

Blind Source Separation Algorithms Using Hyperbolic and Givens Rotations for High-Order QAM Constellations

Syed A. W. Shah, Karim Abed-Meraim, *Senior Member, IEEE*, and Tareq Y. Al-Naffouri*, *Member, IEEE*,

Abstract—This paper addresses the problem of blind demixing of instantaneous mixtures in a multiple-input multiple-output communication system. The main objective is to present efficient blind source separation (BSS) algorithms dedicated to moderate or high-order QAM constellations. Four new iterative batch BSS algorithms are presented dealing with the multimodulus (MM) and alphabet matched (AM) criteria. For the optimization of these cost functions, iterative methods of Givens and hyperbolic rotations are used. A pre-whitening operation is also utilized to reduce the complexity of design problem. It is noticed that the designed algorithms using Givens rotations gives satisfactory performance only for large number of samples. However, for small number of samples, the algorithms designed by combining both Givens and hyperbolic rotations compensate for the ill-whitening that occurs in this case and thus improves the performance. Two algorithms dealing with the MM criterion are presented for moderate order QAM signals such as 16-QAM. The other two dealing with the AM criterion are presented for high-order QAM signals. These methods are finally compared with the state of art batch BSS algorithms in terms of signal-to-interference and noise ratio, symbol error rate and convergence rate. Simulation results show that the proposed methods outperform the contemporary batch BSS algorithms.

Index Terms—blind source separation, constant modulus algorithm, multimodulus algorithm, constellation matched error, alphabet matched algorithm, Givens and hyperbolic rotations

I. INTRODUCTION

BLIND source separation (BSS) is a fundamental signal processing technology that has been intensively used in many systems including biomedical, audio and industrial applications [1]. In the context of overdetermined multiple-input multiple-output (MIMO) systems, BSS aims to find a separation matrix using the received signals and *a priori* information about the statistics or the nature of transmitted source signals. Usually, in a communication system, the modulation technique being used is known *a priori*. One can utilize such information to recover the same properties in the output signal and thus estimate the source signal blindly. Various BSS cost functions can be found in literature [1], [2] depending upon the types of source signals. For general non-Gaussian sources, efficient separation methods based on high order cumulant optimization have been introduced as

S. A. W. Shah and T. Y. Al-Naffouri are with the EE department, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia (e-mail: awaiswahab100@gmail.com). K. Abed-Meraim is with the University of Orléans, PRISME Lab. Orléans, France (e-mail: karim.abed-meraim@univ-orleans.fr). T. Y. Al-Naffouri is also associated with the EE department, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia (e-mail: tareq.alnaffouri@kaust.edu.sa).

Joint Approximate Diagonalization of Eigenmatrices (JADE) [3] and many others in [4]–[7]. However, when considering communication signals, more appropriate dedicated criteria and methods can be used instead. Among them, the constant modulus (CM) criterion for phase/frequency modulated signals such as PSK/FSK and multimodulus (MM) criterion for QAM signals have attracted great interest.

The CM criterion [8] restricts the squared modulus of the output to be a constant, but such algorithms even work for non CM signals. They lead to a number of constant modulus algorithms (CMA) used for blind equalization [9], [10], blind beamforming [11], [12] and BSS [13], [14]. On the other hand, the MM criterion [15] takes into account the knowledge of square QAM constellation. Its respective cost function deals with the real and imaginary parts of the signal separately and leads to numerous multimodulus algorithms (MMA) used for the application of blind equalization [15] and BSS [16]–[18]. MMA outperforms the CMA for the case of square QAM, which is used in many modern communication systems such as LTE [19] and WiMAX [20]. For such advanced systems requiring high data rate, high-order modulations having better spectral efficiency are used such as 64-QAM. For these high-order modulations, MMA leads to considerable amount of residual errors and does not ensure low symbol error rate (SER). Thus, in order to improve the performance of BSS algorithms for high-order QAM signals, a number of alphabet matched (AM) penalty terms [21]–[23] were suggested. All of these AM cost functions were found to have good local convergence properties and therefore require a good initialization [23]. Thus, alphabet matched algorithms (AMA) should be used along with either CMA or MMA, as both of them have good global convergence properties.

Out of numerous CMA solutions, the algebraic one named Analytical Constant Modulus Algorithm (ACMA) [13] provides an exact separation in the noise-free case. To overcome the drawback of numerical complexity of ACMA, two batch¹ BSS algorithms namely Givens CMA (G-CMA) and Hyperbolic G-CMA (HG-CMA) were presented in [14], which outperform ACMA. The adaptive versions of ACMA and G-CMA were presented in [24] and [25], respectively. Similarly, for the MM criterion, an adaptive MMA algorithm using stochastic gradient minimization technique was presented in [16], which outperforms the Multi-User Kurtosis (MUK) algorithm [26].

¹Batch refers to the algorithms which act on the collection of samples of a signal, as opposed to the adaptive algorithms which update the parameter estimate with the arrival of every sample.

Seeing the popularity of ACMA, the same analytical approach was used for MM signals and thus an Analytical Multimodulus Algorithm (AMMA) was presented in [17], which outperforms the MMA algorithm in [16]. In terms of blind equalization considering a single source, a number of AMA were presented using a combination of CM/MM and AM cost functions either in hybrid (i.e., with combined cost functions) [23] or dual mode (i.e., CM/MM for initialization followed by AM for the estimation refining) [27]. It is shown in [28] that hybrid and dual mode have nearly the same performance. An adaptive blind equalization algorithm by combining CM and AM cost functions was presented in [29] that separates all the sources using multi-stage cascaded equalizers, where the number of equalizers were equal to the number of sources.

A. Contributions

In this paper, we propose four new batch BSS algorithms utilizing the MM and AM criteria for MIMO systems. The major contribution includes the optimization of MM/AM criterion using Givens and hyperbolic rotation parameters for the case of multiple sources. Two algorithms are designed by minimization of MM and AM cost functions using Givens rotations and named as Givens MMA (G-MMA) and Givens AMA (G-AMA), respectively. The other two algorithms are designed using both Givens and hyperbolic rotations and thus named as Hyperbolic G-MMA (HG-MMA) and Hyperbolic G-AMA (HG-AMA). To overcome the difficulties encountered for the optimization of parameters of complex Givens and hyperbolic rotations (see section II-C), we have considered real representations of these complex matrices (i.e., real Givens and hyperbolic matrices) that lead to closed form formulas for the optimal rotation parameters. To the best of our knowledge, this is the first paper which presents batch BSS algorithms for MM and AM criteria.

Previously, stochastic gradient techniques were used for the minimization of MM and AM criteria [16], [21]–[23], [29], thus all of these algorithms are adaptive and slow in convergence. Therefore, we compare our algorithms with batch BSS algorithms designed for CM signals and general non-Gaussian sources such as ACMA, G-CMA, HG-CMA and JADE, in terms of signal to interference and noise ratio (SINR), SER, and convergence rate.

B. Paper Organization and Notations

This paper is organized as follows. In section II, the data model and a brief overview of BSS principle are presented. Then, the used criteria as well as Givens and hyperbolic rotations are defined. The derivation of proposed algorithms G-MMA, HG-MMA, G-AMA and HG-AMA is presented in sections III, IV, V, and VI, respectively. Section VII includes some comments to highlight the important features of the proposed algorithms. Simulation results are presented in section VIII and section IX concludes the paper.

Following are the notations used in this paper. \mathbf{x} denotes a column vector where its i th entry is denoted by x_i . The real and imaginary parts of x are denoted by x_R and x_I . The matrix and its (i, j) -th entry are denoted by \mathbf{X} and x_{ij} , respectively. If

the matrix consists of only real elements, then it is represented as $\tilde{\mathbf{X}}$. \mathbf{I} represents the identity matrix. $(\cdot)^T$ and $(\cdot)^H$ are used to represent matrix/vector transpose and complex conjugate transpose, respectively. \underline{x} denotes the pre-filtered variables. ι is used to denote $\sqrt{-1}$. $E[\cdot]$ is the mathematical expectation operator and $|\cdot|$ denotes the modulus function.

II. PROBLEM FORMULATION

Consider a MIMO system consisting of N_t sources, each having a single antenna element and a receiver equipped with an array of N_r antennas². All sources transmit their signals over the same band of frequencies. Each transmitted source signal $s(i) = s_R(i) + \iota s_I(i)$ is drawn from a known L -ary square QAM constellation where $s_R(i), s_I(i) \in \{\pm 1, \pm 3, \dots, \pm(\sqrt{L} - 1)\}$. The unknown source signal $\mathbf{s}(i) = [s_1(i) \ \dots \ s_{N_t}(i)]^T$ is passed through a flat fading channel represented by an unknown mixing matrix $\mathbf{A} \in \mathbb{C}^{N_r \times N_t}$ whose elements a_{mn} denotes the channel path between transmitter n and receiver m . The received signal with the added noise can be mathematically represented as

$$\mathbf{y}(i) = \mathbf{A}\mathbf{s}(i) + \mathbf{n}(i) \quad (1)$$

where $\mathbf{n}(i) = [n_1(i) \ \dots \ n_{N_r}(i)]^T$ is the white noise vector of covariance $\sigma_n^2 \mathbf{I}_{N_r}$. Here, we assume that the mixing matrix \mathbf{A} is of full column rank which implies that $N_r \geq N_t$.

The objective is to recover the source signals $\mathbf{s}(i)$ without prior knowledge of the channel or without the use of training sequences (pilots). This is accomplished using BSS which relies on the observation vector $\mathbf{y}(i)$ only and also uses some source's structural information. In order to recover the source signals (up to a permutation and scaling factors [14]), we apply a $(N_t \times N_r)$ separation matrix \mathbf{W} according to

$$\mathbf{z}(i) = \mathbf{W}\mathbf{y}(i) = \mathbf{W}\mathbf{A}\mathbf{s}(i) + \mathbf{W}\mathbf{n}(i) = \mathbf{G}\mathbf{s}(i) + \bar{\mathbf{n}}(i) \quad (2)$$

where $\mathbf{z}(i) = [z_1(i) \ \dots \ z_{N_t}(i)]^T$ is the estimated source signal vector, $\mathbf{G} = \mathbf{W}\mathbf{A}$ is the $(N_t \times N_t)$ global system matrix and $\bar{\mathbf{n}}(i) = \mathbf{W}\mathbf{n}(i)$ is the filtered noise vector.

In this paper, we consider batch BSS algorithms in which N_s samples of the received signal are collected and then a separation matrix is applied on the received data packet $\mathbf{Y} = [\mathbf{y}(1) \ \dots \ \mathbf{y}(N_s)]$, so that (1) and (2) can be rewritten as

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{N}, \quad \mathbf{Z} = \mathbf{W}\mathbf{Y} \quad (3)$$

where \mathbf{Z} , \mathbf{S} and \mathbf{N} are defined in a way similar to the definition of \mathbf{Y} . In what follows, we seek a separation matrix in the form $\mathbf{W} = \mathbf{V}\mathbf{B}$, where \mathbf{B} is a $(N_t \times N_r)$ pre-whitening matrix that can be computed from a covariance matrix as in [12], or simply a $(N_t \times N_r)$ projection matrix onto the signal subspace (since pre-whitening is needed only for the G-MMA/G-AMA but not for the HG-MMA/HG-AMA methods). Our main contribution lies in designing efficient methods for the computation of the matrix \mathbf{V} in order to minimize cost functions suitable for high-order QAM signals.

