Low-complexity channel estimation for time division duplex massive multi-user multi-input multi-output systems

Muamer Hawej | Yousef R. Shayan

Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

Correspondence
Muamer Hawej, Department of Electrical and Computer Engineering, Concordia University, Montreal PQ H3G 1M8, Canada.
Email: m.malhwaij@yahoo.com

Abstract
This paper addresses the problem of minimum mean square error channel estimator for time division duplex massive multi-user multi-input multi-output systems. It is noteworthy that, the minimum mean square error has been previously proposed for multi-cell massive multi-user multi-input multi-output channel estimation. However, the minimum mean square error estimator suffers from high computational complexity due to the large dimension of the covariance matrix inversion. In this study, low-complexity channel estimator for time division duplex massive multi-user multi-input multi-output networks is designed by using the low-rank matrix approximation techniques. The proposed estimator is referred to as an approximate minimum mean square error estimator. Furthermore, the computational complexity of the proposed approximate minimum mean square error estimator is analysed and compared to the minimum mean square error and least square estimators. The normalised mean square error and the uplink achievable sum-rate performance criteria are used to evaluate the performance of the proposed estimator. Finally, the simulation results show the effectiveness of the proposed estimator under two different scenarios: noise-limited and pilot contamination. These simulation results are compared to the conventional minimum mean square error and least square estimators.

1 | INTRODUCTION

In recent years, multi-cell massive multi-user multi-input multi-output (MIMO) technology has attracted significant research interests in wireless communication system [1]. It has been recently proposed for the fifth-generation (5G) mobile communication network [2, 3]. In this technology, each cell has one base station (BS) which is supplied with hundreds of antenna elements, and simultaneously serving tens of users with single antenna each. Compared to the classical MIMO system, the massive multi-user MIMO system has more antennas at each BS. These increasing numbers of antennas in such a system can offer several benefits, such as better link reliability, much improvement in both spectral and power efficiencies, and simplicity of the system implementation [4]. These advantages can be only achieved if the channel state information (CSI) is precisely known at the BS. In the real scenario, however, the perfect CSI is not available at the BS. Therefore, it will be estimated either at the BS when the time division duplex (TDD) mode is considered for massive MIMO or at the user terminal when the frequency division duplex mode is used [5]. Due to the reciprocity property, the TDD mode has been considered as an operating system for massive MIMO [6].

The channel estimation in multi-cell massive multi-user MIMO networks is one of the significant challenges due to the so-called pilot contamination problem [6]. This problem occurs in the uplink pilot transmission phase when the same frequency band is used in all cells, which results in the channel estimation corruption [7]. Therefore, many channel estimation approaches are proposed to alleviate pilot contamination in massive MIMO systems [8].

In literature, the conventional pilot-based channel estimation methods such as the least square (LS) and minimum mean square error (MMSE) are used in classical MIMO systems [9]. In massive MIMO systems, the LS channel estimation performance is so bad due to the noise and pilot contamination problems [10, 11]. On the other hand, the MMSE is considered as an optimal channel estimator scheme for massive MIMO systems when the statistic information of the desired and interference channels is known at the BS [12]. In other words, this information is not always available at the BS. Moreover, the MMSE is of a considerable computational complexity due to...
the large dimensions of a matrix inversion which are scaled with the number of BS antennas. In practice, however, the number of multiplication operations required to design the MMSE channel estimator scheme is high. Thus, the conventional MMSE estimator is hard to implement in the real massive MIMO systems.

The optimal rank reduction is a popular theory that is used to reduce the computational complexity of channel estimators [13, 14]. One of the optimal rank reduction theory applications is the low-rank matrix approximation (LRMA) technique. This technique has been considered for channel estimation in millimetre-wave massive MIMO system [15–18]. Also, it has been previously proposed for channel estimation in a low frequency band (<6 GHz) massive multi-user MIMO system [19, 20]. In this work, to take advantage of the LRMA approximation techniques, we propose to design a low-complexity channel estimator for a multi-cell TDD massive multi-user MIMO system.

