Semi-Blind Source Separation based on Multi-Modulus Criterion: Application for Pilot Contamination Mitigation in Massive MIMO Communications Systems

Ouahbi Rekik, Anissa Mokraoui
L2TI, Institut Galilée,
Université Paris 13 Sorbonne Paris Cité
Villetaneuse, France
firstname.lastname@univ-paris13.fr

Abdelhamid Ladaycia
L2TI, Institut Galilée,
Université Paris 13 Sorbonne Paris Cité
Villetaneuse, France
abdelhamid.ladaycia@univ-paris13.fr

Karim Abed-Meraim
PRISME
Université d'Orléans
Orléans, France
karim.abed-meraim@univ-orleans.fr

***Abstract*—This paper presents an efficient semi-blind approach for demixing instantaneous mixtures, in a massive Multiple-Input Multiple-Output (MIMO) communications system, with pilot contamination. A hybrid cost function is defined based on the Multi-Modulus (MM) criterion, for the blind part, and the Least Squares (LS) criterion for the pilot-based part. A simple but efficient semi-blind block gradient descent procedure is put forward, in which the step size, which globally minimizes the cost function along the search direction, is algebraically computed at each iteration for each user. Simulation results show that the use of pilots enhance the convergence and allow to mitigate the inherent ambiguity of fully-blind methods. Moreover, the optimal step size notably accelerates the convergence and can further reduce the impact of local extrema. Furthermore, the proposed approach is very promising for source recovering when pilot contamination occurs in massive MIMO communications systems.**

***Index Terms*—semi-blind source separation, multi-modulus, pilot contamination, Least squares (LS), Least mean Fourth (LF).**

I. Introduction

Source separation and channel equalization in digital communications aim at recovering the unknown data of the different users transmitted through a distorting propagation medium. With no training sequences nor prior knowledge on the channel, Blind Source Separations (BSS) is an efficient alternative that has been widely investigated in the literature [1] [2]. BSS processes the received signal based on a priori knowledge about the statistics or the nature of the transmitted signals, through the optimization of an appropriate cost function. Various BSS cost functions have been proposed in literature (e.g [1] and references therein) depending upon the type of source signals. Among them; the Constant-Modulus (CM) criterion for phase/frequency modulated signals such as PSK/FSK and Multi-Modulus (MM) criterion, where it has been shown that it outperforms the CM one for the case of square QAM constellation [3], which is used in many modern communications systems such as LTE and WiMAX. The optimization of the cost function can be performed through a closed-form solution, such as Analytical Constant Modulus Algorithm (ACMA) [4] and Analytical Constant Power Algorithm (ACPA) [5], when the channel accepts a noiseless AR model and the FIR equalizer is sufficiently long. Otherwise, an iterative solution can be adopted, such as the gradient descent technique, the Newton's method, or Givens ans Shear's rotations-based techniques [6].

On the other hand, some pilot sequences are often available in most communications technologies, thus, exploiting this available information should notably improve the source recovering performance of BSS by incorporating a pilot-based Least Squares (LS) criterion in a semi-blind scheme. In particular, this approach is shown to be an efficient solution to the pilot contamination problem in massive MIMO systems (e.g. [7]).

In this context, this paper proposes a semi-blind source separation technique for instantaneous mixtures, which can help mitigate the problem of pilot contamination. A hybrid cost function is defined based on the MM criterion for the unknown data and on the LS criterion for the pilots. The latter helps to overcome the inherent ambiguity of the blind process. An iterative-based minimization of the aforementioned cost function is performed through the gradient descent rule. A batch-based full estimation procedure is adopted so that, all sources are separated simultaneously. Moreover, for improving the convergence speed, an optimized step-size procedure is introduced.