²In fact, N_r may represent the number of physical channels in the case N_r sensors are used at the receiver, or 'virtual' channels if fractional sampling (i.e., oversampling of the known symbol period) is used to exploit the communication signal's cyclostationarity as shown in [30].

The first step of algorithm design is the selection of a suitable cost function. Various cost functions can be found in the literature depending upon the properties/types of source signals. The latter can be classified into two categories: (i) cost functions and methods using explicitly some signal's statistics (e.g., higher order cumulants) such as JADE [3] and many others in [4]–[7], and (ii) contrast functions and methods involving non-linear criteria which implicitly rely on the signal's higher order statistics like CMA [8], MMA [15] and in [31]. In the considered case of square QAM signals, this work deals with the second class of methods by using the following cost functions.

A. Cost Functions

For low order square QAM signals, we design MMA using MM cost function, however for high-order square QAM signals, we design AMA using AM cost function as shown next.

1) *Multimodulus (MM) Cost Function*: For multimodulus signals e.g., square QAM, one proposes to estimate the matrix \mathbf{V} by minimizing the MM criterion defined in [15] as

$$\mathcal{J}_{\text{MMA}}(\mathbf{V}) = \sum_{j=1}^{N_t} \mathbb{E} \left[(z_{j,R}^2(i) - R_R)^2 + (z_{j,I}^2(i) - R_I)^2 \right] \quad (4)$$

where $R_R = R_I = \mathbb{E}[|s_R(i)|^4]/\mathbb{E}[|s_R(i)|^2]^2$ are dispersion constants of the real and the imaginary parts, respectively. This cost function was designed such that its minimization can be interpreted as fitting the signal into a square shaped signal. Thus, it contains structural information of QAM signals and also has an inherent ability to restore the phase of the signal. Moreover, the MM cost function has several advantages over the CM one [32] and leads to: i) faster convergence algorithms [33], [34], ii) carrier phase recovery [35], iii) less undesirable minima [36] and iv) ease in hardware implementation [37].

2) *Alphabet Matched (AM) Cost Function*: The AM cost function is used to enhance local convergence property. Similar to MM cost functions, it deals with the real and imaginary parts of the signal separately, where the information of the square QAM signal is embedded in the constellation matched error (CME) term. A variety of AM cost functions were available in the literature [21], [22], [38], [39]. Due to the reasons presented in section V, we have selected the one presented in [40] as

$$\mathcal{J}_{\text{AMA}}(\mathbf{V}) = \sum_{j=1}^{N_t} \mathbb{E} [g(z_{j,R}(i)) + g(z_{j,I}(i))] \quad (5)$$

where $g(x)$ is the CME term defined as

$$g(x) = 1 - \sin^{2n} \left(x \frac{\pi}{2d} \right) \quad (6)$$

where $n \in \mathbb{N}$ and $2d$ is the minimum distance between alphabet points. The choice of the CME term is not unique, and many other smooth functions with minima at the constellation points can be used. However, the CME in (6) satisfies a number of properties that shape the high-order square QAM signals including: i) it does not favor alphabet members over others, thus it has a uniform behavior, ii) it is locally symmetric

around each alphabet point, and iii) it places the highest penalty at the maximum deviation i.e., the midpoint between two alphabet points and does not place any penalty for zero errors i.e., at the alphabet points.

Remark: For simplicity, we have considered in this work the case where all sources belong to the same square QAM constellation. However, the proposed algorithms can be used in the non-square QAM case as well as in the case of mixed QAM constellations (i.e. the source signals belong to different QAM alphabets). In the latter case, we just need to use different non-linear functions (i.e. function $g(\cdot)$ in (5)) corresponding to the considered constellations.

In the sequel, instead of considering a hybrid function with MM and AM, it is proposed here to optimize the MM cost function at first to get an initial estimate of the separation matrix, then to refine this estimate by optimizing the AM criterion in (5). The next step is to devise an efficient method for the optimization of the aforementioned cost functions.

To guarantee a fast convergence with relatively easy implementation, we propose to decompose the separation matrix \mathbf{V} into a product of elementary rotations, similar to Jacobi-like algorithms [41], [42], used for matrix diagonalization. Hence, \mathbf{V} is derived using a sequence of Givens and hyperbolic rotations, whose parameters are computed by minimizing the MM/AM criteria.

B. Review of Givens and Hyperbolic (Shear) Rotations

1) *Givens Rotations*: The unitary Givens rotation $\mathcal{G}_{p,q}(\theta, \alpha)$ is an $(m \times m)$ identity matrix except for the four entries $\mathcal{G}_{pp}, \mathcal{G}_{qq}, \mathcal{G}_{pq}$ and \mathcal{G}_{qp} given by

$$\begin{bmatrix} \mathcal{G}_{pp} & \mathcal{G}_{pq} \\ \mathcal{G}_{qp} & \mathcal{G}_{qq} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & e^{i\alpha} \sin(\theta) \\ -e^{-i\alpha} \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (7)$$

where $\theta \in [-\pi/2, \pi/2]$ and $\alpha \in [-\pi/2, \pi/2]$ are angle parameters with $\alpha = 0$ for the real case.

2) *Hyperbolic Rotations*: The non-unitary Hyperbolic rotation $\mathcal{H}_{p,q}(\gamma, \beta)$ is an $(m \times m)$ identity matrix, except for the four elements $\mathcal{H}_{pp}, \mathcal{H}_{qq}, \mathcal{H}_{pq}$ and \mathcal{H}_{qp} given by

$$\begin{bmatrix} \mathcal{H}_{pp} & \mathcal{H}_{pq} \\ \mathcal{H}_{qp} & \mathcal{H}_{qq} \end{bmatrix} = \begin{bmatrix} \cosh(\gamma) & e^{i\beta} \sinh(\gamma) \\ e^{-i\beta} \sinh(\gamma) & \cosh(\gamma) \end{bmatrix} \quad (8)$$

where $\gamma \in [-\Gamma, \Gamma], \Gamma > 0$ and $\beta \in [-\pi/2, \pi/2]$. Similar to Givens rotations, in the real case, $\beta = 0$.

C. Motivation for using Real Givens and Hyperbolic Rotations

For large number of sources N_t , the difficulty to estimate \mathbf{V} increases. Thus, to simplify the estimation process, similar to Jacobi-like algorithms [41], [42], we propose to decompose \mathbf{V} into a product of $N_t(N_t - 1)$ elementary Givens rotations as

$$\mathbf{V} = \prod_{N_{\text{Sweeps}}} \prod_{1 \leq p, q \leq N_t} \mathcal{G}_{p,q}(\theta, \alpha) \quad (9)$$

where N_{Sweeps} denotes the number of iterations. Parameters θ and α are computed in order to minimize the MM criterion (4). Consider a unitary transformation $\mathbf{Z} = \mathcal{G}_{p,q} \mathbf{Y}$, which

according to (7) only changes the rows ' $j = p$ ' and ' $j = q$ ' of $\underline{\mathbf{Y}}$ so that

$$\begin{aligned} z_{ji} &= \underline{y}_{ji} \text{ for } j \neq p, q \\ z_{pi} &= \cos(\theta)\underline{y}_{pi} + e^{i\alpha} \sin(\theta)\underline{y}_{qi} \\ z_{qi} &= -e^{-i\alpha} \sin(\theta)\underline{y}_{pi} + \cos(\theta)\underline{y}_{qi} \end{aligned} \quad (10)$$

By omitting the constant terms of \mathbf{Z} independent of (θ, α) , (4) can be approximated by

$$\mathcal{J}_{\text{MMA}}(\mathcal{G}_{pq}) = \sum_{i=1}^{N_s} \left[(z_{pi,R}^2 - R_R)^2 + (z_{qi,R}^2 - R_R)^2 + (z_{pi,I}^2 - R_I)^2 + (z_{qi,I}^2 - R_I)^2 \right] \quad (11)$$

where each term $z_{pi,R}^2$, $z_{qi,R}^2$, $z_{pi,I}^2$, and $z_{qi,I}^2$ equals to $g_i^1 \cos(2\theta) + g_i^2 \cos(2\theta) \cos(2\alpha) + g_i^3 \cos(2\theta) \sin(2\alpha) + g_i^4 \sin(2\theta) \cos(\alpha) + g_i^5 \sin(2\theta) \sin(\alpha) + g_i^6 \cos(2\alpha) + g_i^7 \sin(2\alpha) + g_i^8$ and $g_i^j, j = 1, \dots, 8$ are constant terms depending upon the entries of $\underline{\mathbf{Y}}$. As we can see, further analytical simplification and thus the solution of (11) is quite complicated. Similar is the case with hyperbolic rotations. These difficulties motivated us to come up with a different solution explained below³.

III. GIVENS MMA (G-MMA)

Until now, we have been working in the complex domain and to deal with the previously mentioned challenges, we will now work in the real domain. Hence, matrix $\underline{\mathbf{Y}}$ is converted into a real matrix $\hat{\mathbf{Y}}$ containing real and imaginary parts in separate rows as defined in (12). Moreover, a special structure of matrix \mathbf{V} is introduced and maintained while applying the rotations. The transformed real received signal and output signal can now be written as $\hat{\mathbf{Y}}$ and $\hat{\mathbf{Z}} = \mathbf{V}\hat{\mathbf{Y}}$, respectively, where

$$\hat{\mathbf{Y}} = \begin{bmatrix} \underline{\mathbf{Y}}_R \\ \underline{\mathbf{Y}}_I \end{bmatrix} (2N_t \times N_s), \quad \hat{\mathbf{V}} = \begin{bmatrix} \mathbf{V}_R & -\mathbf{V}_I \\ \mathbf{V}_I & \mathbf{V}_R \end{bmatrix} (2N_t \times 2N_t) \quad (12)$$

Similarly, $\hat{\mathbf{S}}$ and $\hat{\mathbf{Z}}$ are now $(2N_t \times N_s)$ real matrices, which can be represented in a way similar to the definition of $\hat{\mathbf{Y}}$ in (12). In order to find the required matrix $\hat{\mathbf{V}}$, considering Lemma 1 of [43], the following sequence of real Givens rotations are used as a counterpart of (9)

$$\hat{\mathbf{V}} = \prod_{N_{\text{Sweeps}}} \prod_{\substack{1 \leq p, q \leq N_t \\ p \neq q}} \mathcal{G}_{p,q}(\theta) \mathcal{G}_{p+N_t, q+N_t}(\theta) \mathcal{G}_{p, q+N_t}(\dot{\theta}) \mathcal{G}_{q, p+N_t}(\dot{\theta}) \prod_{1 \leq p \leq N_t} \mathcal{G}_{p, p+N_t}(\ddot{\theta}) \quad (13)$$

The rotations $\mathcal{G}_{p,q}(\theta)$ and $\mathcal{G}_{p+N_t, q+N_t}(\theta)$ are applied successively using the same angle parameter (θ) . Similarly, the rotations $\mathcal{G}_{p, q+N_t}(\dot{\theta})$ and $\mathcal{G}_{q, p+N_t}(\dot{\theta})$ are applied with another angle parameter $(\dot{\theta})$. Note that, these rotations are paired in this way to preserve the structure of $\hat{\mathbf{V}}$ given in (12) [43]. The rotation $\mathcal{G}_{p, p+N_t}(\ddot{\theta})$ is applied to deal with the phase shift introduced by the diagonal entries of the mixing matrix

³We have presented this work partly in [18].