The contributions of this work are summarised as follows:

- A novel low-complexity AMMSE channel estimator has been designed by using the iterative weighted nuclear norm (IWNNN) approximation technique to reduce the computational complexity of the conventional MMSE estimator.
- The computational complexity of the proposed estimator is analysed and compared to the conventional LS and MMSE estimators.
- The NMSE and uplink ASR performance metrics are used to evaluate the proposed AMMSE estimator under the noise and pilot contamination scenarios.
- Finally, the simulation results show that the proposed estimator has a better estimation performance compared to the LS. However, it is shown a bit of performance loss compared to the MMSE.

The notations are used in this work as follows. We use the uppercase and lowercase boldface to denote matrices and vectors, respectively. The inverse, transpose and conjugate transpose of a matrix $X$ are denoted by $(X)^{-1}$, $(X)^T$ and $(X)^H$, respectively. $I_M$ denotes the $M \times M$ identity matrix. $\| \cdot \|_F$ stands for the Frobenius norm of a matrix $(\cdot)$, and diag $(\cdot)$ denotes a diagonal matrix $(\cdot)$.

The rest of this work is organised as follows. The multi-cell massive multi-user MIMO system and channel models are introduced in Section 2. In Section 3, the LS, MMSE and proposed AMMSE estimators for TDD massive multi-user MIMO system are explained. A brief computational complexity of the AMMSE estimator is analysed and explained in Section 4. In Section 5, the simulation results are illustrated and discussed. Finally, the conclusion is summarised in Section 6.

## 2 | SYSTEM AND CHANNEL MODELS

### 2.1 | Massive multi-user multi-input multi-output system model

A multi-cell TDD massive multi-user MIMO cellular network with $L$ cells is considered. In this system, each cell has one BS with hundreds of antenna elements, $M$, and simultaneously serves tens of $K$ users with single antenna each [3]. In this study, we suppose that the BS in the cell-$j$ is the target BS- $j$ unless otherwise specified. Moreover, the orthogonal pilot sequences are used in the cell-$j$ to avoid inter-user interference among $K$ users. However, due to the limited coherence time interval in the TDD system, these pilot sequences are reused by other users in each neighbouring cells, which results in the pilot contamination [7].

To illustrate the above idea, we study the CSI estimation when the proposed AMMSE is used at each BS. Thus, the uplink pilot phase is considered where all users from all cells simultaneously transmit their pilot signals to their BSs. Hence, at each time, $t$, the uplink received pilot signals $y^t(\cdot) \in \mathbb{C}^M$ from all $K$ in vector form at the BS-$j$ is denoted by

$$y^t(\cdot) = \sqrt{\rho_{\text{tr}}} H_j^{(t)} x^{(t)} + \sqrt{\rho_{\text{tr}}} \sum_{l \neq j} H_l^{(t)} x^{(t)} + n^t(\cdot), \quad (1)$$

where $x^{(t)} = [x_1^{(t)} x_2^{(t)} \cdots x_K^{(t)}]^T$ is the transmitted pilot signal vector from all $K$ users inside each cell to own BS, $\rho_{\text{tr}}$ is the average power used by each user and $n^t(\cdot) \in \mathbb{C}^{M \times 1}$ is the additive white Gaussian noise (AWGN) vector with zero mean and unit variance. In (1), $H_j^{(t)} \in \mathbb{C}^{M \times K}$ is the channel matrix between the desired $K$ users and the BS-$j$ which is defined as

$$H_j^{(t)} \overset{\Delta}{=} [h_j^1 h_j^2 \cdots h_j^K], \quad (2)$$

and $H_l^{(t)} \in \mathbb{C}^{M \times K}$ is the channel matrix between the interfering users in the adjacent cells $l \neq j$ and the BS-$j$ which is defined as

$$H_l^{(t)} \overset{\Delta}{=} [h_l^1 h_l^2 \cdots h_l^K]. \quad (3)$$

During the training phase, the total received pilot signals at the BS-$j$ in matrix form is given as

$$y^t = \sqrt{\rho_{\text{tr}}} H_j^t X + \sqrt{\rho_{\text{tr}}} \sum_{l \neq j} H_l^t X + N^t. \quad (4)$$

Here $X = [x(1) x(2) \cdots x(\tau)] \in \mathbb{C}^{K \times \tau}$ is the orthogonal pilot signals transmitted from all $K$ users inside each cell to own BS, and $\tau$ is the length of each pilot sequence. In (4), $N^t \in \mathbb{C}^{M \times \tau}$ is a noise matrix with independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance $\sigma_n^2$ elements.

