SAGE Algorithm Based MAP Channel Estimation for Multi-Cell Massive MIMO Systems

Şenol Sançar
Department of Electrical and Electronics Engineering
Istanbul Sabahattin Zaim University
Istanbul, TURKEY
senol.sancar@izu.edu.tr

Bahattin Karakaya
Department of Electrical and Electronics Engineering
Istanbul University
Istanbul, TURKEY
bahattin@istanbul.edu.tr

***Abstract*—This paper represents an efficient space-alternating generalized expectation-maximization (SAGE) algorithm based maximum a posteriori (MAP) channel estimation method for multi-cell massive multiple input multiple output (MIMO) systems. MAP channel estimation method requires conjugate transpose of a $\tau \times K$ pilot matrix where τ is the number of pilot symbols per user and K is the number of single antenna users. Conjugate transpose of the large-size matrix increases computational complexity. The proposed method estimates the channel iteratively and converges to the same mean square error (MSE) performance of the MAP estimator with the increasing number of iterations. Consequently, the proposed method with low-rank approximation avoids conjugate transpose of the large-size matrix and hence reduces the computational complexity significantly.**

***Index Terms*—Channel Estimation, Pilot Contamination, Multi-Cell Massive MIMO, SAGE-Algorithm, Maximum A Posteriori, Least Squares**

I. Introduction

Multi-cell massive MIMO systems are up-and-coming technology for future's cellular systems. In the beginning, massive MIMO systems were focused on point-to-point MIMO which two devices communicate with each other using multiple antennas, afterward the focus has shifted towards multi-cell massive MIMO systems [1]. In multi-cell massive MIMO systems, a base station (BS) equipped with multiple antennas simultaneously serves a number of single antenna users.

In multi-cell massive MIMO systems, frequency reuse causes rise to inter-cell interference (ICI) in channel estimation and it is known as pilot contamination [2]. Channel state information (CSI) at the BS plays a key role in multi-cell massive MIMO systems because it is critical for achieving high system performance.

Time division duplex (TDD) is considered a better method to acquire CSI in multi-cell massive MIMO systems compared to frequency division duplex (FDD) because TDD requires channel estimation in only uplink training and the estimated channel is reciprocal which means the estimated channel can be used in both uplink and downlink transmissions [3].

Multi-cell massive MIMO systems provide a number of advantages such as spectral efficiency, and power efficiency. Spectral efficiency is achieved by serving several users simultaneously through spatial multiplexing. Transmit power efficiency is inversely proportional to the number of BS antennas if BS has perfect-CSI and the transmit power is inverse proportional to the square root of BS antenna numbers if BS has estimated-CSI, compared to single antenna systems [4].

In this paper, an efficient SAGE algorithm based MAP channel estimation method is proposed, which exhibits the same MSE performance with MAP estimator after a number of iterations. The superiority of the proposed method is reducing the computational complexity. The MAP estimator requires conjugate transpose of a large-size matrix and when the size of the pilot matrix increases the computational complexity increases too. The proposed method avoids this problem. For the proposed system model MAP estimator is equivalent to minimum mean square error (MMSE) estimator.

This paper is organized as follows: In Section II, the system model is introduced. In Section III, the conventional channel estimation methods are introduced. In Section IV, the iterative channel estimation method which reduces computational complexity is analyzed. In Section V, the performance of the algorithm is investigated under different circumstances by simulative means. Finally, in Section VI the paper is concluded.

Notation: In this paper, $(.)^H$ denotes the conjugate transpose. $\|.\|^2$ denotes the Frobenius norm and $\mathbf{I_K}$ denotes the $K \times K$ identity matrix.

II. System Model

It is considered a multi-cell massive MIMO system with L cells which each cell contains a BS with M antennas and K $(K << M)$ users. We assume that all BSs propagate in same frequency band and lth BS receives signals from all users in all cells. g_{ilkm} represents the channel coefficients from user k in cell i to the antenna m of the BS in cell l.

$$g_{ilkm} = \sqrt{\beta_{ilk}} h_{ilkm} \tag{1}$$

where $\sqrt{\beta_{ilk}}$ represents the geometric attenuation and shadow fading which is non-negative constant and known a priori, h_{ilkm} represents the fast fading coefficient from user k in cell i to the antenna m of the BS in cell l with a zero mean