A. The angle parameters (θ) , $(\dot{\theta})$ and $(\ddot{\theta})$ are computed in such a way to minimize the MM criterion (11), using above explained iterative method. For that, we express the MM cost function in terms of the angle parameter (θ) . Now, consider a unitary transformation $\hat{\mathbf{Z}} = \mathcal{G}_{p,q} \hat{\mathbf{Y}}$, which according to (7) only changes the rows ' p ' and ' q ' of $\hat{\mathbf{Y}}$ so that

$$\begin{aligned} \hat{z}_{ji} &= \hat{y}_{ji} \text{ for } j \neq p, q \\ \hat{z}_{pi} &= \cos(\theta)\hat{y}_{pi} + \sin(\theta)\hat{y}_{qi} \\ \hat{z}_{qi} &= -\sin(\theta)\hat{y}_{pi} + \cos(\theta)\hat{y}_{qi} \end{aligned} \quad (14)$$

Similarly, the rotation⁴ $\mathcal{G}_{p+N_t, q+N_t}$ with the same angle parameter (θ) modifies the rows ' $p + N_t$ ' and ' $q + N_t$ ' in a similar way as shown in (14). Now, (11) can be rewritten in terms of (θ) as (omitting the terms of $\hat{\mathbf{Z}}$ that are independent of (θ) and assuming for simplicity that $R_R = R_I = R$)

$$\mathcal{J}_{\text{MMA}}(\theta) = \sum_{i=1}^{N_s} \left[(z_{pi}^2 - R)^2 + (z_{qi}^2 - R)^2 + (z_{p+N_t, i}^2 - R)^2 + (z_{q+N_t, i}^2 - R)^2 \right] \quad (15)$$

Using (14) and double angle identities we can write

$$\begin{aligned} z_{pi}^2 &= \mathbf{t}_i^T \mathbf{v} + \frac{1}{2} (\hat{y}_{pi}^2 + \hat{y}_{qi}^2) \\ z_{qi}^2 &= -\mathbf{t}_i^T \mathbf{v} + \frac{1}{2} (\hat{y}_{pi}^2 + \hat{y}_{qi}^2) \end{aligned} \quad (16)$$

where

$$\begin{aligned} \mathbf{v} &= [\cos(2\theta) \quad \sin(2\theta)]^T \\ \mathbf{t}_i &= \left[\frac{1}{2} (\hat{y}_{pi}^2 - \hat{y}_{qi}^2) \quad \hat{y}_{pi} \hat{y}_{qi} \right]^T \end{aligned} \quad (17)$$

This allows us to express the first two terms in (15) as

$$(z_{pi}^2 - R)^2 + (z_{qi}^2 - R)^2 = 2\mathbf{v}^T \mathbf{t}_i \mathbf{t}_i^T \mathbf{v} + 2 \left(\frac{\hat{y}_{pi}^2 + \hat{y}_{qi}^2}{2} - R \right)^2 \quad (18)$$

Similarly, the terms $z_{p+N_t, i}^2$ and $z_{q+N_t, i}^2$ are obtained by replacing \mathbf{t}_i with $\hat{\mathbf{t}}_i$ corresponding to indices ' $p + N_t$ ' and ' $q + N_t$ ' in (17). Disregarding the constant terms in (18), we can express $\mathcal{J}_{\text{MMA}}(\theta)$ as a quadratic form

$$\mathcal{J}_{\text{MMA}}(\theta) = \mathbf{v}^T \sum_{i=1}^{N_s} [\mathbf{t}_i \mathbf{t}_i^T + \hat{\mathbf{t}}_i \hat{\mathbf{t}}_i^T] \mathbf{v} = \mathbf{v}^T \mathbf{T} \mathbf{v} \quad (19)$$

The solution $\mathbf{v}^\circ = [v_1^\circ \quad v_2^\circ]^T$ that minimizes (19) is given by the unit norm eigenvector of \mathbf{T} corresponding to its smallest eigenvalue, so using (17), we can write

$$\cos(\theta) = \sqrt{\frac{1 + v_1^\circ}{2}} \quad \text{and} \quad \sin(\theta) = \frac{v_2^\circ}{\sqrt{2(1 + v_1^\circ)}} \quad (20)$$

Using (20), the computation of $\mathcal{G}_{p,q}$ and $\mathcal{G}_{p+N_t, q+N_t}$ follows directly from (7). Givens rotations $\mathcal{G}_{p, q+N_t}(\dot{\theta})$ and $\mathcal{G}_{q, p+N_t}(\dot{\theta})$ are found similarly and applied successively on $\hat{\mathbf{Y}}$ to compute the filtered separation matrix $\hat{\mathbf{V}}$ according to (13). The Givens rotation $\mathcal{G}_{p, p+N_t}(\ddot{\theta})$ for ' $p = q$ ' can be similarly found by

⁴For simplicity, we keep the notation $\hat{\mathbf{Y}}$ unchanged even though the matrix is modified after each rotation.

TABLE I: Givens MMA (G-MMA) Algorithm

<p>Initialization: $\hat{\mathbf{V}} = \mathbf{I}_{2N_t}$</p> <p>1. Pre-whitening: $\underline{\mathbf{Y}} = \mathbf{B}\mathbf{Y}$ $\mathcal{O}(N_s N_r)$</p> <p>2. Construct real matrix $\underline{\mathbf{Y}}$ using (12)</p> <p>3. Givens Rotations: $(20N_s N_t^2) + \mathcal{O}(N_s N_t)$/Sweep</p> <p>for $n = 1, \dots, N_{\text{Sweeps}}$ do</p> <p> for $p = 1, \dots, N_t$ do</p> <p> for $q = p, \dots, N_t$ do</p> <p> if $p = q$ then</p> <p> a) Compute $\mathcal{G}_{p,p+N_t}$ using (21), (20) and (7) for $\ddot{\theta}$ $(6N_s)$</p> <p> b) $\underline{\mathbf{Y}} := \mathcal{G}_{p,p+N_t} \underline{\mathbf{Y}}$ $(4N_s)$</p> <p> c) $\hat{\mathbf{V}} := \mathcal{G}_{p,p+N_t} \hat{\mathbf{V}}$</p> <p> else</p> <p> d) Compute $\mathcal{G}_{p,q}$ & $\mathcal{G}_{p+N_t,q+N_t}$ using (19), (20) and (7) for same (θ) $(12N_s)$</p> <p> e) $\underline{\mathbf{Y}} := \mathcal{G}_{p,q} \mathcal{G}_{p+N_t,q+N_t} \underline{\mathbf{Y}}$ $(8N_s)$</p> <p> f) $\hat{\mathbf{V}} := \mathcal{G}_{p,q} \mathcal{G}_{p+N_t,q+N_t} \hat{\mathbf{V}}$</p> <p> repeat (d to f) for $(p, q + N_t)$ & $(q, p + N_t)$ using same (θ) $(20N_s)$</p> <p> end if</p> <p> end for</p> <p> end for</p> <p>4. Estimate the complex sources from $\hat{\mathbf{Y}}$ using (3) and (12).</p>

following the above explained method. By replacing ‘ q ’ with ‘ $p + N_t$ ’ in (17) and (18), the cost function (4) (with the constant terms omitted) can be written as

$$\mathcal{J}_{\text{MMA}}(\ddot{\theta}) = \mathbf{v}^T \sum_{i=1}^{N_s} [\mathbf{t}_i \mathbf{t}_i^T] \mathbf{v} = \mathbf{v}^T \mathbf{T} \mathbf{v} \quad (21)$$

Hence, the solution \mathbf{v}° is the least unit norm eigenvector of \mathbf{T} and $\mathcal{G}_{p,p+N_t}(\ddot{\theta})$ is computed using (20) and (7). Matrix $\hat{\mathbf{V}}$ is initialized as $\hat{\mathbf{V}} = \mathbf{I}_{2N_t}$ and the overall algorithm is summarized in Table I.

IV. HYPERBOLIC G-MMA (HG-MMA)

For a small number of samples N_s , the pre-whitening operation is not effective and thus the transformed mixing matrix \mathbf{A} may be far from unitary. In this case, the performance of G-MMA deteriorates and thus the J-unitary real hyperbolic rotations are applied alternatively along with the Givens rotations to overcome this limitation. This results in an algorithm named Hyperbolic Givens MMA (HG-MMA). So, now the matrix $\hat{\mathbf{V}}$ can be decomposed into a product of elementary hyperbolic rotations $\mathcal{H}_{p,q}$, Givens rotations $\mathcal{G}_{p,q}$, and normalization transformation $\mathcal{N}_{p,q}$ as follows

$$\hat{\mathbf{V}} = \prod_{N_{\text{Sweeps}}} \prod_{\substack{1 \leq p, q \leq N_t \\ p \neq q}} \Gamma_{p,q}(\theta, \gamma) \Gamma_{p+N_t, q+N_t}(\theta, \gamma) \Gamma_{p, q+N_t}(\dot{\theta}, \dot{\gamma}) \Gamma_{q, p+N_t}(\dot{\theta}, -\dot{\gamma}) \prod_{1 \leq p \leq N_t} \mathcal{G}_{p, p+N_t}(\ddot{\theta}) \quad (22)$$

where $\Gamma_{p,q} = \mathcal{N}_{p,q} \mathcal{G}_{p,q} \mathcal{H}_{p,q}$. Similar to the Givens rotations, the hyperbolic rotations $\mathcal{H}_{p,q}$ and $\mathcal{H}_{p+N_t, q+N_t}$ are applied using the same parameter (γ) while $\mathcal{H}_{p, q+N_t}$ and $\mathcal{H}_{q, p+N_t}$ are applied using another same but opposite parameter $(\dot{\gamma})$ and $(-\dot{\gamma})$, respectively. We will consider dispersion parameters R_R and R_I be equal to 1 and use $\mathcal{N}_{p,q}$ for normalization. Below we give a brief of finding the hyperbolic and the normalization transformation parameters to minimize the MM criterion (4).