II. Communications System Model and Problem Formulation

Consider a MIMO system consisting of N_t sources (transmitters), each having a single antenna, and a receiver equipped with N_r antennas. All sources transmit their signals over the same band of frequencies. Each transmitted source signal is drawn from an M-ary square QAM constellation, then passed through a flat fading channel represented by an unknown mixing matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, which is assumed to be of full column rank so that $N_r > N_t$. Thus, for the case of instantaneous mixtures, the noisy received signal is given by:

$$\mathbf{y}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k), \tag{1}$$

where $\mathbf{y}(k) = [y_1(k), ..., y_{N_r}(k)]^T$; $\mathbf{s}(k) = [s_1(k), ..., s_{N_t}(k)]^T$ refers to the transmitted signals and $\mathbf{n}(k) = [n_1(k), ..., n_{N_r}(k)]^T$ is an additive white Gaussian noise with a covariance matrix $\sigma_n^2 \mathbf{I}_{N_r}$.

The objective of source separation and channel equalization is to recover the transmitted data symbols, by applying a separation matrix $\mathbf{W} \in \mathbb{C}^{N_r \times N_t}$ to the observed (received) signals as follows: $\mathbf{z}(k) = \mathbf{W}^H \mathbf{y}(k)$. It can be noticed that each row $\mathbf{w}_i$ can extract one source signal. In batch processing approach, N_s samples of the received data are collected before processing, so that the matrix formulation of the problem is given by:

$$\mathbf{Y} = \mathbf{H}\mathbf{S} + \mathbf{N}, \quad \mathbf{Z} = \mathbf{W}^H \mathbf{Y}, \tag{2}$$

where the received signals are given by $\mathbf{Y} \in \mathbb{C}^{N_r \times N_s}$; the transmitted signals are represented by $\mathbf{S} \in \mathbb{C}^{N_t \times N_s}$; the additive white noise is given by $\mathbf{N} \in \mathbb{C}^{N_r \times N_s}$; whereas $\mathbf{Z} \in \mathbb{C}^{N_t \times N_s}$ is the estimated signals. Moreover, in semi-blind approaches, both pilots and data are used, hence, without loss of generality, the pilots are assumed to appear at the beginning of the transmitted frames in a block-type arrangement, thus each frame is formed by N_p pilots followed by N_d data samples, so that $N_s = N_p + N_d$ and $\mathbf{Y} = [\mathbf{Y}_p, \mathbf{Y}_d]$.

III. Semi-blind source separation

In this paper, the semi-blind source separation approach is based on the MM criterion for the unknown data and on the LS criterion for the pilots. Indeed to take advantage of both pilots and data, a hybrid cost function denoted $J_{\text{SB}}(\mathbf{W})$ is defined as follows:

$$J_{\text{SB}}(\mathbf{W}) = (1 - \alpha) J_{\text{B}}(\mathbf{W}) + \alpha J_{\text{LS}}(\mathbf{W}), \tag{3}$$

where $J_{\text{B}}(\mathbf{W})$ stands for a fully blind cost function; $J_{\text{LS}}(\mathbf{W})$ refers to the use of pilots; and α is a real constant, taking values in the interval $[0, 1]$, considered as the weight given to the blind and the training-based parts of the semi-blind cost function.

In what follows, an iterative method based on the gradient descent is adopted to minimize the cost function, given by (3), according to:

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \mu \mathbf{G}_n, \quad n = 0, 1..., \tag{4}$$

where $\mathbf{W}_{n+1}$ (respectively $\mathbf{W}_n$) represents the updated (respectively the old) value of the matrix $\mathbf{W}$; μ is a small positive value, called step size, that determines the speed of convergence; and $\mathbf{G}_n$ is the gradient of the cost function, at the n-th iteration, given by:

$$\mathbf{G}_n = \nabla J_{\text{SB}}(\mathbf{W}_n) = (1 - \alpha) \nabla J_{\text{B}}(\mathbf{W}_n) + \alpha \nabla J_{\text{LS}}(\mathbf{W}_n). \tag{5}$$

As adopted in [5], the iterative procedure is stopped as soon as:

$$\frac{\|\mathbf{W}_{n+1} - \mathbf{W}_n\|}{\|\mathbf{W}_n\|} < \frac{0.1\mu}{\sqrt{N_s}}. \tag{6}$$

Also, a maximum number of iterations can be defined for stopping the iteration process.