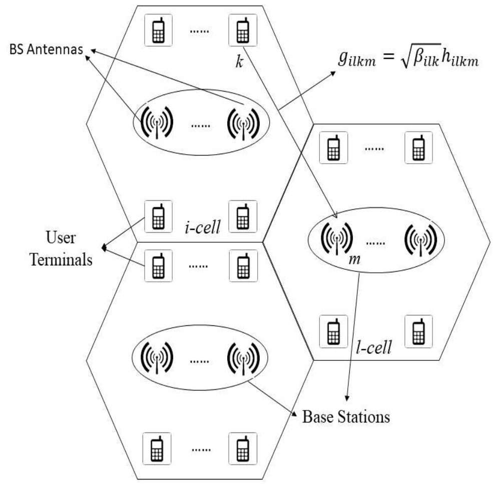

Fig. 1. The channel coefficient between user *k* in the cell *i* and BS antenna *m* in the cell *l*.

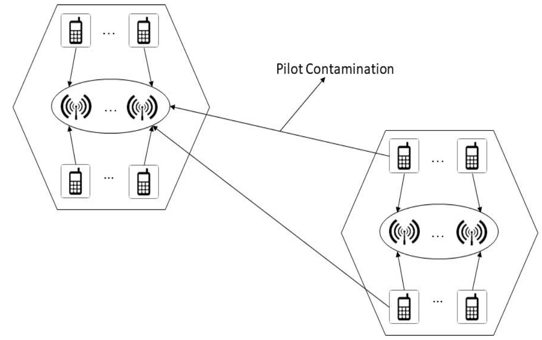

Fig. 2. Schematic view of pilot contamination.

and unit variance $\mathcal{CN}(0,1)$ distribution. The channel matrix G_{il} can be written as

$$G_{il} = D_{il}^{\frac{1}{2}} H_{il}, \tag{2}$$

where $H_{il} \in \mathbb{C}^{K \times M}$ is fast fading coefficients matrix between the *K* users in the *i*th cell and the *l*th BS, $[H_{il}]_{km} = h_{ilkm}$ and $D_{il} \in \mathbb{R}^{K \times K}$ is diagonal matrix whose diagonal elements are $[D_{il}]_{kk} = \beta_{ilk}$ [5], [6].

A. Pilot Contamination

In multi-cell massive MIMO systems, pilot sequences are transmitted from users to the BS in the uplink to estimate the channel. Each column of the pilot sequences within the same cell and in the neighboring cells should be orthogonal.

$$\Phi_i^H \Phi_i = \tau \mathbf{I_K}, i = 1, ..., L, \tag{3}$$

where $\Phi_i \in \mathbb{C}^{\tau \times K}$ pilot matrix. K is number of users and τ is number of pilot symbols per user. When the pilot matrices are orthogonal in all cells a BS can obtain uncontaminated estimation of the channel matrices. Unfortunately, the number of pilot sequences are limited, as a result, limits the number of users which can be served. To serve more users non-orthogonal pilot sequences are used in neighboring cells [2]. The use of non-orthogonal pilot sequences causes pilot contamination.

B. Uplink Training

The users transmit an uplink training sequence of τ pilot symbols and each BS estimates the channels of its users. We consider that users in different cells transmit the same set of pilot sequences at the same time. The pilot sequences of *K* users in each cell are represented as $\Phi_i \in \mathbb{C}^{\tau \times K}$ and pilot sequences are orthogonal $\Phi_i^H \Phi_i = \tau \mathbf{I_K}, \forall i$. Then $\tau \times M$ received training matrix at the BS *l* is represented by Y_l as

$$Y_l = \sqrt{p_u} \sum_{i=1}^{L} \Phi_i G_{il} + N_l, \tag{4}$$

where p_u is the power of each training symbols. $N_l \in \mathbb{C}^{\tau \times M}$ is an additive white Gaussian noise (AWGN) matrix with zero mean and unit variance $\mathcal{CN}(0,1)$ distribution [5].