A. Computation of Hyperbolic and Givens rotations

Let us consider one hyperbolic transformation $\hat{\mathbf{Z}} = \mathcal{H}_{p,q} \hat{\mathbf{Y}}$, which modifies $\hat{\mathbf{Y}}$ according to

$$\begin{aligned} \hat{z}_{ji} &= \hat{y}_{ji} \text{ for } j \neq p, q \\ \hat{z}_{pi} &= \cosh(\gamma) \hat{y}_{pi} + \sinh(\gamma) \hat{y}_{qi} \\ \hat{z}_{qi} &= \sinh(\gamma) \hat{y}_{pi} + \cosh(\gamma) \hat{y}_{qi} \end{aligned} \quad (23)$$

Now, using hyperbolic double angle identities we obtain

$$\begin{aligned} \hat{z}_{pi}^2 &= \mathbf{r}_i^T \mathbf{u} + \frac{1}{2} (\hat{y}_{pi}^2 - \hat{y}_{qi}^2) \\ \hat{z}_{qi}^2 &= \mathbf{r}_i^T \mathbf{u} - \frac{1}{2} (\hat{y}_{pi}^2 - \hat{y}_{qi}^2) \end{aligned} \quad (24)$$

where

$$\begin{aligned} \mathbf{u} &= [\cosh(2\gamma) \quad \sinh(2\gamma)]^T \\ \mathbf{r}_i &= \left[\frac{1}{2} (\hat{y}_{pi}^2 + \hat{y}_{qi}^2) \quad \hat{y}_{pi} \hat{y}_{qi} \right]^T \end{aligned} \quad (25)$$

Similar expressions can be derived for $\hat{z}_{p+N_t, i}^2$ and $\hat{z}_{q+N_t, i}^2$. Substituting these expressions in (15) and omitting the terms that are independent of (γ) yields

$$\begin{aligned} \mathcal{J}_{\text{MMA}}(\gamma) &= \mathbf{u}^T \left[\sum_{i=1}^{N_s} \mathbf{r}_i \mathbf{r}_i^T + \dot{\mathbf{r}}_i \dot{\mathbf{r}}_i^T \right] \mathbf{u} - 2\mathbf{u}^T \left[\sum_{i=1}^{N_s} \mathbf{r}_i + \dot{\mathbf{r}}_i \right] \\ &= \mathbf{u}^T \mathbf{R} \mathbf{u} - 2\mathbf{u}^T \mathbf{r} \end{aligned} \quad (26)$$

where $\dot{\mathbf{r}}_i = \left[\frac{1}{2} (\hat{y}_{p+N_t, i}^2 + \hat{y}_{q+N_t, i}^2) \quad \hat{y}_{p+N_t, i} \hat{y}_{q+N_t, i} \right]^T$. The optimization problem in (26) can be solved using either Lagrange multiplier method (exact solution) or by taking linear approximation of hyperbolic sine and cosine around zero (approximate solution). Both methods are discussed below.

1) *Exact Solution:* We consider the constrained optimization

$$\min_{\mathbf{u}} \mathcal{F}(\mathbf{u}) = \mathbf{u}^T \mathbf{R} \mathbf{u} - 2\mathbf{r}^T \mathbf{u} \quad \text{s.t.} \quad \mathbf{u}^T \mathbf{J}_2 \mathbf{u} = 1 \quad (27)$$

where $\mathbf{J}_2 = \text{diag}([1 \quad -1])$ corresponding to $\cosh^2(2\gamma) - \sinh^2(2\gamma) = 1$. The Lagrangian of (27) can be written as

$$\mathcal{L}(\mathbf{u}, \lambda) = \mathbf{u}^T \mathbf{R} \mathbf{u} - 2\mathbf{r}^T \mathbf{u} + \lambda (\mathbf{u}^T \mathbf{J}_2 \mathbf{u} - 1) \quad (28)$$

The solution of this Lagrangian is given by

$$\mathbf{u} = (\mathbf{R} + \lambda \mathbf{J}_2)^{-1} \mathbf{r} \quad (29)$$

Using (29), the constraint equation results in a 4th order polynomial equation

$$\mathbf{r}^T (\mathbf{R} + \lambda \mathbf{J}_2)^{-1} \mathbf{J}_2 (\mathbf{R} + \lambda \mathbf{J}_2)^{-1} \mathbf{r} = 1 \quad (30)$$

Of the four roots of (30), we use the real value⁵ of λ that results in the minimum value of $\mathcal{L}(\mathbf{u}, \lambda)$ with a vector \mathbf{u} satisfying $u_1 > 0$. We then solve for $\mathbf{u}^\circ = [u_1^\circ, u_2^\circ]^T$ from (29) and solve for the hyperbolic sine and cosine of (γ) as

$$\cosh(\gamma) = \sqrt{\frac{1 + u_1^\circ}{2}} \quad \text{and} \quad \sinh(\gamma) = \frac{u_2^\circ}{\sqrt{2(1 + u_1^\circ)}} \quad (31)$$

which allows us to construct the hyperbolic rotations $\mathcal{H}_{p,q}$ and $\mathcal{H}_{p+N_t, q+N_t}$ defined in (8).

⁵In the case, when the set of solutions is empty, we set by default $\lambda = 0$.

For the remaining hyperbolic rotations $\mathcal{H}_{p,q+N_t}$ and $\mathcal{H}_{q,p+N_t}$, the optimization problem in (26) is conducted for the other hyperbolic parameter $(\dot{\gamma})$, where $\mathbf{r}_i = \left[\frac{1}{2}(\underline{y}_{pi}^2 + \underline{y}_{q+N_t,i}^2) \quad \underline{y}_{pi}\underline{y}_{q+N_t,i} \right]^T$ and $\dot{\mathbf{r}}_i = \left[\frac{1}{2}(\underline{y}_{qi}^2 + \underline{y}_{p+N_t,i}^2) \quad -\underline{y}_{qi}\underline{y}_{p+N_t,i} \right]^T$. Then, the modified optimization problem is minimized using the same method as explained above. This provides the solution $\mathbf{u}^{\circ} = [\dot{u}_1^{\circ}, \dot{u}_2^{\circ}]^T$ and the hyperbolic angles are obtained using (31) for hyperbolic parameter $(\dot{\gamma})$. The computation of the hyperbolic rotations $\mathcal{H}_{p,q+N_t}(\dot{\gamma})$ and $\mathcal{H}_{q,p+N_t}(-\dot{\gamma})$ follows directly from (31) and (8). Note that these rotations are applied using same but opposite hyperbolic angle parameter $\dot{\gamma}$.

2) *Approximate Solution*: In this approach, we will consider the linear approximation of hyperbolic sine and cosine around zero given by $\sinh(2\gamma) \approx 2\sinh(\gamma)$ and $\cosh(2\gamma) \approx \cosh(\gamma)$. This approximation is valid in the vicinity of the desired solution (i.e., optimal separation matrix) in which case, one can demonstrate that the expression of parameter γ given below is close to zero. Now, let us define the elements of symmetric matrix \mathbf{R} and vector \mathbf{r} used in (26) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (32)$$

Using (25), (32) and neglecting the terms independent of γ , the cost function (26) can be rewritten as

$$\mathcal{J}_{\text{MMA}}(\gamma) = \cosh(4\gamma) \frac{r_{11} + r_{22}}{2} + \sinh(4\gamma)r_{12} - 2 \cosh(2\gamma)r_1 - 2 \sinh(2\gamma)r_2 \quad (33)$$

Setting the derivative of (33) w.r.t γ to zero and using the previous approximation, we obtain

$$\sinh(2\gamma)(r_{11} + r_{22} - r_1) - \cosh(2\gamma)(r_2 - r_{12}) = 0 \quad (34)$$

and thus the solution γ is

$$\gamma = \frac{1}{2} \operatorname{arctanh} \left(\frac{r_2 - r_{12}}{r_{11} + r_{22} - r_1} \right) \quad (35)$$

In a similar way, the hyperbolic rotation parameter $\dot{\gamma}$ can be found by using appropriate \mathbf{R} and \mathbf{r} as explained in section IV-A1. The hyperbolic rotations are computed using (35) and (8) and applied accordingly as explained in section IV-A1. After applying the hyperbolic rotations, Givens rotations are applied in a similar way as explained in section III and then normalization rotations are applied as explained below.

B. Calculating the normalization transformations

The normalization is applied to compensate for the dispersion parameters R_R and R_I . Let us consider that we have transformed only one row ' p ' of matrix $\underline{\mathbf{Y}}$, which corresponds to the transformation of rows ' p ' and ' $p + N_t$ ' for matrix $\underline{\mathbf{Y}}$. In this case, the normalization transformation \mathcal{N}_p is an identity matrix except for the two diagonal elements $\mathcal{N}_{pp} = \mathcal{N}_{p+N_t,p+N_t} = \lambda_p$ and the approximated MM cost function (11) (with the constant terms omitted) becomes

$$\mathcal{J}_{\text{MMA}}(\lambda_p) = \sum_{i=1}^{N_s} \left[\left(\left(\lambda_p \underline{y}_{pi} \right)^2 - 1 \right)^2 + \left(\left(\lambda_p \underline{y}_{p+N_t,i} \right)^2 - 1 \right)^2 \right] \quad (36)$$

TABLE II: Hyperbolic Givens MMA (HG-MMA) Algorithm

Initialization: $\hat{\mathbf{V}} = \mathbf{I}_{2N_t}$	
Subspace projection or approximate pre-whitening if $N_r > N_t$	
$\mathcal{O}(N_s N_r^2)$	
1. Create real matrix $\underline{\mathbf{Y}}$ using (12)	
2. Hyperbolic, Givens & Normalization Rotations:	$(40N_s N_t^2) +$
$\mathcal{O}(N_s N_t)$	
for $n = 1, \dots, N_{\text{Sweeps}}$ do	
for $p = 1, \dots, N_t$ do	
for $q = p, \dots, N_t$ do	
if $p = q$ then	
a) Apply Givens rotation using (a to c) of Table I	$(10N_s)$
else	
b) Compute $\mathcal{H}_{p,q}$ & $\mathcal{H}_{p+N_t,q+N_t}$ using (31) and (8) for	$(12N_s)$
(γ)	
c) $\underline{\mathbf{Y}} := \mathcal{H}_{p,q} \mathcal{H}_{p+N_t,q+N_t} \underline{\mathbf{Y}}$	$(8N_s)$
d) $\hat{\mathbf{V}} := \mathcal{H}_{p,q} \mathcal{H}_{p+N_t,q+N_t} \hat{\mathbf{V}}$	
e) Apply Givens rotation using (d to f) of Table I	$(20N_s)$
repeat steps (b to e) for $(p, q + N_t)$ & $(q, p + N_t)$ using	
$(\theta, \dot{\gamma})$ & $(\theta, -\dot{\gamma})$, respectively	$(40N_s)$
end if	
end for	
end for	
f) Compute \mathcal{N} using (37)	$(6N_s N_t)$
g) $\underline{\mathbf{Y}} := \mathcal{N} \underline{\mathbf{Y}}$	$(2N_s N_t)$
h) $\hat{\mathbf{V}} := \mathcal{N} \hat{\mathbf{V}}$	
end for	

Taking the derivative of (36) w.r.t λ_p and setting the result to zero gives optimal normalization parameter

$$\lambda_p = \sqrt{\frac{\sum_{i=1}^{N_s} \underline{y}_{pi}^2 + \underline{y}_{p+N_t,i}^2}{\sum_{i=1}^{N_s} \underline{y}_{pi}^4 + \underline{y}_{p+N_t,i}^4}}, \quad \forall p \quad (37)$$

In our simulations, we observed that the normalization rotation is not necessary at each step and can be performed only once per sweep. In this case, the diagonal entries of matrix \mathcal{N} are $\mathcal{N}_{pp} = \mathcal{N}_{p+N_t,p+N_t} = \lambda_p$ given as in (37) where $1 \leq p \leq N_t$. HG-MMA is presented in Table II.

V. GIVENS AMA (G-AMA)

For the design of AM algorithms, the AM cost function (5) is selected because of the following reasons: i) it satisfies all three properties presented in section II-A2, which are sufficient conditions to shape the cost function for high-order square QAM signals, ii) it is the simplest among all other AM cost functions and computationally less expensive. Note that, the number of computations in this one is independent of the number of alphabet points as opposed to AM cost functions in [22], [38], iii) it deals with the real and imaginary parts of the output signal, separately. Thus, it is relatively easier to optimize using real Givens and hyperbolic rotations. In this section, G-MMA is used as an initialization followed by optimization of AM cost function (with $n = 1$ for CME term in (6)) using real Givens and hyperbolic rotations, which results in algorithms G-AMA and HG-AMA. The combination of MMA and AMA is not new and recently used by Labeled *et al.* [44] for the problem of blind equalization.