In the current work, we adopt a batch processing based on the use of block iterative implementation, as opposed to stochastic algorithms. The latter approaches approximate the gradient by using a one-sample estimate, which leads to dropping the expectation operator. Consequently, these methods generally lead to a slow convergence. By contrast, batch-based methods approximate the gradient from a block of the received samples repeatedly at each iteration. This more precise gradient estimate improves convergence speed and accuracy [8]. Moreover, all the sources are simultaneously estimated, so that the accumulated errors of the deflation-based methods are avoided [9].

A. Multi-Modulus criterion

In the current work, the blind process is based on the MM criterion, which penalizes the deviation of the real and imaginary parts of the equalized signals from the squared constellation shape as follows:

$$J_{\text{MM}}(\mathbf{W}) = \sum_{i=1}^{N_t} \frac{1}{N_d} \sum_{k=N_p+1}^{N_s} [(z_{i,R}^2(k) - R_R)^2 + (z_{i,I}^2(k) - R_I)^2], \tag{7}$$

where $z_{i,R} = real(\mathbf{w}_i^H \mathbf{Y}_d)$ (respectively $z_{i,I} = imag(\mathbf{w}_i^H \mathbf{Y}_d)$) is the real (respectively imaginary) part of the (i,k)-th element of the recovered signal; $R_R = E[s_R^4(k)]/E[s_R^2(k)]$ and $R_I = E[s_I^4(k)]/E[s_I^2(k)]$ are the real and imaginary dispersion constants.

The gradient of the MM criterion, for the i-th user, is defined as follows:

$$\begin{aligned}\nabla J_{\text{MM}}(\mathbf{w}_i) = \frac{1}{N_d} \sum_{k=N_p+1}^{N_s} [\mathbf{y}_d(k) \big((z_{i,R}^2(k) - R_R) z_{i,R}(k) \\ - j(z_{i,I}^2(k) - R_I) z_{i,I}(k)\big)].\end{aligned} \tag{8}$$

In what follows, the CM criterion is used for comparison. It is given by [4]:

$$J_{\text{CM}}(\mathbf{W}) = \sum_{i=1}^{N_t} \frac{1}{N_d} \sum_{k=N_p+1}^{N_s} (\mid z_i(k) \mid^2 - R)^2. \tag{9}$$

The gradient of the CM criterion, for each user, is defined as follows:

$$\nabla J_{\text{CM}}(\mathbf{w}_i) = \frac{1}{N_d} \sum_{k=N_p+1}^{N_s} [\mathbf{y}_d(k) \big(z_i(k)^* (\mid z_i(k) \mid^2 - R)\big)]. \tag{10}$$

Note that the multiplicative constants of the previous gradient formulas are omitted because they are absorbed by the step size as given in (4).

B. Pilot-based criterion

In the current work, the use of the pilots is through the LS criterion, which is based on the error between the transmitted pilot symbols and their estimates, according to:

$$J_{\text{LS}}(\mathbf{W}) = \sum_{i=1}^{N_t} \frac{1}{N_p} \sum_{k=1}^{N_p} \mid z_i(k) - s_{p_i}(k) \mid^2, \tag{11}$$

where $z_i(k)$ and $s_{p_i}(k)$ stand for the k-th estimated and transmitted pilot sample of the i-th source.

The gradient of the pilot-based LS criterion, for the i-th user, is defined as follows:

$$\nabla J_{\text{LS}}(\mathbf{w}_i) = \frac{1}{N_p} \sum_{k=1}^{N_p} [\mathbf{y}_p(k) \left((z_i(k) - s_{p_i}(k))^*\right)]. \tag{12}$$

It is clear that the MM and CM functions are 4-th order ones whereas the LS is a quadratic function. In some works (e.g. [10] and references therein), the 4-th order CM cost function is approximated by a quadratic function. To get a 'homogeneous' hybrid criterion, an alternative pilot-based cost function would be the Least mean Fourth (LF) [11] given by:

$$J_{\text{LF}}(\mathbf{W}) = \sum_{i=1}^{N_t} \frac{1}{N_p} \sum_{k=1}^{N_p} \mid z_i(k) - s_{p_i}(k) \mid^4 . \tag{13}$$

As can be seen in the sequel, the latter cost function leads to a slight improvement of the source separation quality.