### 2.2 | Massive multi-user multi-input multi-output channel model

In this study, we are mainly focussed on the finite scattering propagation environment where the number of physical objects is limited. Hence, the realistic finite scattering channel model is
adopted, which is recently considered for multi-cell multi-user massive MIMO technologies [3, 12, 19, 20]. The channel vector between the $k$th user in each cell and the BS-$j$ is given by

$$H_j^k = \hat{\beta}_j^k \sqrt{\rho_j^k} \sum_{p=1}^{P} a(\theta_p^k) g_{kp}^j,$$

where $\hat{\beta}_j^k$ and $g_{kp}^j$ are the path loss coefficient and fading gain and between the $k$th user in each cell and the BS-$j$, respectively. For each $k$th user, $\hat{\beta}_j^k$ is denoted by [12] as

$$\hat{\beta}_j^k = \sqrt{\frac{\alpha}{(d_j^k)^\xi}},$$

where $d_j^k$ is the distance between the $k$th user in the cell-$j$ and the BS-$j$, $\xi$ is the path loss exponent factor and $\alpha$ is a constant dependent on the signal-to-noise ratio (SNR) at cell edge. In (5), $a(\theta_p)$ is the $M \times 1$ steering vector between each $k$th user and the BS-$j$ associated with each path $p$, which is defined as

$$a(\theta_p) = \left[ 1, e^{-j2\pi \frac{D}{\lambda} \cos(\theta_p)}, \ldots, e^{-j2\pi (M-1) \frac{D}{\lambda} \cos(\theta_p)} \right]^T, \quad (7)$$

where $\lambda$ is the signal wavelength, $D$ is the fixed distance between two BS antenna elements and $\theta_p \in [-\pi/2, \pi/2]$ is the angle of arrival (AoA) associated with each path $p$, which is adopted, the AoA are perfectly known at each BS with the knowledge of the covariance matrices $\Sigma_{m}$ of the desired channels.

$$H_j^k = AG_j^k (D_j^k)^{1/2}, \quad (8)$$

where $A \triangleq [a(\theta_1) \ldots a(\theta_P)]$ is the $M \times P$ steering matrix with full rank, and $D_j^k \triangleq \text{diag}(\hat{\beta}_j^k, \hat{\beta}_j^k, \ldots, \hat{\beta}_j^k)$ is a $K \times K$ diagonal matrix with the path loss coefficients on its main diagonal.

In (8), $G_j^k \triangleq [g_{j1}^k \ldots g_{jK}^k]$ is the $P \times K$ Rayleigh fading channel matrix gain whose entries are i.i.d. with zero mean and unit variance $\sigma^2$. Under this channel model, all users in all cells can be collectively written in matrix form as

$$\begin{align*}
H_j^k = \left[ H_j^1 \ H_j^2 \ \cdots \ H_j^K \right],
\end{align*}$$

### 3.1 Least square channel estimator

The LS channel estimator is a low-complexity approach that can be achieved by multiplying the received signal $y_j$ in (4) with the known pilot sequence $X$ [9]. For massive multi-user MIMO, the LS channel estimation matrix $\hat{H}_{LS}$ is denoted by [12] as

$$\hat{H}_{LS} = \frac{1}{\sqrt{\rho_j^k}} X^H (XX^H)^{-1}, \quad (9)$$

where $X$ is the orthogonal pilot sequences, which are assumed to be known at the BS-$j$. By substituting $y_j$ into (9) and applying the orthogonality property of $XX^H = \Gamma I_K$, we rewrite (9) as

$$\hat{H}_{LS} = H_j^k + \sum_{l \neq j} H_l^k + \frac{1}{\tau} I_{K}, \quad (10)$$

As it appears in the second term of (10), the LS channel estimation performance is mainly corrupted by the inter-cell interference channels (pilot contamination). This effect degrades the channel estimation performance, and therefore, the LS channel estimator may not be used in the real massive MIMO systems.

