform much better than LS estimators in such conditions. However, MMSE estimators present an issue that prevents its adoption in practical systems, namely the assumption there is perfect knowledge of the cross-cell large scale coefficients. The proposed estimator does not present any of the issues mentioned above and additionally provides performance that improves as the number of paths of the channels increases, reaching that of the MMSE estimator.**

Index Terms— **Massive MU-MIMO, channel estimation, pilot contamination, multipath.**

I. Introduction

The employment of large numbers of antennas at base stations in cellular networks can lead to extremely high gains in both, spectral and energy efficiency. The higher spectral efficiency is attained by serving several users simultaneously through spatial multiplexing and the increase in energy efficiency is mostly due to the array gain provided by the large set of antennas.

Faithful channel estimation is of fundamental importance to communications systems. Without optimum channel estimation such systems can not attain its expected performance.

For communications systems that employ Massive MIMO technology, channel estimation is of utmost importance once both precoding and decoding strongly depend on optimum channel estimation. In cellular massive MIMO systems, the reuse of frequency by other cells causes interference in channel estimation, degrading it. This problem is known in the literature as pilot contamination [1]. Time division duplexing (TDD) is usually considered in the literature due to pilot overhead cost [2]. When TDD is employed, users transmit pilots in the uplink and a base station (BS) estimates the channels normally based on least squares (LS) [1] or minimum mean square error (MMSE) [3, 4] methods. In the great majority of works out there inter and intracell large-scale fading coefficients are assumed to be perfectly known when applying the MMSE method [5].

978-1-5090-5137-3/16/$31.00

A myriad of works in the literature tackles the problem of channel estimation only for flat fading channels [2]–[11], which sometimes does not reflect the true nature of communications channels in practice.

Given the extensive body of literature in flat-fading channel estimation for massive MIMO TDD systems, the main contribution of this paper is in understanding multipath multi-cell massive MIMO TDD systems with channel training. Specifically, the main contribution is to demonstrate the pilot contamination problem associated with uplink training in multipath fading channels, understand its impact on the operation of multi-cell massive MIMO TDD cellular systems, and propose a practical channel estimator to mitigate the problems caused by pilot contamination.

The present work analyses the performance of some linear estimators and proposes an efficient and practical one that does not require knowledge of inter-cell large-scale fading coefficients for the case when the communications channels present multipath fading behavior, $i.e.$, frequency-selective fading. It is an extension of the work presented in [11] to multipath fading channels. The proposed approach instead of individually estimate the large-scale coefficients estimates a parameter that is the sum of them plus a normalized noise variance. That parameter is substituted back into the MMSE estimator. Simulation results show that the performance of the proposed channel estimation method approaches that of the ideal MMSE estimator asymptotically, $i.e.$, $M \rightarrow \infty$.

This work is divided into four parts: First, it presents the signal model adopted for this study. Then, it introduces the proposed channel estimator for multipath channels. Next, some numerical results are presented in order to support the effectiveness of the proposed estimator against the well know estimators, namely LS and MMSE. Finally, we present some conclusions and give the directions for future work.

II. Signal Model

We consider a multi-cell system with L cells where each cell has a BS with co-located M antennas elements and K randomly located single antenna users. We assume multipath Rayleigh fading channels which are considered as being static within a frame and independent across users and antennas. Let $g_{ilkm}^{(p)}$ represent the complex gain of the pth path of the channel from user k in cell l to the antenna m of the BS in cell i. We can write $g_{ilkm}^{(p)} = \sqrt{\beta_{ilk}} h_{ilkm}^{(p)}$ where $\sqrt{\beta_{ilk}}$ (the same coefficient for all antennas, *i.e.*, m) is the large-scale coefficient (for both path loss and log-normal shadowing) and $h_{ilkm}^{(p)}$ is the small-scale coefficient with a $\mathcal{CN}(0,1)$ distribution.