For notational simplicity, the desired channel matrix can be represented as $G_{ll} \in \mathbb{C}^{K \times M}$ and the pilot sequence can be represented as $\Phi_l \in \mathbb{C}^{\tau \times K}$, all interfering cell pilot sequences are collected in $\tilde{\Phi} = [\Phi_1, ..., \Phi_L] \in \mathbb{C}^{\tau \times (L-1)K}, \forall i \neq l$ and channel matrices collected in $\tilde{G} = [G_{1l}^T, ..., G_{Ll}^T]^T \in \mathbb{C}^{(L-1)K \times M}, \forall i \neq l$ leading to

$$Y_l = \sqrt{p_u} \Phi_l G_{ll} + \sqrt{p_u} \tilde{\Phi} \tilde{G} + N_l. \tag{5}$$

III. Channel Estimation Methods

A. MMSE Channel Estimation

We assume that the BS uses MMSE channel estimation [7]. The MMSE estimator is a scaled form of LS estimator and MMSE can significantly outperform LS estimator with the presence of prior information at the BS. The estimated channel matrix $\hat{G}_{ll}^{mmse}$ can be written as [8]

$$\hat{G}_{ll}^{mmse} = \sqrt{p}_u D_{ll} \left(\mathbf{I_K} + \tau p_u \sum_{i=1}^{L} D_{il} \right)^{-1} \Phi_l^H Y_l. \tag{6}$$

B. MAP Channel Estimation

If the priors are known, MAP estimator can be formulated for multi-cell massive MIMO systems [8]. In the MAP estimation approach we choose $\hat{G}_{ll}$ and $\hat{\tilde{G}}$ to maximize the posteriori probability density function (PDF) or

$$(\hat{G}_{ll}, \hat{\tilde{G}}) = \underset{(G_{ll}, \tilde{G})}{argmax} [p(G_{ll}, \tilde{G} | Y_l)]. \tag{7}$$

To find the MAP estimator, we must equivalently maximize $p(Y_l | G_{ll}, \tilde{G}) p(G_{ll}, \tilde{G})$. Hence, the MAP estimator is

$$(\hat{G}_{ll}, \hat{\tilde{G}}) = \underset{(G_{ll}, \tilde{G})}{argmax} [p(Y_l | G_{ll}, \tilde{G}) p(G_{ll}, \tilde{G})], \tag{8}$$

or equivalently

$$(\hat{G}_{ll}, \hat{\tilde{G}}) = \underset{(G_{ll},\tilde{G})}{argmax}[\ln p(Y_l|G_{ll},\tilde{G}) + \ln p(G_{ll},\tilde{G})]. \quad (9)$$

Since the noise is AWGN, PDF of received signal Y_l can be expressed as

$$p(Y_l|G_{ll},\tilde{G}) \propto exp[- \parallel Y_l - \sqrt{p_u}\Phi_l G_{ll} - \sqrt{p_u}\tilde{\Phi}\tilde{G} \parallel^2]. \quad (10)$$

The prior PDF of the channel coefficients can be expressed as

$$p(G_{ll},\tilde{G}) \propto \exp(-\sum_{i=1}^{L} G_{il}^H D_{il}^{-1} G_{il}). \quad (11)$$

For notational simplicity superior equation can be written as

$$p(G_{ll},\tilde{G}) \propto \exp(-G_{ll}^H D_{ll}^{-1} G_{ll} - \tilde{G}^H \tilde{D}^{-1}\tilde{G}), \quad (12)$$

where $\tilde{D} = diag[D_{il}, ...D_{Ll}] \in \mathbb{R}^{(L-1)K\times(L-1)K}, \forall i \neq l$.

The log-likelihood function of received signal and prior information of channel can be written as

$$\ln p(Y_l|G_{ll},\tilde{G}) = - \parallel Y_l - \sqrt{p_u}\Phi_l G_{ll} - \sqrt{p_u}\tilde{\Phi}\tilde{G} \parallel^2, \quad (13)$$

$$\ln p(G_{ll},\tilde{G}) = -G_{ll}^H D_{ll}^{-1} G_{ll} - \tilde{G}^H \tilde{D}^{-1}\tilde{G}. \quad (14)$$

Taking derivatives in eq. (9) with respect to G_{ll} and $\tilde{G}$ and equating the resulting equations to zero, we have

$$\begin{bmatrix} \Phi_l^H\Phi_l + \frac{D_{ll}^{-1}}{p_u} & \Phi_l^H\tilde{\Phi} \\ \tilde{\Phi}^H\Phi_l & \tilde{\Phi}^H\tilde{\Phi} + \frac{\tilde{D}^{-1}}{p_u} \end{bmatrix} \begin{bmatrix} G_{ll} \\ \tilde{G} \end{bmatrix} = \frac{1}{\sqrt{p_u}} \begin{bmatrix} \Phi_l^H \\ \tilde{\Phi}^H \end{bmatrix} Y_l. \quad (15)$$