IV. Optimal step size

Exact line search optimization technique has been successfully used recently for optimizing the step size of the steepest-descent gradient-based algorithms for channel identification/equalization and independent component analysis (e.g. [5] [12]), where the update rule is expressed as given in (4) but with a variable step size. Indeed, it can be observed that $J_{SB}(\mathbf{W}_{n+1})$ is a polynomial function of the step size parameter μ, thus, it is possible to perform a steepest descent of the objective function by finding the optimal step size:

$$\mu_{opt} = arg \min_{\mu} J_{SB}(\mathbf{W}_n - \mu \mathbf{G}_n). \tag{14}$$

Consequently, μ_{opt} is the appropriate root of the derivative of $J_{SB}(\mathbf{W}_{n+1})$ w.r.t. μ, which is a 3^{rd}-degree polynomial given by:

$$\nabla J_{SB}(\mathbf{W}_{n+1}) = (1-\alpha) p_B(\mu) + \alpha p_{LS}(\mu), \tag{15}$$

where

$$\begin{aligned} p_B(\mu) = p_{MM}(\mu) =& (\boldsymbol{\beta}_{3_R} + \boldsymbol{\beta}_{3_I})\mu^3 + (\boldsymbol{\beta}_{2_R} + \boldsymbol{\beta}_{2_I})\mu^2 + \\ & (\boldsymbol{\beta}_{1_R} + \boldsymbol{\beta}_{1_I})\mu + \boldsymbol{\beta}_{0_R} + \boldsymbol{\beta}_{0_I}, \end{aligned} \tag{16}$$

so that

$$\begin{aligned} \boldsymbol{\beta}_{3_R} &= \frac{2}{N_d} \sum_{k=N_p+1}^{N_s} (\mathbf{a}_R^2), \quad \boldsymbol{\beta}_{2_R} = \frac{3}{N_d} \sum_{k=N_p+1}^{N_s} (\mathbf{a}_R \mathbf{b}_R), \\ \boldsymbol{\beta}_{1_R} &= \frac{2}{N_d} \sum_{k=N_p+1}^{N_s} (\mathbf{a}_R \mathbf{c}_R + \mathbf{b}_R^2), \\ \boldsymbol{\beta}_{0_R} &= \frac{1}{N_d} \sum_{k=N_p+1}^{N_s} (\mathbf{b}_R \mathbf{c}_R), \end{aligned} \tag{17}$$

and

$$\begin{aligned} &\mathbf{a}_R = \mid \mathbf{G_r} \mid^2, \quad \mathbf{b}_R = -2real\{\mathbf{W_r} \mathbf{G_r}^H\}, \\ &\mathbf{c}_R = \mid \mathbf{W_r} \mid^2 - R, \end{aligned} \tag{18}$$

$$\mathbf{G_r} = real\{\mathbf{G}^H \mathbf{Y}_d\} \quad \text{and} \quad \mathbf{W_r} = real\{\mathbf{W}^H \mathbf{Y}_d\}.$$

For $\boldsymbol{\beta}_{x_I}, x = 0, ...3$ the same equations used in (17) still valid but with $\mathbf{G_I} = imag\{\mathbf{G}^H \mathbf{Y}_d\}$ and $\mathbf{W_I} = imag\{\mathbf{W}^H \mathbf{Y}_d\}$.
Also we have

$$p_{LS}(\mu) = \alpha_1 \mu + \alpha_0, \tag{19}$$

where

$$\alpha_1 = \frac{1}{N_p} \sum_{k=1}^{N_p} (\mid \mathbf{G}^H \mathbf{Y}_p \mid^2), \tag{20}$$

$$\alpha_0 = \frac{-1}{N_p} \sum_{k=1}^{N_p} (real\{\mathbf{G}^H \mathbf{Y}_p^* (\mathbf{Y}_p - \mathbf{S}_p)\}), \tag{21}$$

where $\mathbf{S}_p$ contains the pilots of all users.