The multipath channels are modeled as $1 \times P$ vectors, $\mathbf{g}_{ilkm} = \left[g_{ilkm}^{(0)}, g_{ilkm}^{(1)}, ..., g_{ilkm}^{(P-1)}\right]$ where P is the number of paths (or coefficients) of the channel. The overall $M \times KP$ channel matrix is denoted by $\mathbf{G}_{il}$ as showed below. Each column of $\mathbf{G}_{il}$ is denoted by $\mathbf{g}_{ilk}$ and represents the channels from the kth user to the M antennas at the ith BS in the lth cell. For detection and precoding, BS i needs to know the channels of the users in cell i, namely $\{\mathbf{g}_{iik} : \forall k\}$.

$$\begin{aligned}\mathbf{G}_{il} &= \begin{bmatrix} \mathbf{g}_{il11} & \mathbf{g}_{il21} & \cdots & \mathbf{g}_{ilK1} \\ \mathbf{g}_{il12} & \mathbf{g}_{il22} & \cdots & \mathbf{g}_{ilK2} \\ \mathbf{g}_{il13} & \mathbf{g}_{il23} & \cdots & \mathbf{g}_{ilK3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{g}_{il1M} & \mathbf{g}_{il2M} & \cdots & \mathbf{g}_{ilKM} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{g}_{il1} & \mathbf{g}_{il2} & \cdots & \mathbf{g}_{ilK} \end{bmatrix} \\ &= \begin{bmatrix} \beta_{il1}\mathbf{h}_{il1} & \beta_{il2}\mathbf{h}_{il2} & \cdots & \beta_{ilK}\mathbf{h}_{ilK} \end{bmatrix}.\end{aligned} \tag{1}$$

The same way as in the literature, we treat $\{\beta_{ilk}\}$ as being deterministic during the channel estimation [1, 5, 6, 11].

A. Uplink Training

Each user transmits a pilot sequence so that the BS can estimate the M multipath channels from that user to it. We assume users of different cells transmit the same set of pilots at the same time (a typical scenario in massive MIMO) and the pilot reuse factor is of one. The multipath channels are modeled as Finite Response Filters (FIR), this way, the received pilot sequences (symbols) at the BS are represented in matrix form by a $M \times N$ matrix $\mathbf{Y}_i$ as

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

where N is the length of the pilot sequence, 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 i.i.d. elements of $\mathcal{CN}(0,1)$.

The pilot signals of K users are represented by a $N \times KP$ matrix $\mathbf{S}$ of the form $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_K]$. As we are dealing with multipath channels, each of the pilot sequences is convoluted with each one of the M channels. Therefore, in order to express convolution in matrix form we write each of the pilots as a $N \times P$ matrix with orthogonality property $\mathbf{S}_k^H \mathbf{S}_k = N\mathbf{I}_p$, where $\mathbf{S}_k$ is of the form

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

In order to obtain circular matrices as showed in (3) users must transmit its pilot sequences twice making it a $2N$ sequence. This is direct consequence of the orthogonality property mentioned above. The pilots are created by applying cyclic shifts to Zadoff-Chu sequences with length N, where N is a prime number. These sequences exhibits the useful property that cyclically shifted versions of themselves are orthogonal to one another [12]. We assume that $N > KP$, this assumption must hold true so that all users have valid pilots, *i.e.*, it is possible to construct orthogonal pilot sequences from the same Zadoff-Chu sequence for every user in the cell.

B. LS Channel Estimator

As showed earlier, $\mathbf{S}_k$ denotes the kth pilot matrix of $\mathbf{S}$, therefore, for estimation of the channel $\mathbf{g}_{ilk}$ at BS i, a sufficient statistic is given by

$$\mathbf{z}_{ik} = \frac{1}{\sqrt{q}N} \mathbf{Y}_i \mathbf{S}_k = \sum_{l=1}^{L} \mathbf{g}_{ilk} + \frac{\mathbf{N}_i \mathbf{S}_k}{\sqrt{q}N}. \tag{4}$$

where $\mathbf{z}_{ik}$ is a $M \times P$ matrix with a $\mathcal{CN}(\mathbf{0}_{M \times P}, M\zeta_{ik}\mathbf{I}_P)$ distribution where

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

while the noise part of (4) has a $\mathcal{CN}(\mathbf{0}_{M \times P}, \frac{M}{qN}\mathbf{I}_P)$ distribution.