Since the pilot sequences are orthogonal and same in all cells, then the superior equation becomes

$$\hat{G}_{ll}^{map} = \sqrt{p_u} D_{ll} \left(\mathbf{I_K} + \tau p_u \sum_{i=1}^{L} D_{il} \right)^{-1} \Phi_l^H Y_l. \quad (16)$$

Since $p(Y_l|G_{ll},\tilde{G})$ is jointly Gaussian in G_{ll}, $\tilde{G}$ and Y_l, the MAP estimate is equal to MMSE estimate for the proposed model [9].

IV. Iterative Channel Estimation

By expectation-maximization (EM) algorithm in each iteration, all parameters which are unknown calculated simultaneously and the parameters iteratively converge to the LS estimate of the channel coefficients. EM algorithm decreases computational complexity of the system [10]. To reduce the convergence time of the EM algorithm SAGE algorithm is recommended by [11]. SAGE algorithm preserves the stability of EM algorithm and able to improve the convergence rate significantly. Instead of estimating all channel coefficients at once, SAGE algorithm calculates the channel coefficients sequentially. Only a subset of unknown channel coefficients is calculated at each iteration.

Y_k represents the received data component which is transmitted by the *k*th user through the channel g_{ilkm} as follows

$$Y_k = \sqrt{p_u}\sum_{i=1}^{L} \phi_{ik} g_{ilkm} + N_k, \quad (17)$$

where the Gaussian noise matrix N_k symbolizes a part of N_l and it is defined by $\sum_{k=1}^{K} N_k = N_l$ and ϕ_{ik} is the *k*th column of pilot matrix.

For notational simplicity, the desired channel matrix can be represented as $g_{llkm} \in \mathbb{C}^{1\times M}$ and the pilot sequence can be represented as $\phi_{lk} \in \mathbb{C}^{\tau\times 1}$, all interfering cell pilot sequences are collected in $\tilde{\phi} = [\phi_{1k}, ..., \phi_{Lk}] \in \mathbb{C}^{\tau\times(L-1)}, \forall i \neq l$ and channel matrices collected in $\tilde{g} = [g_{1lkm}^T, ..., g_{Llkm}^T]^T \in \mathbb{C}^{(L-1)\times M}, \forall i \neq l$ leading to

$$Y_k = \sqrt{p_u}\phi_{lk} g_{llkm} + \sqrt{p_u}\tilde{\phi}\tilde{g} + N_k. \quad (18)$$

The incomplete (original) received data at the *l*th BS can be written as

$$Y_l = \sum_{k=1}^{K} Y_k. \quad (19)$$

A. SAGE Algorithm Based MAP Channel Estimation

A drawback of directly solving eq. (9) is calculating conjugate transpose of $\tau \times K$ pilot matrix. For large values of τ and K computational complexity increases extremely. When SAGE algorithm based MAP estimator is applied to the system, the computational complexity decreases significantly and the proposed method achieves the same MSE performance as the MAP estimator after a number of iterations.

The algorithm of the proposed method can be written as follows

Initialization: For $1 \leq k \leq K$

$$\hat{z}_k^{(0)} = \frac{1}{\tau\sqrt{p_u}} \beta_{llk}^{-1} (1 + \tau p_u \sum_{i=1}^{L} \beta_{ilk}) \phi_{lk} \hat{g}_{llkm}^{(0)}. \quad (20)$$

At the *q*th iteration (q=0,1,2...): For $k = 1 + mod[q, K]$, compute

$$\hat{Y}_k^{(q)} = \hat{z}_k^{(q)} + [Y_l - \sum_{k=1}^{K} \hat{z}_k^{(q)}]. \quad (21)$$

$$(\hat{g}_{llkm}^{(q+1)}, \tilde{g}^{(q+1)}) = \underset{(g_{llkm},\tilde{g})}{argmax}[p(g_{llkm}, \tilde{g}|\hat{Y}_k^{(q)})]. \quad (22)$$