Remark: For the CM criterion, an optimal step size can be calculated by considering the polynomial:

$$p_{CM}(\mu) = \boldsymbol{\beta}_{3_C} \mu^3 + \boldsymbol{\beta}_{2_C} \mu^2 + \boldsymbol{\beta}_{1_C} \mu + \boldsymbol{\beta}_{0_C}, \tag{22}$$

where $\boldsymbol{\beta}_{3_C}, \boldsymbol{\beta}_{2_C}, \boldsymbol{\beta}_{1_C}$ and $\boldsymbol{\beta}_{0_C}$ are calculated in the same way as in (17) but by considering $\mathbf{G_C} = \mathbf{G}^H \mathbf{Y}_d$ and $\mathbf{W_C} = \mathbf{W}^H \mathbf{Y}_d$.

Finally, μ_{opt} is chosen as the real-valued root that minimizes the cost function $J_{SB}(\mathbf{W}_{n+1})$.

It is important to note that an optimal step size can be calculated for each user, so that, the upgrading rule can be expressed as follows:

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \mathbf{G}_n diag(\boldsymbol{\mu}), \tag{23}$$

where $\boldsymbol{\mu} = [\mu_1, ..., \mu_{N_t}]$ contains all the optimal step size values of all users.

V. SB-MM source separation under the effect of pilot contamination

This section considers a multi-cell massive MIMO system composed of N_c cells. In such a case, the system model is given by:

$$\mathbf{y}(k) = \sum_{l=1}^{N_c} \mathbf{H}_l \mathbf{s}_l(k) + \mathbf{n}(k), \tag{24}$$

where the first cell represents the cell of interest while the others are the interfering neighboring cells.

Traditionally, a channel estimation is performed before recovering the transmitted data during the uplink phase according to the Time Division Duplexing (TDD) protocol. However, pilots of the cell of interest are interfering with pilots of adjacent cells leading to the phenomenon of pilot contamination [7]. Recent work (e.g. [13]) show that semi-blind approaches represent a potential solution to this problem. Thus, we propose to recover the data of users located in the cell of interest by means of the proposed semi-blind source separation method. Interestingly, exploiting the pilots, through the LS or LF criterion included in the semi-blind cost function (i.e. equation (3)), allows to recover only the data of the cell of interest, from the observed mixture signal, and removing the inherent ambiguity of the blind-based source separation techniques.

Indeed, massive MIMO systems are characterized by their large size antennas, so that we assume that N_r is larger than the total number of users, i.e. $N_{t_1} + \cdots + N_{t_{N_c}}$, coming from the N_c cells and hence the global system matrix $\mathbf{H}$ is left invertible. In that case, the target separation matrix $\mathbf{W}$ is of size $N_r \times N_{t_1}$.

VI. Performance analysis and discussions

This section highlights the performance of the proposed semi-blind source separation and its effectiveness under the effect of pilot contamination. The unknown data is drawn from a 4-QAM modulation with equal weights for the blind and pilot-based criteria, i.e. $\alpha = 0.5$ and a fixed step size $\mu = 0.01$. The results are averaged over 500 Monte Carlo runs, where the channel is randomly generated (with i.i.d. normal Gaussian entries) at each run. The performance is assessed through an average Symbol Error Rate (SER).

As a reference, the conventional LS solution (with only pilots) is calculated using the available training pilots: $\mathbf{W}_{LS} = \mathbf{Y}_p^{\#} \mathbf{S}_p^H$, where $\mathbf{Y}_p^{\#} = (\mathbf{Y}_p \mathbf{Y}_p^H)^{-1} \mathbf{Y}_p$ is the pseudo inverse of the received pilots $\mathbf{Y}_p$. Accordingly, we refer to the LS solution with both data and pilots used for training as the "MMSE lower bound" introduced here just for benchmarking.