The least squares estimator is given by [13]

$$\hat{\mathbf{g}}_{iik}^{\text{ls}} = \mathbf{z}_{ik}. \tag{6}$$

The MSE per antenna of the LS estimator is given by

$$\eta_{ik}^{\text{ls}} = \frac{1}{M} \mathbb{E}\left\{ \|\hat{\mathbf{g}}_{iik}^{\text{ls}} - \mathbf{g}_{iik}\|^2 \right\} = (\zeta_{ik} - \beta_{iik})\mathbf{I}_P. \tag{7}$$

As known, the LS estimator has larger MSE than the MMSE estimator, however, it does not need to know the large-scale fading coefficients, β_{ilk}.

Differently from [11], the MSE equation for the LS estimator does not result in a scalar value, instead, it results in an identity matrix, $\mathbf{I}_P$, which is multiplied by the same MSE found in that work. This result is due to the P coefficients of the multipath channels. When $P = 1$ equation (9) reduces to equation (5) in [11].

Remark 1: Due to pilot contamination, as $q \to \infty$, $\eta_{ik}^{\text{ls}} \to \left(\sum_{l=1, l \neq i}^{L} \beta_{ilk}\right) \mathbf{I}_P$.

C. MMSE Channel Estimator

The Bayesian MMSE estimator requires knowledge of statistics of parameters to be estimated as well as those of noise and interference. Many massive MIMO works acquire channel knowledge based on MMSE estimation [2, 4], and they assume perfect knowledge of all large-scale fading coefficients, *i.e.*, $\{\beta_{ilk}, 1 \leq i, l \leq L, 1 \leq k \leq K\}$ which may not be justifiable in practice. With the perfect knowledge of $\{\beta_{ilk}\}$, the ideal MMSE estimator is given by [13]

$$\hat{\mathbf{g}}_{iik}^{\text{mmse}} = \frac{\beta_{iik}}{\zeta_{ik}} \mathbf{z}_{ik}. \tag{8}$$

The MSE per antenna of the ideal MMSE channel estimator is given by [13]

$$\eta_{ik}^{\text{mmse}} = \frac{1}{M}\mathbb{E}\left\{\|\hat{\mathbf{g}}_{iik}^{\text{mmse}} - \mathbf{g}_{iik}\|^2\right\} = \left\{\frac{\beta_{iik}(\zeta_{ik} - \beta_{iik})}{\zeta_{ik}}\right\}\mathbf{I}_P. \tag{9}$$

The MSE per antenna, η_{ik}^{mmse}, decreases with increasing q, decreasing β_{iik}, or decreasing β_{iik} (smaller interference level). Again, the MSE per antenna for the MMSE estimator only differs from the findings in [11] by the identity matrix.

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

III. Proposed Channel Estimator

Acquiring knowledge of inter-cell large-scale fading coefficients may be unjustifiable in practice due to the excessive overhead it would require. For instance, for L cells with K users in each cell, each BS needs to estimate $(L-1)K$ inter-cell large-scale fading coefficients. A simple but effective solution to this issue is based on the observation that what we need is not individual β_{ilk} as assumed in the existing works, but just ζ_{ik}. Thus, our approach is to estimate ζ_{ik} and replace it in the MMSE estimator (8).

The minimum variance unbiased estimator (MVUE) of ζ_{ik} given the observed signal $\mathbf{z}_{ik}$ is given by [13]

$$\hat{\zeta_{ik}} = \frac{\text{Tr}\left(\|\mathbf{z}_{ik}\|^2\right)}{MP}. \tag{10}$$

The $\text{Tr}(.)$ operation removes all inner products between estimated channels of the kth user to the M antennas at the ith BS and as P increases, the term

$$\frac{\text{Tr}\left(\|\mathbf{z}_{ik}\|^2\right)}{P}, \tag{11}$$

in (10) tends to $M\zeta_{ik}$, therefore resulting in a better MVUE estimator. As can be noticed, equation (11) is the average of the P inner products in the main diagonal of $\|\mathbf{z}_{ik}\|^2$.