After using G-MMA for the initialization, the matrix $\hat{\mathbf{V}}$ is updated using the following Givens rotations

$$\hat{\mathbf{V}}^n = \prod_{\substack{1 \leq p, q \leq N_t \\ p \neq q}} \mathcal{G}_{p, q+N_t}(\hat{\theta}) \mathcal{G}_{q, p+N_t}(\hat{\theta}) \mathcal{G}_{p, q}(\theta) \mathcal{G}_{p+N_t, q+N_t}(\theta) \hat{\mathbf{V}}^{n-1} \quad (38)$$

where $n = n_0 + 1, \dots, N_{\text{Sweeps}}$, where N_{Sweeps} is the number of iterations of G-AMA until convergence and n_0 is the number of iterations of G-MMA for initialization⁶. Let us express the AM cost function in terms of the angle parameter θ which is computed such that $\mathcal{J}_{\text{AMA}}(\theta)$ is minimized. Using similar derivations as before, one can write

$$\mathcal{J}_{\text{AMA}} = \sum_{i=1}^{N_s} [g(\hat{z}_{pi}) + g(\hat{z}_{qi}) + g(\hat{z}_{p+N_t, i}) + g(\hat{z}_{q+N_t, i})] \quad (39)$$

where the first two terms in (39) can be defined with $n = 1$ in (6) as

$$g(\hat{z}_{pi}) = 1 - \sin^2 \left\{ \left(\cos(\theta) \hat{y}_{pi} + \sin(\theta) \hat{y}_{qi} \right) \left(\frac{\pi}{2d} \right) \right\} \\ g(\hat{z}_{qi}) = 1 - \sin^2 \left\{ \left(-\sin(\theta) \hat{y}_{pi} + \cos(\theta) \hat{y}_{qi} \right) \left(\frac{\pi}{2d} \right) \right\} \quad (40)$$

and the last two terms are obtained by replacing ‘ p ’ and ‘ q ’ with ‘ $p + N_t$ ’ and ‘ $q + N_t$ ’ in (40), respectively. The bounded non-linear optimization problem can now be stated as

$$\min_{\theta} \mathcal{J}_{\text{AMA}} \quad \text{s.t. } \theta \in [-\pi/4, \pi/4] \quad (41)$$

The optimization problem in (41) can be solved either by using Nelder–Mead simplex algorithm [45] that can be termed as ‘exact solution’ in the sense, with a line search, one can get the optimal solution up to some chosen accuracy, or by using Taylor series approximation of trigonometric functions around zero, which will be referred to as ‘approximate solution’. This approximation can be justified using Figure 1, which plots the values of AMA cost function \mathcal{J}_{AMA} in (39) vs. θ for some random received pre-whitened signal $\hat{\mathbf{Y}}$ after 5 sweeps of G-MMA with $N_t = 3, N_r = 5, N_s = 300$, SNR = 30dB and normalized 64-QAM constellation. Note that the optimum θ° is very close to zero. Thus, in the following section, we show that for a certain range of θ close to zero, the approximation fits very well with the original values of the AMA cost function.

A. Exact Solution

For the exact solution, the objective function in (39) is passed to the Nelder–Mead simplex algorithm⁷ along with $\theta_0 = 0.001$ as a starting point and bounds $\theta \in [-\pi/4, \pi/4]$, in order to find optimum θ° for the minimization of (41). Givens rotation matrices $\mathcal{G}_{p, q}(\theta^\circ)$ and $\mathcal{G}_{p+N_t, q+N_t}(\theta^\circ)$ are then applied to update $\hat{\mathbf{V}}$ according to (38). The remaining Givens rotations $\mathcal{G}_{p, q+N_t}(\hat{\theta})$ and $\mathcal{G}_{q, p+N_t}(\hat{\theta})$ can be found

⁶As per observations from the rate of convergence for G-MMA, it converges in $n_0 = 5$ for the considered cases. However, one can choose the number of sweeps as the one corresponding to an almost flat variation of \mathcal{J}_{MMA} .

⁷In this work, MATLAB optimization toolbox ‘**fminsearch**’ and ‘**fminsearchbd**’ is used which is an implementation of Nelder–Mead simplex algorithm.

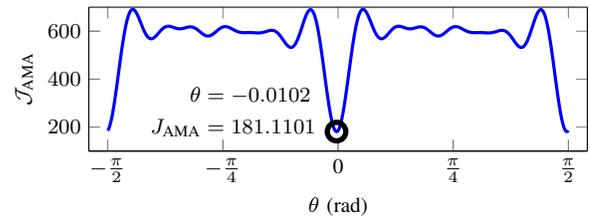


Fig. 1: \mathcal{J}_{AMA} vs. θ for random received pre-whitened signal after 5 sweeps of G-MMA with $N_t = 3, N_r = 5, N_s = 300$, SNR = 30dB and normalized 64-QAM constellation.

similarly by replacing subscripts accordingly in (39) and (40) and then computing optimum θ° . Then, the separation matrix $\hat{\mathbf{V}}$ is updated again according to (38). This process is repeated until convergence.

B. Approximate Solution

As observed from Figure 1, the optimum θ° is very close to zero, thus the Taylor series approximation around zero can be applied. Here, we will consider the approximation up to the 4th order using the following approximate identities

$$\sin(\theta) \approx \theta - \frac{\theta^3}{6}, \quad \cos(\theta) \approx 1 - \frac{\theta^2}{2} + \frac{\theta^4}{24} \quad (42)$$

Now, using the approximation in (42) to ‘ $\cos(\theta)$ ’ and ‘ $\sin(\theta)$ ’ in (40) and expanding the terms, results in

$$g(\hat{z}_{pi}) \approx \frac{1}{48d^4} c_4^{pi} \theta^4 + \frac{1}{12d^3} c_3^{pi} \theta^3 + \frac{1}{4d^2} c_2^{pi} \theta^2 \\ - \frac{1}{2d} c_1^{pi} \theta + \frac{1}{2} c_0^{pi} \quad (43)$$

where

$$c_4^{pi} = 4\pi^2 d^2 \hat{y}_{qi}^2 \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) + \pi^4 \hat{y}_{qi}^4 \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ - 3\pi^2 d^2 \hat{y}_{pi}^2 \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) - \pi d^3 \hat{y}_{pi} \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ - 6\pi^3 d \hat{y}_{pi} \hat{y}_{qi}^2 \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ c_3^{pi} = \pi d^2 \hat{y}_{qi} \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) + \pi^3 \hat{y}_{qi}^3 \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ + 3\pi^2 d \hat{y}_{pi} \hat{y}_{qi} \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ c_2^{pi} = \pi d \hat{y}_{pi} \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) - \pi^2 \hat{y}_{qi}^2 \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ c_1^{pi} = \pi \hat{y}_{qi} \sin\left(\frac{\pi \hat{y}_{pi}}{d}\right) \\ c_0^{pi} = 1 + \cos\left(\frac{\pi \hat{y}_{pi}}{d}\right) \quad (44)$$

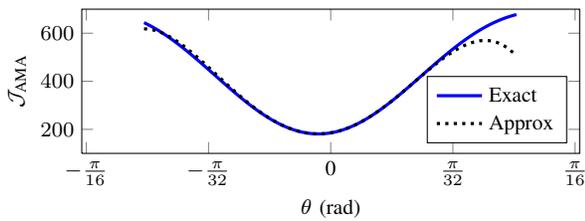


Fig. 2: Comparison of exact and approximated G-AMA cost functions after 5 sweeps of G-MMA with $N_t = 3$, $N_r = 5$, $N_s = 300$, SNR = 30dB and normalized 64-QAM constellation

Similarly $g(\hat{z}_{qi})$ can be approximated by

$$g(\hat{z}_{qi}) \approx \frac{1}{48d^4}c_4^{qi}\theta^4 - \frac{1}{12d^3}c_3^{qi}\theta^3 + \frac{1}{4d^2}c_2^{qi}\theta^2 + \frac{1}{2d}c_1^{qi}\theta + \frac{1}{2}c_0^{qi} \quad (45)$$

where all the coefficients are obtained by replacing ‘ p ’ with ‘ q ’ and ‘ q ’ with ‘ p ’ in (44). The 3rd term $g(\hat{z}_{p+N_t,i})$ of (39) has the same approximation as given in (43), where the coefficients are obtained by replacing ‘ p ’ with ‘ $p+N_t$ ’ and ‘ q ’ with ‘ $q+N_t$ ’ in (44). The last term $g(\hat{z}_{q+N_t,i})$ of (39) is approximated as (45), where the coefficients are obtained by replacing ‘ p ’ with ‘ $q+N_t$ ’ and ‘ q ’ with ‘ $p+N_t$ ’ in (44).

Now, using (43) and (45) in cost function (39) results in the 4th order polynomial equation

$$\mathcal{J}_{\text{AMA}} \approx \frac{1}{48d^4}C_4\theta^4 + \frac{1}{12d^3}C_3\theta^3 + \frac{1}{4d^2}C_2\theta^2 + \frac{1}{2d}C_1\theta + \frac{1}{2}C_0 \quad (46)$$

where the coefficients in (46) are given by

$$\begin{aligned} C_l &= \sum_{i=1}^{N_s} \left(c_l^{pi} + c_l^{qi} + c_l^{p+N_t,i} + c_l^{q+N_t,i} \right), \quad l \in \{0, 2, 4\} \\ C_3 &= \sum_{i=1}^{N_s} \left(c_3^{pi} - c_3^{qi} + c_3^{p+N_t,i} - c_3^{q+N_t,i} \right) \\ C_1 &= \sum_{i=1}^{N_s} \left(-c_1^{pi} + c_1^{qi} - c_1^{p+N_t,i} + c_1^{q+N_t,i} \right) \end{aligned} \quad (47)$$

Taking the gradient of (46) w.r.t. θ yields

$$\frac{\partial \mathcal{J}_{\text{AMA}}(\theta)}{\partial \theta} \approx \frac{1}{12d^4}C_4\theta^3 + \frac{1}{4d^3}C_3\theta^2 + \frac{1}{2d^2}C_2\theta + \frac{1}{2d}C_1 \quad (48)$$

where the coefficients are the same as defined in (47). Out of the three possible roots of (48), the optimum θ° is selected which is real-valued and results in minimum value of $\mathcal{J}_{\text{AMA}}(\theta)$ in (39).

To illustrate that the approximation in (46) is good enough, we have compared the original cost function and approximated one for a certain range of θ around zero in Figure 2.

Rotations $\mathcal{G}_{p,q+N_t}(\hat{\theta})$ and $\mathcal{G}_{q,p+N_t}(\hat{\theta})$ are similarly found by replacing subscripts accordingly and computing optimum $\hat{\theta}^\circ$. Then, the rotations are applied successively on $\underline{\mathbf{Y}}$.