### 3.2 Minimum mean square error channel estimator

A linear MMSE estimator for massive MU-MIMO channel matrix is given by [12] as

$$\hat{H}_{MMSE} = R_j^f \left( H_j^f + \sum_{l \neq j} R_l^f + \frac{\sigma^2}{\tau \rho_j^k} I_M \right)^{-1}, \quad (11)$$

where $\hat{H}_{LS}$ is the LS channel estimation, and $R_j^f \in C^{M \times M}$ and $R_l^f \in C^{K \times K}$ are the desired and interference covariance channel matrices, respectively. As it appears in (11), the MMSE estimator suffers from the following drawbacks: (i) It needs the knowledge of both covariance matrices $R_j^f$ and $R_l^f \forall i \neq j$, which are not perfectly known at the BS in the real system. (ii) It is of considerable computational complexity since the dimensions of matrix inversion are increased with the number of BS antennas. Thus, in this study, we further proceed with the LRMA techniques, such as the IWN approximation to reduce the computational complexity of the MMSE estimator.

### 3.3 Proposed Approximate minimum mean square error channel estimator

In this section, the eigenvalue decomposition (EVD) and IWN approximation methods are used to design AMMSE channel estimator. The EVD is applied to the conventional MMSE channel estimator in (11). In the Appendix, it is shown...
that the proposed approximate minimum mean square error (AMMSE) channel estimator is

\[
\hat{H}_{\text{AMMSE}}' = U_{\Delta_r} (U_{\Delta_r}^H)^H \hat{H}_{1,\text{LS}}
\]

where \(U_{\Delta_r} \in \mathbb{C}^{M \times r}\) is the desired eigenvector matrix which can be approximated by using the discrete Fourier transform basis \([21]\), and \(\Delta_r\) is a diagonal matrix with entries

\[
\delta_j = \begin{cases} \sigma_j, & i = 1, 2, \ldots, r \\ \sigma_j + \sum_{i \neq j} \sigma_i + \frac{\sigma_j}{\tau}, & i = r + 1, \ldots, M \\ 0, & \end{cases}
\]

where \(r\) is the rank of the desired channel covariance matrix.

As it appears in (12), a diagonal matrix \(\Delta_r\) is only needed to be known at the BS. Therefore, the IWNN approximation based on the low-rank reduction theory is used in this work to compute a diagonal matrix \(\Delta_r\). It is noteworthy that, the IWNN approximation has been applied for massive multi-user MIMO channel estimation in our previous work \([19, 20]\). However, it is also adopted in this work. Hence, the eigenvalue estimation problem in (24) (see Appendix) is formulated as an unconstrained weighted nuclear norm optimisation problem as follows:

\[
\Delta_r = \arg \min_{\Sigma_j} \left\{ \frac{1}{2} \left\| \Sigma_j - \Sigma_j' \right\|_F^2 + \gamma \sum_{j=1}^{r} |w_i| \sigma_j \right\},
\]

where \(\Delta_r\) is a diagonal matrix with an approximate eigenvalues \(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r\) at the main diagonal, and \(\gamma\) is the regularisation parameter. In (14), the regularisation parameter \(\gamma\) is used to control the trade-off between the data fidelity \(\frac{1}{2} \left\| \Sigma_j - \Sigma_j' \right\|_F^2\) and the penalty function, \(\gamma \sum_{j=1}^{r} |w_i| \sigma_j\), which is computed using the formula given by \([22]\) as

\[
\gamma = \sqrt{2(M + L\tau)\tau r_0 \sigma_{\text{max}}^2}.
\]

In (14), \(w_i\) is the weight element which can be obtained using the following formula given by \([12]\) as

\[
w_i^{n+1} = w_i^n + \frac{\mu}{\bar{\sigma}_i^n + \varepsilon} \quad i = 1, 2, \ldots, r,
\]

where \(\bar{\sigma}_i^n\) is the estimated eigenvalue of \(\Delta_r\) in the \(n\)th iteration and \(\mu\) is the step-size parameter which is used to accelerate time convergence. To avoid dividing Equation (16) by zero, a small positive number \(\varepsilon\) is added. The proposed IWNN approximation algorithm used to design the AMMSE channel estimator is summarised in Table 1.