Equation (10) comes from the fact that

$$\mathbb{E}\left\{\|\mathbf{z}_{ik}\|^2\right\} = M\zeta_{ik}\mathbf{I}_P \tag{12}$$

and that the unbiased estimator has

$$\mathbb{E}\{\hat{\zeta_{ik}}\} = \zeta_{ik}. \tag{13}$$

Therefore, inserting (10) into (8) produces the proposed channel estimator, which is defined by

$$\hat{\mathbf{g}}_{iik}^{\text{prop}} = MP\frac{\beta_{iik}}{\text{Tr}\left(\|\mathbf{z}_{ik}\|^2\right)}\mathbf{z}_{ik}. \tag{14}$$

Note that when $P = 1$ equation (14) simplifies to equation (8) of [11]. As can be also noticed, as P increases the proposed estimator approaches the ideal MMSE estimator once there are more values to average over.

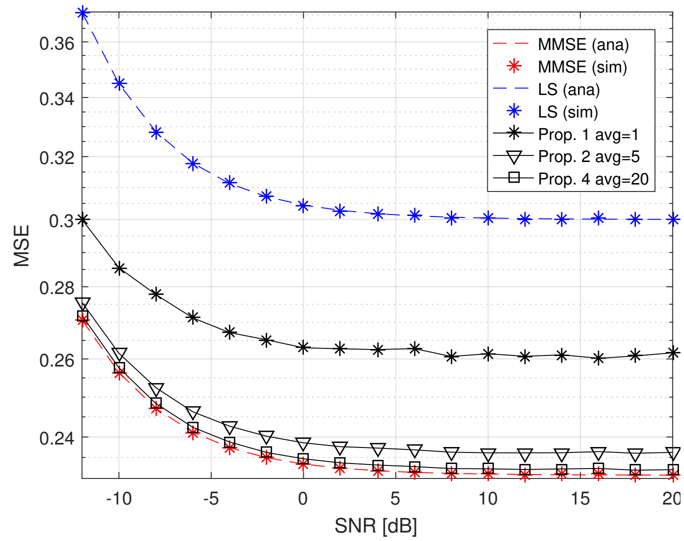

Fig. 1. Channel Estimation MSE versus uplink pilot power.

IV. Numerical Results and Discussion

In this section we compare the performance of the proposed channel estimator with the ideal MMSE estimator and LS estimator. We use a typical multicell structure with $L = 7$ cells, $K = 10$ users in each cell, frequency reuse factor of 1, number of coefficients of the multipath channel $P = 20$, and $N = 223$ pilot symbols. We consider a setup with fixed values for $\{\beta_{ilk}\}$. For the fixed case, we set $\beta_{iik} = 1$ and $\beta_{ilk} = a, \forall\, l \neq i$.

Fig. 1 shows the channel estimation MSE versus uplink pilot power for $a = 0.05$ and $M = 30$. As can be seen, analytical and simulation MSEs match for LS and MSE estimators. With the increase of the pilot power, MSEs of all the methods decrease. There are MSE floors for all three estimators due to pilot contamination (see remarks 1 and 2). At low uplink pilot power, MSE of the proposed estimator is very close to that of the ideal MMSE estimator while the gap between LS

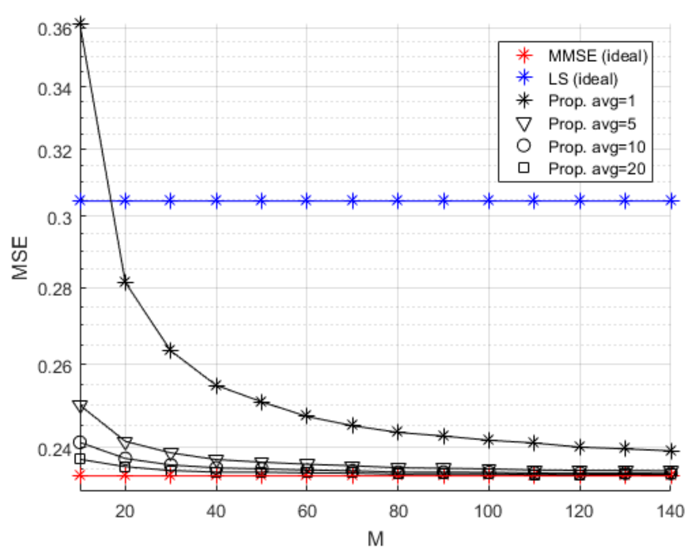

Fig. 2. Channel estimation MSE versus the number of receiving antennas.