Figure 1 illustrates the performance of the blind MM Algorithm (MMA-B), the blind CM Algorithm (CMA-B), the semi-blind MM Algorithm (MMA-SB), and the semi-blind CM Algorithm (CMA-SB) for source separation, in a mono-cell system, with separating matrix randomly initialized. One can notice the enhancement, in terms of SER, of the semi-blind approaches compared to the blind ones. Moreover, better performance is obtained with an optimal step size (MMA-SB-OSS, CMA-SB-OSS) compared to the fixed step size case. By taking into account the nature of the cost function, the MM-based approaches outperform slightly the CM-based ones, since the former takes into account the phase variation.

Since some training pilots are available, an LS-based initialization is used for the separating matrix. In this context the performance of the different approaches are illustrated in Figure 2. One can observe that the blind approaches have been clearly enhanced (compared to the results given in Figure 1), whereas a local minimum convergence problem is observed for the semi-blind ones with a fixed step size at high SNR. By contrast, one can notice that the performance with optimal step size is virtually independent of the initialization, while dramatically reducing the iterations number as illustrated in Table I. Notice that this reduction compensates the additional consumed time of calculating the optimal step size.

As explained in section III-B, a LF criterion is adopted to obtain two homogeneous 4-th order cost functions (i.e. the MM and the LF). As shown in Figure 3, an enhancement of the semi-blind performance with LF (MMA-SB4 and CMA-SB4) is observed for medium and high SNRs as compared to the case of LS (MMA-SB and CMA-SB).

Remark: note that in figures 1, 2 and 3 the very small values of SER ($<< 10^{-5}$) are not plotted.

Figure 4 illustrates the behavior of the semi-blind approaches when increasing the number of pilots, for a given SNR=$5dB$. one can observe that with very few pilots (< 5 in our case), the fully blind case performs better than the semi-blind one as if using too few pilot symbols could 'confuse' the semi-blind approaches and leads to ill convergence. Similar effect has been observed for the semi-blind equalization in [5]. However, For a reasonable number of pilots, the semi-blind methods are able to attain the MMSE lower bound while maintaining good spectral efficiency and an effective data rate.

Figure 5 illustrates the behavior of the blind and semi-blind approaches when increasing the number of data symbols N_d (for MMSE lower bound N_d is kept fixed as reference). It can be clearly shown that with higher number of data symbols (around $N_d = 100$), for a given SNR=5dB, the performance of the blind and semi-blind source separation are enhanced. This result allows to use a small pilot/data size ratio for better throughout.

Figure 6 illustrates the behavior of the proposed solution under the effect of pilot contamination. One can observe that the LS-based equalizers are severely affected by the problem of pilot contamination (LS-contam) compared to the single cell case (LS). However, such a problem can be alleviated by the semi-blind approach especially with MM criterion as compared to CM one (MMA-SB-contam, CMA-SB-contam). This result is very promising for pilot contamination mitigation in massive MIMO communications systems.

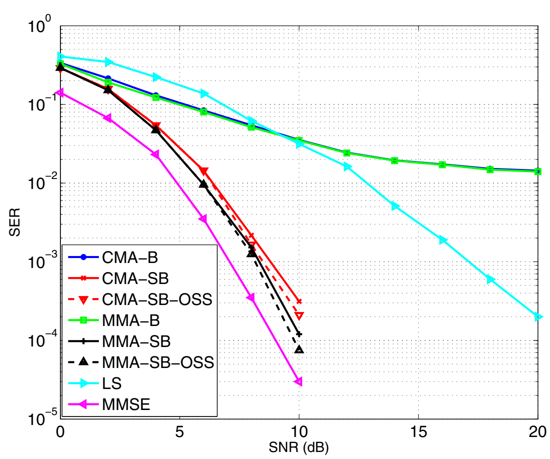

Fig. 1. Average SER vs SNR for blind and semi-blind source separation with random initialization ($N_r = 20, N_t = 4, N_d = 100, N_p = 5$)