### Table 1: IWNN proposed for AMMSE estimator

| Channel estimators | Computational complexity | Estimation performance |
|--------------------|--------------------------|-----------------------|
| LS                 | \(O(Mr)\)                | Low                   |
| AMMSE              | \(O(MrPN)\)              | Moderate              |
| MMSE              | \(O(M^3\tau^3)\)         | High                  |

### 4 ESTIMATOR COMPLEXITY

The main complexity of the MMSE channel estimator in (11) comes from a large dimension of a matrix inversion, which is scaled with the number of BS antennas \(M\) \([23]\). Thus, the total number of multiplication operations required to compute the MMSE channel estimation matrix is \(M^3\). On the other hand, the total number of multiplications required to compute the AMMSE channel estimation matrix is only \(MrPN\), where the total number of iterations is \(N\). Compared to the MMSE estimator, the total number of multiplications required to compute the channel estimation matrix by using the proposed AMMSE estimator is reduced from \(M^3\) to \(MrPN\). However, the proposed AMMSE estimator can be considered as low-complexity channel estimator only if \(N \leq 10\). The asymptotic complexities and the estimation performances of the LS, MMSE and the proposed AMMSE estimators are illustrated in Table 2.

### 5 SIMULATION RESULTS

In this section, we provide the simulation results of the proposed AMMSE estimator for TDD multi-cell massive
MU-MIMO system. Thus, we consider a multi-cell massive MU-MIMO system with \( L = 3 \) cells, and each cell contains the total number of \( K = 20 \) users, and one BS with \( M = 100 \) antennas. We assume the length of the pilot sequence is \( \tau = 20 \), and the number of multi-path is \( P = 10 \). The SNR is set for 0 dB, and the steering vector parameters are selected to have \( D/\lambda = 0.5 \), and \( \theta_p = -\pi/2 + (p-1)\pi/2P \), where \( p = 1, 2, \ldots, P \) [3, 11, 12].

Moreover, the simulation results of the proposed AMMSE estimator are compared to the conventional LS and MMSE estimators under two different scenarios: noise-limited and pilot contamination. To investigate the above, the following two criteria are used to evaluate each estimator performance [12]. The first criteria is a normalised mean square error (NMSE), which is given as

\[
\text{NMSE (dB)} = 10\log_{10}\left( \frac{E \left\| \hat{H}^j - H^j \right\|_F^2}{E \left\| H^j \right\|_F^2} \right),
\]

where \( H \) and \( \hat{H}^j \) are the desired channel and its estimate using the conventional MMSE or AMMSE estimator, respectively. The second criteria is the uplink achievable sum-rate (ASR) of \( K \) users which is defined as

\[
\text{ASR} \leq \sum_{k=1}^{K} \log_2 (1 + \text{SINR}_k),
\]

where \( \text{SINR}_k \) is the received signal-to-interference-plus-noise ratio of the \( k \)th user at the linear detector output. When the linear maximum ratio combing (MRC) detector is assumed at the BS-\( j \), the weight vector of MRC is defined as \( \mathbf{v}_j = \hat{h}_j \).