To find the MAP estimator, we must equivalently maximize $p(Y_k^{(q)}|g_{llkm},\tilde{g})p(g_{llkm},\tilde{g})$. Hence, the MAP estimator is

$$(\hat{g}_{llkm}^{(q+1)}, \hat{\tilde{g}}^{(q+1)}) = \underset{(g_{llkm},\tilde{g})}{argmax}[p(Y_k^{(q)}|g_{llkm},\tilde{g})p(g_{llkm},\tilde{g})], \quad (23)$$

or equivalently

$$(\hat{g}_{llkm}^{(q+1)}, \hat{\tilde{g}}^{(q+1)}) = \underset{(g_{llkm}, \tilde{g})}{argmax}[\ln p(\hat{Y}_k^{(q)}|g_{llkm}, \tilde{g}) + \ln p(g_{llkm}, \tilde{g})]. \quad (24)$$

since the noise is AWGN, PDF of Y_k can be written as

$$p(\hat{Y}_k^{(q)}|g_{llkm}, \tilde{g}) \propto exp[- \parallel \hat{Y}_k^{(q)} - \sqrt{p_u}\phi_{lk} g_{llkm} - \sqrt{p_u}\tilde{\phi}\tilde{g} \parallel^2]. \quad (25)$$

The prior PDF of the channel coefficients can be written as

$$p(g_{llkm}, \tilde{g}) \propto exp[-\sum_{i=1}^{L} g_{ilkm}^H \beta_{ilk}^{-1} g_{ilkm}]. \quad (26)$$

For notational simplicity superior equation can be written as

$$p(g_{llkm}, \tilde{g}) \propto exp[-g_{llkm}^H \beta_{llk}^{-1} g_{llkm} - \tilde{g}^H \tilde{\beta}^{-1} \tilde{g}], \quad (27)$$

where $\tilde{\beta} = diag[\beta_{1lk}, ..., \beta_{Llk}] \in \mathbb{R}^{(L-1)\times(L-1)}, \forall i \neq l$.

The log-likelihood function of Y_k and prior information of channel can be written as

$$\ln p(\hat{Y}_k^{(q)}|g_{llkm}, \tilde{g}) = - \parallel \hat{Y}_k^{(q)} - \sqrt{p_u}\phi_{lk} g_{llkm} - \sqrt{p_u}\tilde{\phi}\tilde{g} \parallel^2, \quad (28)$$

$$\ln p(g_{llkm}, \tilde{g}) = -g_{llkm}^H \beta_{llk}^{-1} g_{llkm} - \tilde{g}^H \tilde{\beta}^{-1} \tilde{g}. \quad (29)$$

Taking derivatives in eq. (24) with respect to g_{llkm}, $\tilde{g}$ and equating the resulting equations to zero, we have

$$\begin{bmatrix} \phi_{lk}^H \phi_{lk} + \frac{\beta_{llk}^{-1}}{p_u} & \phi_{lk}^H \tilde{\phi} \\ \tilde{\phi}^H \phi_{lk} & \tilde{\phi}^H \tilde{\phi} + \frac{\tilde{\beta}^{-1}}{p_u} \end{bmatrix} \begin{bmatrix} g_{llkm} \\ \tilde{g} \end{bmatrix} = \frac{1}{\sqrt{p_u}} \begin{bmatrix} \phi_{lk}^H \\ \tilde{\phi}^H \end{bmatrix} \hat{Y}_k^{(q)}. \quad (30)$$

since pilot sequences are orthogonal and same in all cells, then the superior equation becomes

$$\hat{g}_{llkm}^{(q+1)} = \sqrt{p_u}\beta_{llk} \left(1 + \tau p_u \sum_{i=1}^{L} \beta_{ilk}\right)^{-1} \phi_{lk}^H \hat{Y}_k^{(q)}. \quad (31)$$

As we see in the superior equation, since ϕ_{lk} is a vector, proposed method does not require conjugate transpose of the large-size matrix and consequently, reduces computational complexity significantly.

$$\hat{z}_k^{(q+1)} = \sqrt{p_u}\phi_{lk}\hat{g}_{llkm}^{(q+1)}. \quad (32)$$

For $1 \leq j \leq K$ and $j \neq k$

$$\hat{z}_j^{(q+1)} = \hat{z}_j^{(q)}. \quad (33)$$

B. Initialization

The initial channel coefficients ($\hat{g}_{llkm}^{(0)}$) of SAGE algorithm are important because they determine the convergence speed of the algorithm. For the proposed algorithm the initial channel coefficients are selected as random complex values.