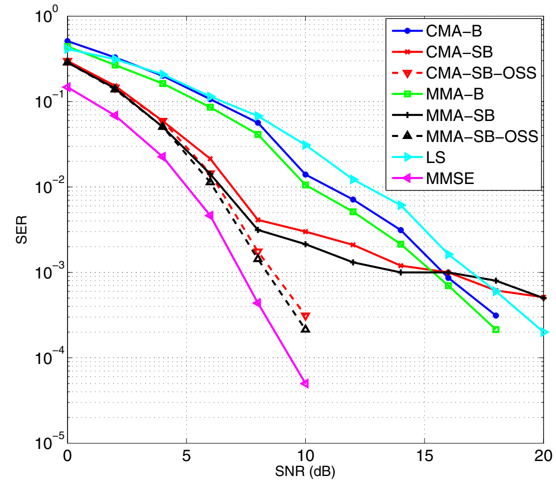

Fig. 2. Average SER vs SNR for blind and semi-blind source separation with pilot-based initialization ($N_r = 20, N_t = 4, N_d = 100, N_p = 5$)

TABLE I
AVERAGE NUMBER OF ITERATIONS FOR CONVERGENCE FOR DIFFERENT SCENARIOS

	B-CMA	SB-CMA	B-MMA	SB-MMA
Fixed Step Size	482	677	616	783
Optimized Step Size	280	392	350	460

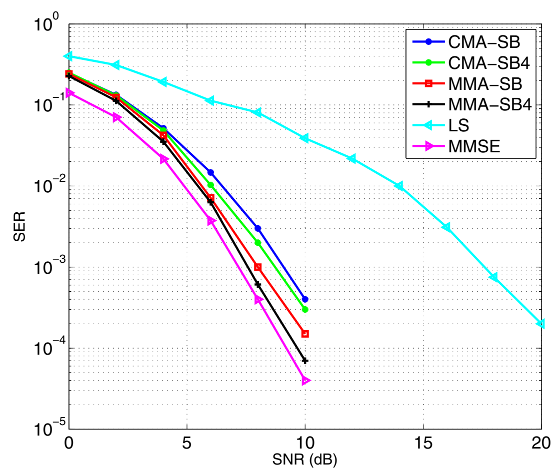

Fig. 3. Average SER vs SNR with LS and LF ($N_r = 20, N_t = 4, N_d = 100, N_p = 5$)

VII. Conclusion

This paper proposed a new approach for semi-blind source separation, which can help mitigate the problem of pilot contamination in massive MIMO communications systems. A hybrid cost function is defined based on the MM criterion for the blind part and the LS or LF criterion for the pilot-based one. A full estimation procedure has been adopted based on the gradient descent rule and an optimized step size. Simulation results shown that the proposed method accelerates the convergence and can further reduce the negative impact of local minima. Moreover, under the effect of pilot contamination, the

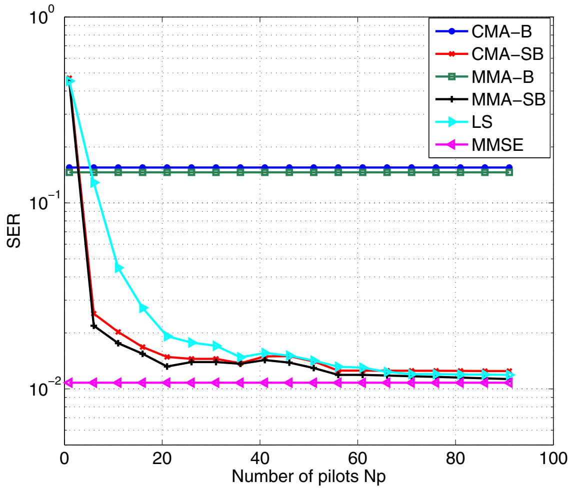

Fig. 4. Average SER vs number of pilots N_p for blind and semi-blind source separation $(N_r = 20, N_t = 4, N_d = 100, SNR = 5dB)$