### 5.1 Noise-limited scenario

In the noise-limited scenario, we study the behaviours of the proposed AMMSE estimator for massive multi-user MIMO under a noisy setting. This scenario is given by setting \( \beta_{jk} = 0 \) for all interfering channels in the adjacent cells, while \( \beta_{jk} = 1 \) for all desired channels in the desired cell-\( j \). The SNR is set for 0 dB at cell-edge. In Figures 1 and 2, the NMSE and uplink ASR of the proposed AMMSE estimator are evaluated under a different number of BS antennas \( M \) and compared to the conventional LS and MMSE estimators. Figure 1 shows that as \( M \) increases, both estimation performances of the AMMSE and MMSE estimators are quickly converged to zero, while the LS estimator does not show any achievement. Moreover, the NMSE of the proposed AMMSE estimator is only about 0.5 dB less than the MMSE estimator when \( M \geq 100 \). Figure 2 shows that the uplink ASRs of the proposed AMMSE and conventional MMSE estimators are improved as \( M \) increases and almost identical to each other for all evaluated values of \( M \). On the other hand, the uplink ASR of the LS estimator is quickly saturated as the number of BS antennas \( M \geq 100 \). As mentioned earlier in Section 4, the main computational complexity of the proposed AMMSE channel estimator comes from increasing the number of iterations \( N \). Therefore, we turn our attention to study this issue. Figures 1 and 2 show that the AMMSE channel estimation performance in terms of the NMSE and uplink ASR is converged to the MMSE performance with only \( N = 5 \) iterations. Thus, it is much smaller than the computational complexity of the MMSE estimator.

### 5.2 Pilot contamination scenario

In this scenario, the behaviours of the proposed AMMSE estimator under two different pilot contamination cases (i.e. \( \beta_{jk} = 0.1 \) and \( \beta_{jk} = 0.9 \)) are studied. For both cases, we set the SNR = 0 dB, and \( \beta_{jk} = 1 \) for all desired channels in the desired
cell-$j$. In the first experiment the impact of the weak pilot contamination on the proposed AMMSE estimator, is studied under a different number of BS antennas, $M$ as shown in Figures 3 and 4. In the second experiment, we examine the impact of strong pilot contamination on the proposed AMMSE estimator under a different number of BS antennas.

To understand the above scenario, we assume that each interfering channel has $\beta_{jk} = 0.9$ which represents the strong pilot contamination case. Figures 5 and 6 show that the NMSE and uplink ASR performances of the proposed AMMSE estimator are improved as $M$ increases, compared to the LS estimator. After one iteration, the NMSE and uplink ASR performances of the proposed AMMSE estimator are about 3 dB and 3.5 bps/Hz less than the MMSE estimator, respectively. In contrast, the uplink ASR obtained by the LS channel estimation is quickly saturated for $M \geq 100$ in both pilot contamination cases as shown in Figures 4 and 6.

6 | CONCLUSION

In this work, a novel low-complexity channel estimator namely “AMMSE” for multi-cell TDD massive multi-user MIMO systems has been proposed. The IWNN approximation based on the low-rank reduction theory is used to design the proposed AMMSE estimator. Compared to the conventional MMSE estimator, the computational complexity of the proposed AMMSE estimator regarding the number of multiplications is reduced from $\mathcal{O}(M^3 \tau^2)$ to $\mathcal{O}(M \tau P N)$. The simulation results show the
agreements between the proposed estimator and the conventional MMSE estimator in terms of the NMSE and the uplink ASR performances. These estimation performances of the proposed AMMSE estimator have been investigated under two different scenarios: noise-limited and pilot contamination. Moreover, the proposed AMMSE channel estimation performance is much better than the LS estimation regarding the NMSE and the uplink ASR.

ACKNOWLEDGEMENTS
This work was mainly supported by Ministry of Higher Education in Libya, and partially by Natural Sciences and Engineering Research Council (NSERC) in Canada.