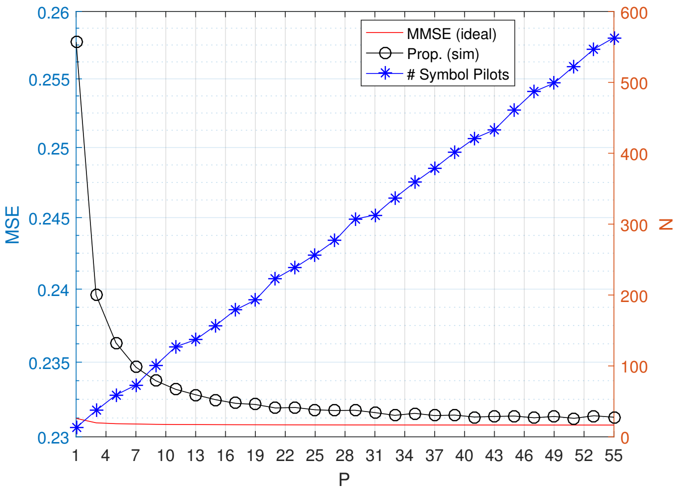

Fig. 3. Channel estimation MSE versus the number of channel coefficients versus number of pilot symbols.

and ideal MMSE estimators is large. In this simulation we vary the number of elements of the main diagonal of $\|\mathbf{z}_{ik}\|^2$ that are averaged to calculate the proposed estimator. The figure shows that the number of elements taken into account directly affects the performance of the proposed estimator, the greater the number of averaged elements, the smaller the MSE gap between the ideal MMSE and the proposed estimator.

In Fig. 2, we compare MSE versus the number of BS antennas M under the setting of $a = 0.05$ and TX SNR $q = 0$ 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. We also compare the influence of the number of elements of $\|\mathbf{z}_{ik}\|^2$ taken in the average.

In Fig. 3, we compare MSE versus the number of channel coefficients, P, versus the pilot symbols length, N, for TX SNR $q = 20$ dB and $M = 30$. With the increase of P, both the MSE of the proposed method approaches that of the ideal MSE estimator and the number of pilot symbols increase in order to provide users with orthogonal sequences.

V. Conclusions

In this work we have proposed a simple and practical channel estimator for multipath multi-cell massive MIMO TDD systems with pilot contamination. The proposed estimator replaces the combined interference, $i.e.$, the summation of the cross-cell large scale coefficients, plus noise power term in the ideal MMSE estimator with a simple and practical estimate. Also, the proposed estimator presents MSE results very close to that of the ideal MMSE estimator without requiring any additional overhead.

For future work, we intend to derived an analytical MSE expression of the proposed estimator which will be useful in system design and performance evaluations.

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] 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.

[3] 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.

[4] 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.

[5] A. Ashikhmi, T. L. Marzetta, and L. Li, “Interference reduction in multi-cell massive MIMO systems I: large-scale fading precoding and decoding”, arXiv preprint arXiv:1411.4182, 2014.

[6] N. Shariati, E. Bjornson, M. Bengtsson and M. Debbah, “Low-Complexity Polynomial Channel Estimation in Large-Scale MIMO with Arbitrary Statistics”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 815-830, Oct. 2014.

[7] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A Coordinated Approach to Channel Estimation in Large-Scale MultipleAntenna Systems”, *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264-273, Feb. 2013.

[8] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, “Pilot beam pattern design for channel estimation in massive MIMO systems”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 787-801, Oct. 2014.

[9] K. T. Truong and R. W. Heath, “Effects of Channel Aging in Massive MIMO Systems”, *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338-351, August 2013.

[10] R. R. Muller, L. Cottatellucci, and M. Vehkapera, “Blind Pilot Decontamination”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773-786, Oct 2014.

[11] Amin Khansefid, Hlaing Minn, “On Channel Estimation for Massive MIMO With Pilot Contamination”, *IEEE Communications Letters.*, vol. 19, no. 9, pp. 1660-1663, Sept. 2015.

[12] David C. Chu, “Polyphase codes with good periodic correlation properties”, *IEEE Transactions on Information Theory*, pp. 531-532, July 1972.

[13] S. M. Kay, “Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory”, New York, NY, USA: Pearson Education, 1993.