In summary, matrix $\hat{\mathbf{V}}$ is initialized as identity matrix, then G-MMA is applied for $n_0 = 5$ followed by the update of matrix $\hat{\mathbf{V}}$ according to (38) by applying Givens rotations on

TABLE III: Givens AMA (G-AMA) Algorithm

<p>Initialization: $\hat{\mathbf{V}} = \mathbf{I}_{2N_t}$ 1. Pre-whitening: $\underline{\mathbf{Y}} = \mathbf{B}\mathbf{Y}$ 2. Construct real matrix $\underline{\mathbf{Y}}$ using (12) 3. Givens Rotations: for $n = 1, \dots, N_{\text{sweeps}}$ do if $n \leq n_0$ then a) Apply G-MMA as given in Table I else for $p = 1, \dots, N_t - 1$ do for $q = p + 1, \dots, N_t$ do b) Find optimum (θ°) using roots of (48) which gives minimum value of (39) c) Compute $\mathcal{G}_{p,q}$ & $\mathcal{G}_{p+N_t,q+N_t}$ using (7) for same (θ°) d) $\underline{\mathbf{Y}} := \mathcal{G}_{p,q} \mathcal{G}_{p+N_t,q+N_t} \underline{\mathbf{Y}}$ e) $\hat{\mathbf{V}} := \mathcal{G}_{p,q} \mathcal{G}_{p+N_t,q+N_t} \hat{\mathbf{V}}$ repeat (b to e) for ($p, q + N_t$) & ($q, p + N_t$) using same (θ°) end for end for end if end for</p>
--

$\underline{\mathbf{Y}}$ using the above method, until convergence. The overall algorithm is summarized in Table III.

VI. HYPERBOLIC G-AMA (HG-AMA)

As stated earlier, for a small number of samples N_s , J-unitary real hyperbolic rotations are applied alternatively along with the Givens rotations to overcome the limitation of ill-whitening. This results in an algorithm named as Hyperbolic Givens AMA (HG-AMA).

For HG-AMA, first of all G-MMA is used for initialization⁸. Then, matrix \mathbf{V} is updated iteratively until convergence using following hyperbolic $\mathcal{H}_{p,q}$ and Givens $\mathcal{G}_{p,q}$ rotations

$$\hat{\mathbf{V}}^n = \prod_{\substack{1 \leq p, q \leq N_t \\ p \neq q}} \Gamma_{p,q+N_t}(\hat{\theta}, \hat{\gamma}) \Gamma_{q,p+N_t}(\hat{\theta}, -\hat{\gamma}) \Gamma_{p,q}(\theta, \gamma) \Gamma_{p+N_t,q+N_t}(\theta, \gamma) \hat{\mathbf{V}}^{n-1} \quad (49)$$

where $\Gamma_{p,q} = \mathcal{G}_{p,q} \mathcal{H}_{p,q}$. Let us express the AM cost function in terms of parameter γ which is computed such that $\mathcal{J}_{\text{AMA}}(\gamma)$ is minimized. Now, using similar derivations as before, one can write

$$\mathcal{J}_{\text{AMA}} = \sum_{i=1}^{N_s} [g(\hat{z}_{pi}) + g(\hat{z}_{qi}) + g(\hat{z}_{p+N_t,i}) + g(\hat{z}_{q+N_t,i})] \quad (50)$$

where the first two terms in (50) can be defined as

$$\begin{aligned} g(\hat{z}_{pi}) &= 1 - \sin^2 \left\{ \left(\cosh(\gamma) \hat{y}_{pi} + \sinh(\gamma) \hat{y}_{qi} \right) \left(\frac{\pi}{2d} \right) \right\} \\ g(\hat{z}_{qi}) &= 1 - \sin^2 \left\{ \left(\sinh(\gamma) \hat{y}_{pi} + \cosh(\gamma) \hat{y}_{qi} \right) \left(\frac{\pi}{2d} \right) \right\} \end{aligned} \quad (51)$$

and the last two terms are obtained by replacing ‘ p ’ and ‘ q ’ with ‘ $p+N_t$ ’ and ‘ $q+N_t$ ’ in (51), respectively.

Figure 3 represents \mathcal{J}_{AMA} in (50) vs. γ after 5 sweeps of G-MMA with $N_t = 3$, $N_r = 5$, $N_s = 300$, SNR = 30dB and normalized 64-QAM constellation. It can be noticed that optimum γ° is very close to zero. Thus, we can apply

⁸HG-MMA can be used here as well in order to avoid the pre-whitening step.

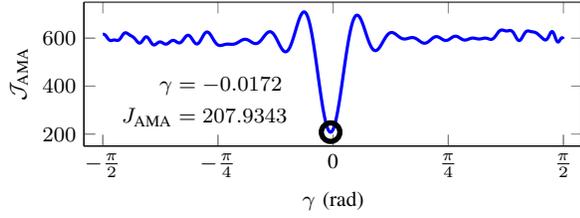


Fig. 3: \mathcal{J}_{AMA} vs. γ for random received pre-whitened signal after $N_{\text{Sweeps}} = 5$ of G-MMA with $N_t = 3, N_r = 5, N_s = 300$, SNR = 30dB and normalized 64-QAM constellation.

Taylor series approximation of hyperbolic and trigonometric functions around zero, in order to find the solution of the optimization problem in (50). Next, two possible ways are detailed to solve this optimization problem.

A. Exact Solution

For the exact solution, the objective function (50) is passed to the Nelder–Mead simplex algorithm with $\gamma_0 = 0.001$ as starting point which minimizes (50) and returns the optimum hyperbolic rotation parameter γ° . $\mathcal{H}_{p,q}(\gamma^\circ)$ and $\mathcal{H}_{p+N_t,q+N_t}(\gamma^\circ)$ are then computed and applied to update \mathbf{V} according to (49).

For rotations $\mathcal{H}_{p,q+N_t}(\dot{\gamma})$ and $\mathcal{H}_{q,p+N_t}(-\dot{\gamma})$, 1st and 4th terms of the objective function in (50) are defined as

$$g(\dot{z}_{pi}) = \cos^2 \left\{ \left(\cosh(\dot{\gamma}) \dot{y}_{pi} + \sinh(\dot{\gamma}) \dot{y}_{q+N_t,i} \right) \left(\frac{\pi}{2d} \right) \right\} \quad (52)$$

$$g(\dot{z}_{q+N_t,i}) = \cos^2 \left\{ \left(\sinh(\dot{\gamma}) \dot{y}_{pi} + \cosh(\dot{\gamma}) \dot{y}_{q+N_t,i} \right) \left(\frac{\pi}{2d} \right) \right\}$$

and the 2nd and 3rd terms of (50) are obtained by replacing $(\dot{\gamma})$ with $(-\dot{\gamma})$ and indices ‘ p ’ and ‘ $q+N_t$ ’ with ‘ q ’ and ‘ $p+N_t$ ’ in (52), respectively. Now, the modified objective function is used to find the optimum $\dot{\gamma}^\circ$. Matrices $\mathcal{H}_{p,q+N_t}$ and $\mathcal{H}_{q,p+N_t}$ are then computed using the above explained method and applied successively on \mathbf{Y} . The process is repeated until convergence.

B. Approximate Solution

Again, we will use here the Taylor series approximation of trigonometric angles given in (42) and hyperbolic angles around zero up to 4th order, which can be written as

$$\sinh(\gamma) \approx \gamma + \frac{\gamma^3}{6}, \quad \cosh(\gamma) \approx 1 + \frac{\gamma^2}{2} + \frac{\gamma^4}{24} \quad (53)$$

Let us consider the first term of (50) given in (51) as

$$g(\dot{z}_{pi}) = 1 - \sin^2 \left\{ \left(\cosh(\dot{\gamma}) \dot{y}_{pi} + \sinh(\dot{\gamma}) \dot{y}_{qi} \right) \left(\frac{\pi}{2d} \right) \right\} \quad (54)$$

Now, applying the hyperbolic angle approximation given in (53) to ‘ $\cosh(\dot{\gamma})$ ’ and ‘ $\sinh(\dot{\gamma})$ ’ in the argument of sine in (54) and expanding the terms, we get

$$g(\dot{z}_{pi}) \approx 1 - \sin^2 \left\{ \left(24\dot{y}_{pi} + 24\dot{y}_{qi}\dot{\gamma} + 12\dot{y}_{pi}\dot{\gamma}^2 + 4\dot{y}_{qi}\dot{\gamma}^3 + \dot{y}_{pi}\dot{\gamma}^4 \right) \left(\frac{\pi}{12d} \right) \right\} \quad (55)$$

Finally, the approximation in (42) is used leading to

$$g(\dot{z}_{pi}) \approx \frac{1}{48d^4} c_4^{pi} \dot{\gamma}^4 + \frac{1}{12d^3} c_3^{pi} \dot{\gamma}^3 - \frac{1}{4d^2} c_2^{pi} \dot{\gamma}^2 - \frac{1}{2d} c_1^{pi} \dot{\gamma} + \frac{1}{2} c_0^{pi} \quad (56)$$

where

$$c_4^{pi} = \pi^4 \dot{y}_{qi}^4 \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right) + 6\pi^3 d \dot{y}_{pi} \dot{y}_{qi}^2 \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right) - 4\pi^2 d^2 \dot{y}_{qi}^2 \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right) - 3\pi^2 d^2 \dot{y}_{pi}^2 \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right) - \pi d^3 \dot{y}_{pi} \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right)$$

$$c_3^{pi} = \pi^3 \dot{y}_{qi}^3 \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right) - \pi d^2 \dot{y}_{qi} \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right) - 3\pi^2 d \dot{y}_{pi} \dot{y}_{qi} \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right)$$

$$c_2^{pi} = \pi^2 \dot{y}_{qi}^2 \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right) + \pi d \dot{y}_{pi} \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right)$$

$$c_1^{pi} = \pi \dot{y}_{qi} \sin \left(\frac{\pi \dot{y}_{pi}}{d} \right)$$

$$c_0^{pi} = 1 + \cos \left(\frac{\pi \dot{y}_{pi}}{d} \right) \quad (57)$$

Similarly, the other terms $g(\dot{z}_{qi})$, $g(\dot{z}_{p+N_t,i})$ and $g(\dot{z}_{q+N_t,i})$ of (50) can be approximated as (56), where the coefficients are obtained by replacing indices accordingly in (57).

Now, using (56) in (50) leads to

$$\mathcal{J}_{\text{AMA}} \approx \frac{1}{48d^4} C_4 \dot{\gamma}^4 + \frac{1}{12d^3} C_3 \dot{\gamma}^3 - \frac{1}{4d^2} C_2 \dot{\gamma}^2 - \frac{1}{2d} C_1 \dot{\gamma} + \frac{1}{2} C_0 \quad (58)$$

where the coefficients in (58) are given by

$$C_l = \sum_{i=1}^{N_s} \left(c_l^{pi} + c_l^{qi} + c_l^{p+N_t,i} + c_l^{q+N_t,i} \right) \quad (59)$$

where $l \in \{0, \dots, 4\}$. Taking the gradient with respect to γ of AMA cost function in (58), we get

$$\frac{\partial \mathcal{J}_{\text{AMA}}(\gamma)}{\partial \gamma} \approx \frac{1}{12d^4} C_4 \dot{\gamma}^3 + \frac{1}{4d^3} C_3 \dot{\gamma}^2 - \frac{1}{2d^2} C_2 \dot{\gamma} - \frac{1}{2d} C_1 \quad (60)$$

Out of the three possible real roots of (60), the optimum ($\dot{\gamma}^\circ$) is selected such that $\mathcal{J}_{\text{AMA}}(\gamma)$ is minimum.

To illustrate that the approximation in (58) is good enough, we have compared the original cost function and approximated one for a certain range of γ around zero in Figure 4.