REFERENCES
1. Zheng, K., et al.: Survey of large-scale MIMO systems. IEEE Commun. Surv. Tutorials 17(3), 1738–1760 (2015)
2. Larsson, E., et al.: Massive MIMO for next-generation wireless systems. IEEE Commun. Mag. 52(2), 186–195 (2014)
3. Ngo, H., et al.: The multi-cell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel. IEEE Trans. Commun. 61(6), 2350–2361 (2013)
4. Lu, L., et al.: An overview of massive MIMO: Benefits and challenges. IEEE J. Sel. Top. Signal Process. 8(5), 742–758 (2013)
5. Osseiran, A., et al.: Scenarios for 5G mobile and wireless communications: The vision of the METIS project. IEEE Commun. Mag. 52(5), 26–35 (2014)
6. Rusek, F., et al.: Scaling up MIMO: Opportunities and challenges with very large arrays. IEEE Signal Process. Mag. 30(1), 40–60 (2013)
7. Figueiredo, F.A.P., et al.: Channel estimation for massive MIMO TDD systems. In: IEEE International Conference on Communications, Cape Town, South Africa, pp. 1–5 (2010)
8. Elijah, O., et al.: A comprehensive survey of pilot contamination in massive MIMO-5G system. IEEE Commun. Surv. Tutorials 18(2), 905–923 (2016)
9. Biguesh, M., Gershman, A.: Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals. IEEE Trans. Signal Process. 54(3), 884–893 (2006)
10. Marzetta, T.L.: Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wireless Commun. 9(11), 3590–3600 (2010)
11. Figueiredo, F.A.P., et al.: Channel estimation for massive MIMO TDD systems assuming pilot contamination and flat fading. EURASIP J. Wireless Commun. Networking 2018(14), 1–10 (2018)
12. Yin, H., et al.: A coordinated approach to channel estimation in large-scale multiple-antenna systems. IEEE J. Sel. Areas Commun. 31(2), 264–273 (2013)
13. Shariat, N., et al.: Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics. IEEE J. Sel. Topics Signal Process. 8(5), 815–830 (2014)
14. Nie, H., et al.: An overview of low-rank channel estimation for massive MIMO systems. IEEE Access 4, 2014–2018 (2016)
15. Alkhateeb, A., et al.: Channel estimation and hybrid precoding for millimeter wave cellular systems. IET Commun. 8(5), 831–846 (2014)
16. Sung, J., et al.: Channel estimation for millimeter-wave multi-user MIMO systems via PARAFAC decomposition. IEEE Trans. Wireless Commun. 15(11), 7501–7516 (2016)
17. Zhou, Z., et al.: Low-rank tensor decomposition-aided channel estimation for millimeter-wave MIMO-OFDM systems. IEEE J. Sel. Areas Commun. 35(7), 1524–1538 (2017)
18. Elahi, P.A., et al.: Low-rank spatial channel estimation for millimeter-wave cellular systems. IEEE Trans. Wireless Commun. 16(5), 2748–2759 (2017)
19. Hawej, M., Shayan, Y.: Pilot decontamination in massive multi-user MIMO systems based on low-rank matrix approximation. IET Commun. 13(5), 594–600 (2018)
20. Hawej, M., Shayan, Y.: Iterative weighted nuclear norm minimization-based channel estimation for massive multi-user MIMO systems. In: IEEE 88th Vehicular Technology Conference (VTC), Chicago, IL, August 2018, pp. 1–5
21. Chen, Z., Yang, C.: Pilot decontamination in wideband massive MIMO systems by exploiting channel sparsity. IEEE Trans. Wireless Commun. 15(7), 5087–5100 (2016)
22. Nguyen, S.L.: Compressive sensing for multi-channel and large-scale MIMO networks. Ph.D. Dissertation, Concordia University (2013)
23. Savaux, V., et al.: Low-complexity approximations for LMMSE channel estimation in OFDM/OQAM. In: 23rd International Conference on Telecommunications (ICT), Thessaloniki, Greece, May 2016, pp. 1–5

APPENDIX
In this Appendix, the AMMSE estimator in (12) is proved by applying the EVD and the low-rank reduction theory to the MMSE estimator in (11). First, the EVD is applied to the covariance matrices $R_j$ and $R_j'$ in (11) as

$$\hat{H}_{\text{MMSE}} = U_j' \Sigma_j' \left( U_j' \right)^H \times \left( U_j' \Sigma_j' \left( U_j' \right)^H + \sum_{j' \neq j} U_j' \Sigma_j' \left( U_j' \right)^H + \frac{\sigma_r^2}{\rho_{tr}} I_M \right)^{-1} \hat{H}_{L,S}. \tag{A1}$$

where $\Sigma_j'$ is the diagonal desired channel matrix containing the eigenvalues $\sigma_{j1} \geq \sigma_{j2} \geq \cdots \geq \sigma_{JM}$ on its diagonal, and $\Sigma_j' \forall j \neq j$ is the diagonal interference channel matrix containing the eigenvalues $\sigma_{j1} \geq \sigma_{j2} \geq \cdots \geq \sigma_{JM}$ on its diagonal. In (19), $U_j' \in C^{M_{x}M_{j}}$ and $U_j' \in C^{(M_{x}M_{j})}$ are the eigenvector matrices where $m_j$ and $m_j$ are the rank of desired and interference channel covariance matrices, respectively.