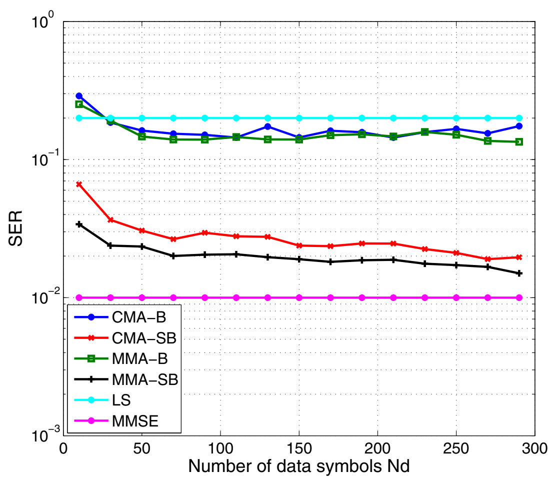

Fig. 5. Average SER vs number of data symbols N_d for blind and semi-blind source separation $(N_r = 20, N_t = 4, N_p = 5, SNR = 5dB)$

proposed method allowed to successfully recover only the data of the cell of interest, from the observed mixture signal. Consequently, such an approach reduces the computational cost at the receiver by avoiding channel estimation and data recovering of signals from the adjacent cells.

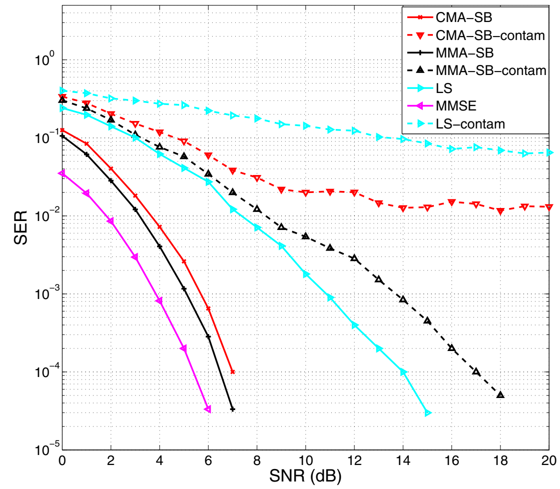

Fig. 6. Average SER vs SNR under pilot contamination with asynchronous cells $(N_r = 20, N_t = 2$ per cell, $N_c = 4, N_d = 100, N_p = 5)$

References

[1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[2] F. R. Pavan, M. T. Silva, and M. D. Miranda, “A numerically robust blind equalization scheme applied to MIMO communications systems,” *Journal of the Franklin Institute*, vol. 355, no. 1, pp. 596–624, 2018.

[3] S. Abrar and A. K. Nandi, “Blind equalization of square-QAM signals: a multimodulus approach,” *IEEE Transactions on Communications*, vol. 58, no. 6, 2010.

[4] A.-J. Van Der Veen and A. Paulraj, “An analytical constant modulus algorithm,” *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, 1996.

[5] V. Zarzoso and P. Comon, “Blind and semi-blind equalization based on the constant power criterion,” *IEEE Transactions on Signal Processing*, vol. 53, no. 11, p. 4363, 2005.

[6] S. A. W. Shah, K. Abed-Meraim, and T. Y. Al-Naffouri, “Multi-modulus algorithms using hyperbolic and givens rotations for blind deconvolution of MIMO systems,” in *International Conference on Acoustics Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 2155–2159.

[7] 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, 2010.

[8] P. Comon, “Block methods for channel identification and source separation,” in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*. IEEE, 2000, pp. 87–92.

[9] N. Fatema, Y. Xiang, and I. Natgunanathan, “Analysis of a semi-blind pilot decontamination method in massive MIMO,” in *27th International Telecommunication Networks and Applications Conference (ITNAC)*. IEEE, 2017, pp. 1–6.

[10] X. Hong and S. Chen, “Recursive least squares semi-blind beamforming for MIMO using decision directed adaptation and constant modulus criterion,” *International Journal of Automation and Computing*, vol. 14, no. 4, pp. 442–449, 2017.

[11] A. Zerguine, “Convergence and steady-state analysis of the normalized least mean fourth algorithm,” *Digital Signal Processing*, vol. 17, no. 1, pp. 17–31, 2007.

[12] V. Zarzoso and P. Comon, “Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size,” *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.

[13] D. Hu, L. He, and X. Wang, “Semi-blind pilot decontamination for massive MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 525–536, 2016.