For the remaining matrices $\mathcal{H}_{p,q+N_t}(\dot{\gamma})$ and $\mathcal{H}_{q,p+N_t}(-\dot{\gamma})$, the AM cost function can be written as (50) and all the terms have the same approximation as given in (56) with the replacement of γ with $\dot{\gamma}$ and indices accordingly. Also, for the 2nd and 3rd terms, the sign of coefficients c_1 and c_3 are opposite to the one shown in (56). Now, using (56) the optimization problem in (50) can be written as

$$\mathcal{J}_{\text{AMA}} \approx \frac{1}{48d^4} C_4 \dot{\gamma}^4 + \frac{1}{12d^3} C_3 \dot{\gamma}^3 - \frac{1}{4d^2} C_2 \dot{\gamma}^2 + \frac{1}{2d} C_1 \dot{\gamma} + \frac{1}{2} C_0 \quad (61)$$

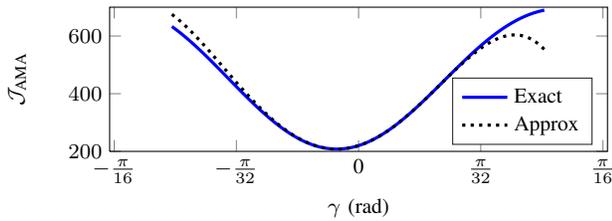


Fig. 4: Comparison of exact and approximated HG-AMA cost functions after $n_0 = 5$ of G-MMA with $N_t = 3$, $N_r = 5$, $N_s = 300$, SNR= 30dB and normalized 64-QAM constellation.

with

$$\begin{aligned}
 C_l &= \sum_{i=1}^{N_s} \left(c_l^{pi} + c_l^{q+N_t,i} + c_l^{qi} + c_l^{p+N_t,i} \right) \quad l \in \{0, 2, 4\} \\
 C_3 &= \sum_{i=1}^{N_s} \left(c_3^{pi} + c_3^{q+N_t,i} - c_3^{qi} - c_3^{p+N_t,i} \right) \\
 C_1 &= \sum_{i=1}^{N_s} \left(-c_1^{pi} - c_1^{q+N_t,i} + c_1^{qi} + c_1^{p+N_t,i} \right)
 \end{aligned} \tag{62}$$

The final solution is obtained by zeroing the gradient of (61). Once we obtain the solution $\hat{\gamma}^\circ$, matrices $\mathcal{H}_{p,q+N_t}(\hat{\gamma}^\circ)$ and $\mathcal{H}_{q,p+N_t}(-\hat{\gamma}^\circ)$ are computed using (8). The separation matrix $\hat{\mathbf{V}}$ is then updated according to (49).

Note that the normalization transformation used to compensate for the dispersion parameters in HG-MMA, is not considered here since the dispersion parameters are not involved in the AM cost function.

In summary, matrix $\hat{\mathbf{V}}$ is initialized as identity matrix then after applying 5 sweeps of G-MMA, matrix $\hat{\mathbf{V}}$ is updated according to (49) by applying Givens and hyperbolic rotations successively on $\hat{\mathbf{Y}}$ using the above explained method, until convergence. The overall algorithm is summarized in Table IV.

VII. PRACTICAL CONSIDERATIONS

We provide here some insight into the proposed algorithms.

A. Numerical Cost

Taking into account the structure of the rotation matrices, the numerical cost of the proposed algorithms are compared with other CMA-like BSS algorithms in terms of the number of flops per sweep in Table V. As can be seen from Table V, the proposed algorithms are much cheaper than ACMA and of the same cost order as G-CMA and HG-CMA. Moreover, the proposed algorithms have very fast convergence (typically less than 10 sweeps) as shown next in the simulation experiments. Also, HG-AMA is more expensive but has better performance than all the other algorithms as can be observed from the simulation results.

All considered BSS algorithms except G-CMA and G-MMA requires a pre-whitening operation which costs

TABLE IV: Hyperbolic Givens AMA (HG-AMA) Algorithm

```

Initialization:  $\hat{\mathbf{V}} = \mathbf{I}_{2N_t}$ 
Subspace projection or approximate pre-whitening if  $N_r > N_t$ 
1. Construct real matrix  $\hat{\mathbf{Y}}$  using (12)
2. Hyperbolic & Givens Rotations:
for  $n = 1, \dots, N_{\text{sweeps}}$  do
  if  $n \leq n_0$  then
    a) Apply G-MMA as given in Table I
  else
    for  $p = 1, \dots, N_t - 1$  do
      for  $q = p + 1, \dots, N_t$  do
        b) Find optimum ( $\hat{\gamma}^\circ$ ) using roots of (60) which gives minimum value of (50)
        c) Compute  $\mathcal{H}_{p,q}$  &  $\mathcal{H}_{p+N_t,q+N_t}$  using (8) for same ( $\hat{\gamma}^\circ$ )
        d)  $\hat{\mathbf{Y}} := \mathcal{H}_{p,q} \mathcal{H}_{p+N_t,q+N_t} \hat{\mathbf{Y}}$ 
        e)  $\hat{\mathbf{V}} := \mathcal{H}_{p,q} \mathcal{H}_{p+N_t,q+N_t} \hat{\mathbf{V}}$ 
        f) Apply Givens rotations using (b to e) of Table III
        repeat steps (b to f) for ( $p, q + N_t$ ) & ( $q, p + N_t$ ) using ( $\hat{\theta}^\circ, \hat{\gamma}^\circ$ ) & ( $\hat{\theta}^\circ, -\hat{\gamma}^\circ$ ), respectively
      end for
    end for
  end if
end for

```

TABLE V: Numerical complexity of different BSS algorithms

BSS Algorithm	Complexity Order
HG-AMA	$140N_s N_t^2 + \mathcal{O}(N_s N_t)$
G-AMA	$70N_s N_t^2 + \mathcal{O}(N_s N_t)$
HG-MMA	$40N_s N_t^2 + \mathcal{O}(N_s N_t)$
G-MMA	$20N_s N_t^2 + \mathcal{O}(N_s N_t)$
HG-CMA	$30N_s N_t^2 + \mathcal{O}(N_s N_t)$
G-CMA	$15N_s N_t^2 + \mathcal{O}(N_s N_t)$
ACMA	$\mathcal{O}(N_s N_t^4)$

$\mathcal{O}(N_s N_r^2)$ flops. Moreover, the numerical cost of every algorithm except ACMA in Table V has to be multiplied by the number of sweeps to obtain the overall cost.

B. Adaptive implementation

The numerical cost of the designed batch algorithms increases linearly with the sample size N_s . Furthermore, in real life environments, systems are time varying and hence the separation matrix \mathbf{W} has to be re-estimated or updated along the time axis. Thus, for slowly time varying systems, this update can be obtained by using adaptive estimation methods. Utilizing a sliding window technique as in [14], one can achieve such source separation in an adaptive manner with a numerical cost proportional to $\mathcal{O}(\hat{N}_s N_t^2)$ where \hat{N}_s is the window size (instead of total sample size N_s).

C. Complex implementation

As shown in section II-C, the real matrix representation has been introduced to overcome the difficulties encountered for the optimization of parameters of complex Givens and hyperbolic rotations. However, we can observe that the obtained results can be cast into complex matrix forms using the

following straightforward relations:

$$\begin{aligned}
 \mathcal{G}_{p,q}(\theta) \mathcal{G}_{p+N_t,q+N_t}(\theta) \dot{\mathbf{Y}} &\iff \mathcal{G}_{p,q}(\theta, 0) \mathbf{Y} \\
 \mathcal{H}_{p,q}(\gamma) \mathcal{H}_{p+N_t,q+N_t}(\gamma) \dot{\mathbf{Y}} &\iff \mathcal{H}_{p,q}(\gamma, 0) \mathbf{Y} \\
 \mathcal{G}_{p,q+N_t}(\hat{\theta}) \mathcal{G}_{q,p+N_t}(\hat{\theta}) \dot{\mathbf{Y}} &\iff \mathcal{G}_{p,q}(\theta, -\frac{\pi}{2}) \mathbf{Y} \\
 \mathcal{H}_{p,q+N_t}(\hat{\gamma}) \mathcal{H}_{q,p+N_t}(\hat{\gamma}) \dot{\mathbf{Y}} &\iff \mathcal{H}_{p,q}(\gamma, -\frac{\pi}{2}) \mathbf{Y}
 \end{aligned} \quad (63)$$

where all matrices on the left side of (63) are real and the right ones are complex. Somehow, we have replaced the two degrees of freedom of $\mathcal{G}_{p,q}(\theta, \alpha)$ (resp. $\mathcal{H}_{p,q}(\gamma, \beta)$) by the two free parameters θ and $\hat{\theta}$ (resp. γ and $\hat{\gamma}$). This way we have avoided the non-linear optimization discussed in section II-C.

D. Performance

The main advantage of the proposed algorithms resides in their fast convergence in terms of the number of sweeps (typically less than 10 sweeps are needed for convergence) and also in terms of sample size (typically $N_s = \mathcal{O}(10N_t)$ is sufficient for the algorithm's convergence). Comparatively, the ACMA method requires $N_s = \mathcal{O}(10N_t^2)$ samples for its convergence and standard CMA-like methods need even more samples to converge to their steady state.

Another important issue is the robustness of the proposed algorithms to the presence of an interference source which would not be a QAM source. Our conjecture is that, for the MMA part, the separation occurs (because the sources are sub-Gaussian [46]) but the quality would be degraded. For the AMA part, the degradation might be more significant depending on the interference signal nature. The thorough investigation of the robustness will be considered in future works.

VIII. SIMULATION RESULTS

In order to evaluate the performance of the proposed algorithms, simulation results are presented in this section. Due to the lack of any batch BSS algorithm dealing with the MM criterion, we do the comparison with batch BSS algorithms dealing with the CM criterion and general non-Gaussian signals such as ACMA, G-CMA, HG-CMA and JADE w.r.t. convergence rate, SER and SINR defined by

$$\text{SINR} = \frac{1}{N_t} \sum_{j=1}^{N_t} \text{SINR}_j \quad (64)$$

with

$$\text{SINR}_j = \frac{|g_{jj}\mathbf{s}_j|^2/N_s}{\sum_{l,l \neq j} |g_{jl}\mathbf{s}_l|^2/N_s + \mathbf{w}_j \mathbf{R}_n \mathbf{w}_j^H} \quad (65)$$

where SINR_j is the signal to interference and noise ratio at the j th output, defined similarly in [47], with $g_{ij} = \mathbf{w}_i \mathbf{a}_j$, where \mathbf{w}_i and \mathbf{a}_j are the i th row vector and j th column vector of separation matrix \mathbf{W} and mixing matrix \mathbf{A} , respectively. \mathbf{R}_n is the noise covariance matrix and \mathbf{s}_j is the $(1 \times N_s)$ source vector at j th input.

We consider a MIMO system consisting of 5 transmitters and 7 receivers ($N_t = 5, N_r = 7$) with the data model

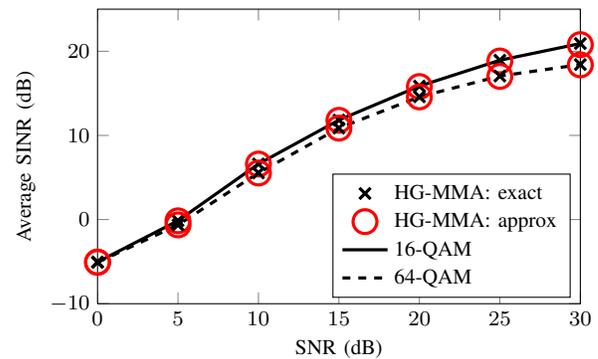


Fig. 5: Average SINR of exact and approximate solution of HG-MMA vs. SNR for $N_t = 5, N_r = 7, N_s = 100$ and $N_{\text{Sweeps}} = 10$ considering both 16-QAM and 64-QAM.