The worst-case scenario in massive multi-user MIMO systems is when all users from all cells share the same AoAs which are received at the BS-$j$. Under this assumption, the correlation among these channels increases. Therefore, all channel users will have the same steering matrix with $U_j' = U_j'$ and different eigenvalues. Thus, Equation (A1) is simplified as

$$\hat{H}_{\text{MMSE}} = U_j' \Sigma_j' \left( U_j' \right)^H \times \left( U_j' \Sigma_j' + \sum_{j \neq j'} U_j' \Sigma_j' \left( U_j' \right)^H + \frac{\sigma_r^2}{\rho_{tr}} I_M \right)^{-1} \hat{H}_{L,S}. \tag{A2}$$
Moreover, applying the matrix inversion identity $A(BA + I)^{-1} = (AB + I)^{-1}A$ results in

$$\hat{H}^j_{\text{MMSE}} = U_j \Sigma_j \left( U_j^H U_j \left( \Sigma_j + \sum_{i \neq j} \Sigma_i \right) + \frac{\sigma_n^2}{\tau \rho_{tr}} I_M \right)^{-1} \times (U_j^H) \hat{H}^j_{\text{LS}}.$$  

(A3)

In the massive MIMO system, we have $(U_j^H) U_j = I_M$ when $M$ is a large number [12]. By substituting this feature into (21), we rewrite (21) as

$$\hat{H}^j_{\text{MMSE}} = U_j \Sigma_j \left( \Sigma_j + \sum_{i \neq j} \Sigma_i + \frac{\sigma_n^2}{\tau \rho_{tr}} I_M \right)^{-1} (U_j^H) \hat{H}^j_{\text{LS}}.$$  

(A4)

which is equivalent to

$$\hat{H}^j_{\text{MMSE}} = U_j \Sigma_j \left( \Sigma_{\text{LS}} \right)^{-1} (U_j^H) \hat{H}^j_{\text{LS}},$$  

(A5)

where $\Sigma_{\text{LS}}$ is a diagonal matrix with LS estimate eigenvalues $\hat{\sigma}_{j1}^2 \geq \hat{\sigma}_{j2}^2 \geq ... \geq \hat{\sigma}_{jM}^2$ at the main diagonal which can be expressed as

$$\Sigma_{\text{LS}} = \Sigma_j + \sum_{i \neq j} \Sigma_i + \frac{\sigma_n^2}{\tau \rho_{tr}} I_M.$$  

We further simplify (23) as

$$\hat{H}^j_{\text{MMSE}} = U_j \Delta (U_j^H) \hat{H}^j_{\text{LS}},$$  

(A7)

where

$$\Delta = \Sigma_j \left( \Sigma_j + \sum_{i \neq j} \Sigma_i + \frac{\sigma_n^2}{\tau \rho_{tr}} I_M \right)^{-1}$$

= \text{diag} \left( \frac{\sigma_{j1}}{\sigma_{j1} + \sum_{i \neq j} \sigma_{i1} + \frac{\sigma_n^2}{\tau \rho_{tr}}}, ..., \frac{\sigma_{jM}}{\sigma_{jM} + \sum_{i \neq j} \sigma_{iM} + \frac{\sigma_n^2}{\tau \rho_{tr}}} \right).$$  

(A8)

The theory of low-rank reduction in [14] then is applied to introduce the proposed AMMSE channel estimator as

$$\hat{H}^j_{\text{AMMSE}} = U_j \left[ \Delta_r 0 \right] (U_j^H) \hat{H}^j_{\text{LS}},$$  

(A9)

where $\Delta_r$ is the $r \times r$ upper left corner of the matrix $\Delta$. 

The expression is simplified further by considering the case when $\Delta_r = I_r$, resulting in

$$\hat{H}^j_{\text{AMMSE}} = U_j \left[ I_r 0 \right] (U_j^H) \hat{H}^j_{\text{LS}}.$$  

(A10)