C. Complexity Calculation

In this section, a short description of the computational complexity of the proposed SAGE-MAP estimator and MAP estimator is given. The MAP channel estimator given in (16) requires $K^2\tau^2 + K\tau M + 2K^2M + K$ complex multiplications. On the other hand, the SAGE-MAP estimator requires $q\tau^2 + K\tau M + 2q\tau M$ complex multiplication. Therefore, it is shown that the computational complexity of MAP estimator is higher than SAGE-MAP estimator. If it is defined the total numbers of multiplication required for the MAP and SAGE-MAP estimator as MAP_{Total} and $SAGE-MAP_{Total}$ respectively, then complexity rate can be summarized as given in Table 1.

TABLE I
COMPLEXITY TABLE

K	τ	M	q	$SAGE-MAP_{Total}/MAP_{Total}$
10	10	60	9	0.63
10	10	100	9	0.72
10	20	60	9	0.58
10	20	100	9	0.74

V. SIMULATIONS

In this section, we compare that after how many iterations MSE performance of SAGE-based MAP estimator converges to the MSE performance of MAP estimator under different circumstances. We have a typical multi-cell structure with $L = 7$ cells and $\tau = K$ pilot symbols. We fix the large-scale fading coefficients (β_{ilk}) as $\beta_{llk} = 1$ and $\beta_{ilk} = a$, $\forall i \neq l$.

In fig. 3, we show the MSE performance of MAP estimator and convergence speed of SAGE-MAP estimator for a different number of iterations in a different shadow fading (a) values. We set $M = 100$, $K = 10$ and $pu = 10dB$. It can be seen from the figure that when $'a'$ (the effect of pilot contamination) increases the system performance of MAP estimator degrades and comes close to initial SAGE-MAP estimate values but the number of iterations does not change. For $a = 0.05$ the SAGE-MAP estimate converges to the MAP estimate within the 9 iterations and for $a = 0.03$ SAGE-MAP converges to the MAP estimate again within 9 iterations.

In fig. 4, we show the MSE performance of MAP estimator and convergence speed of SAGE-MAP estimator for a different number of iterations in different signal to noise ratio (SNR) values. We set $M = 100$, $K = 10$ and $a = 0.05$. It can be seen from the figure that for small SNR values the system performance of MAP estimator degrades and comes close to initial SAGE-MAP estimate values but the number of iterations does not change. For pu = 0dB the SAGE-MAP estimate converges to the MAP estimate within the 9 iterations and for pu = 20dB SAGE-MAP converges to the MAP estimate again within 9 iterations.

In fig. 5, we compare the MSE performance of MMSE estimator, MAP estimator and SAGE-based MAP estimator

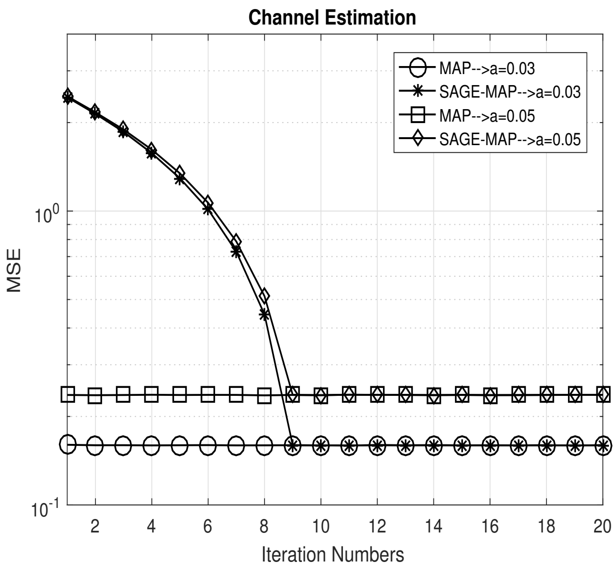

Fig. 3. Convergence of MSE versus number of iterations for different ”a” values.

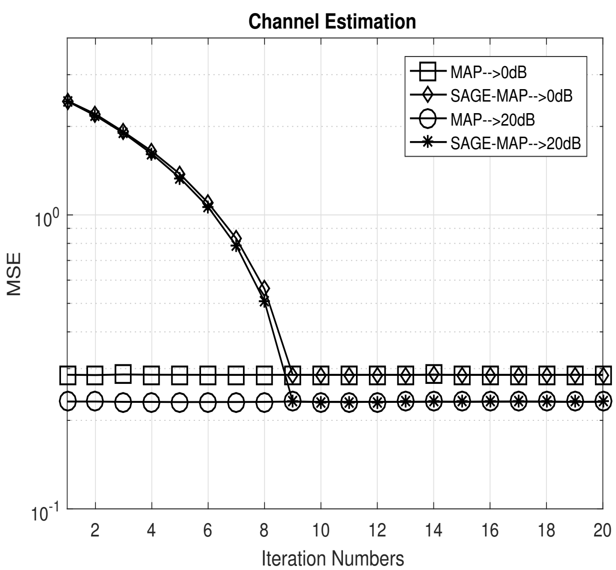

Fig. 4. Convergence of MSE versus number of iterations for different ”SNR” values. (pu = SNR)

for different SNR values. We set $M = 100$, $K = 10$, $a = 0.05$ and $iteration - number = 9$. It can be seen from the figure that MAP estimate, MMSE estimate and SAGE-based MAP estimate have the same MSE performance.

VI. Conclusion

In this paper, an efficient SAGE algorithm based MAP channel estimation method for multi-cell massive MIMO systems is derived. It is shown that the proposed method achieves the same MSE performance as MAP estimator with the increasing number of iterations. MAP estimator requires conjugate transpose of a large-size matrix, thus this situation increases the computational complexity. The proposed method avoids conjugate transpose of the large-size matrix and consequently decreases computational complexity significantly and coincides with the same MSE performance with MAP estimator.

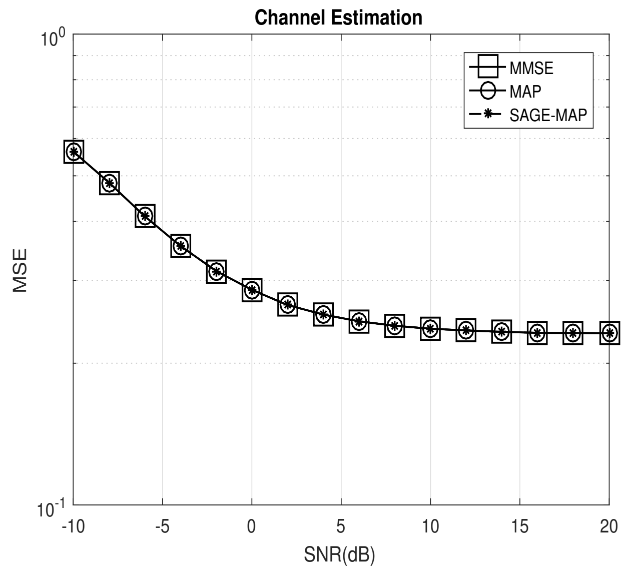

Fig. 5. MSE comparison of different channel estimation methods.

References

[1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An Overview of Massive MIMO: Benefits and Challenges,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, oct 2014.

[2] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, nov 2010.

[3] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, “A Comprehensive Survey of Pilot Contamination in Massive MIMO5G System,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 905–923, 2016.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, apr 2013.

[5] J. Jose and A. Ashikhmin, “Pilot contamination problem in multi-cell TDD systems,” *Theory, 2009. ISIT*, pp. 2184–2188, jun 2009.

[6] A. Khansefid and H. Minn, “On Channel Estimation for Massive MIMO with Pilot Contamination,” *IEEE Communications Letters*, vol. 19, no. 9, pp. 1660–1663, sep 2015.

[7] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot Contamination and Precoding in Multi-Cell TDD Systems,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, aug 2011.

[8] S. Kay, “Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1),” 1993.

[9] D. Neumann, M. Joham, and W. Utschick, “Channel Estimation in Massive MIMO Systems,” mar 2015.

[10] M. Feder and E. Weinstein, “Parameter Estimation of Superimposed Signals Using the EM Algorithm,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, apr 1988.

[11] J. A. Fessler and A. O. Hero, “Space-Alternating Generalized Expectation-Maximization Algorithm,” *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, 1994.