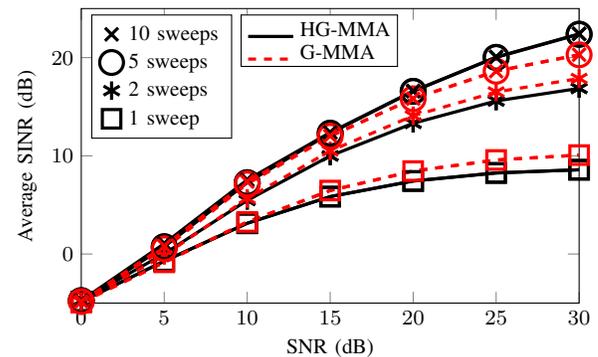


Fig. 6: Average SINR of HG-MMA and G-MMA vs. SNR for different N_{Sweeps} considering $N_t = 5, N_r = 7, N_s = 150$ and 16-QAM constellation.

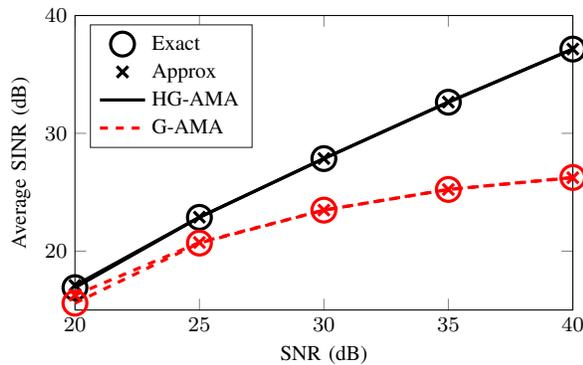
given in section II. Every uncoded data symbol transmitted by each source is drawn from 16-QAM, 64-QAM and 256-QAM constellations. The resulting signals are then passed through matrix \mathbf{A} , generated randomly at each Monte Carlo run with controlled conditioning and with i.i.d complex Gaussian variable entries of zero mean and unity variance. The noise variance is adjusted according to specified signal to noise ratio (SNR). The results are averaged over 1000 Monte Carlo runs.

A. Experiment 1: Exact vs. Approximate Solution of HG-MMA

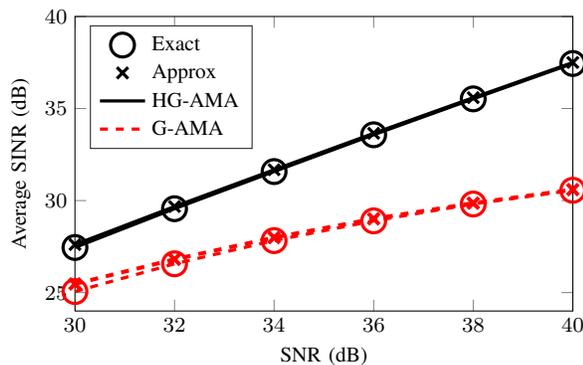
In Figure 5, we compare the exact and approximate solution of HG-MMA in terms of SINR vs. SNR for 16-QAM and 64-QAM constellations. The number of sweeps N_{Sweeps} and samples N_s are set equal to 10 and 100, respectively (in a more general context, one can use instead a stopping criterion to finish the iterative process). We notice that both the exact and approximate solutions have the same performance for the considered constellations. Therefore, in the following simulations for the HG-MMA, we will use the approximate solution, as it is cheaper and easier to implement.

B. Experiment 2: Finding Optimum Number of Sweeps for G-MMA and HG-MMA

In Figure 6, we examine the effect of the number of sweeps on the SINR of the G-MMA and HG-MMA for the case of



(a) 64-QAM, $N_s = 200$



(b) 256-QAM, $N_s = 500$

Fig. 7: Average SINR of exact and approximate solution of HG-AMA and G-AMA vs. SNR for $N_t = 5$, $N_r = 7$, $N_{\text{Sweeps}} = 10$.

$N_s = 150$ and 16-QAM. We notice that the performance of proposed algorithms increases with the number of sweeps and remains almost unchanged after 5 sweeps. So, in the following simulations we will fix the number of sweeps to 5.

C. Experiment 3: Exact vs. Approximate Solution of G-AMA and HG-AMA

Now, we compare the performance of exact and approximate solutions presented for G-AMA and HG-AMA in terms of SINR vs. SNR. Figure 7a and 7b shows the plots for $N_s = 200$, 64-QAM and $N_s = 500$, 256-QAM constellations, respectively. The number of sweeps N_{Sweeps} is fixed at 10, where we used 5 sweeps of G-MMA followed by 5 sweeps of AMAs. From Figure 7, we notice that both the exact and approximate solutions have the same performance. Therefore, in the following simulations for the G-AMA and HG-AMA, we will use the approximate solution, as it is cheaper and easier to implement.

D. Experiment 4: Finding Optimum Number of Sweeps for G-AMA and HG-AMA

In Figure 8, we examine the effect of the number of sweeps N_{Sweeps} on the SINR of the G-AMA and HG-AMA for the case of $N_s = 200$ and 64-QAM. We notice that the performance of proposed algorithms increases with the number of sweeps

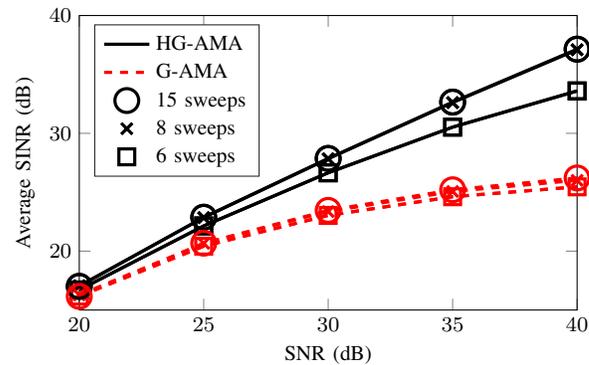


Fig. 8: Average SINR of HG-AMA and G-AMA vs. SNR for different N_{Sweeps} considering $N_t = 5$, $N_r = 7$, $N_s = 200$ and 64-QAM constellation.

and remains almost unchanged after 8 sweeps (5 G-MMA + 3 AMA sweeps). So, in the following simulations we will fix the number of sweeps to 8.

E. Experiment 5: Comparison of Rate of Convergence

In Figure 9, we have compared the convergence rate of the proposed and benchmarked algorithms. The SNR is fixed at 30 dB and N_s is selected as 200 and 500 for 64-QAM and 256-QAM, respectively. It can be noticed that G-AMA and HG-AMA converge in 8 sweeps, while all other algorithms converge in 5 sweeps. Even though the proposed algorithms G-AMA and HG-AMA require 3 extra sweeps, the performance is much better than all the other algorithms. Moreover, the performance of HG-MMA and G-MMA is better than the HG-CMA and G-CMA.

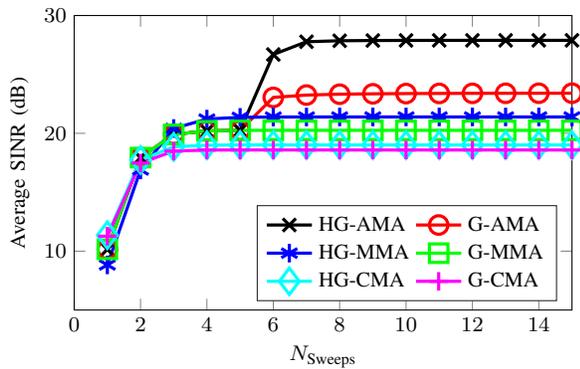
By combining the numerical complexity of the batch BSS algorithms presented in Table V and the SINR comparison of the algorithms in Figure 9, one can clearly measure the separation quality as a function of the number of flops.

F. Experiment 6: Effect of the Number of Samples

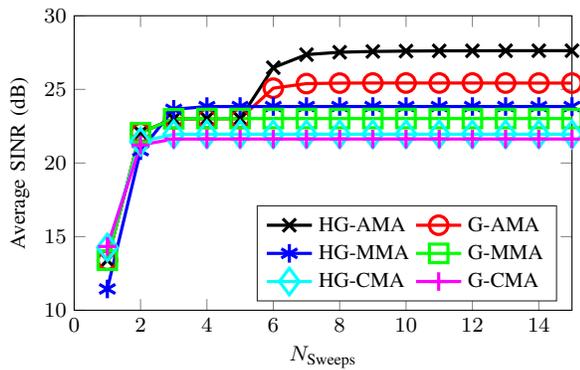
Figure 10a and 10b, show the SINR of our proposed algorithms vs. the number of samples N_s for 64-QAM and 256-QAM constellations, respectively. The SNR and the total number of sweeps N_{Sweeps} are fixed at 30 dB, and 8, respectively. We notice that as expected, the larger the number of samples the better the performance is. However, we observe a threshold point after which the gain is not significant as the SINR will be essentially limited by the SNR value. It can be seen that the performance of designed AM and MM algorithms are better than CM and JADE algorithms. Also, HG-AMA takes the lead among all other algorithms.

G. Experiment 7: Comparison based on SER

Figure 11a and 11b depict the SER of AM, MM, CM and JADE algorithms vs. SNR for the case of 64-QAM and 256-QAM constellations, respectively. The number of samples $N_s = 300$ and $N_s = 900$ are considered for the case of 64-QAM and 256-QAM, respectively. As noticed previously,



(a) 64-QAM, $N_s = 200$



(b) 256-QAM, $N_s = 500$

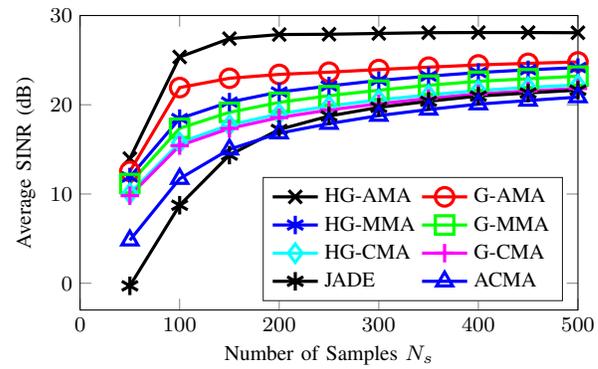
Fig. 9: Average SINR of iterative batch BSS algorithms vs. N_{Sweeps} for $N_t = 5$, $N_r = 7$ and $\text{SNR} = 30\text{dB}$.

the performance of the HG-AMA is significantly better than all the other algorithms. Further, JADE algorithm which is applicable for general non-Gaussian sources works better than ACMA and close to G-CMA for QAM signals. Similar to other figures, same pattern of performance is observed i.e., the HG-AMA takes the lead followed by the HG-MMA, G-AMA, G-MMA and then by the HG-CMA, G-CMA, JADE and ACMA. By observing these figures, we can say that HG-AMA is the only algorithm which works well for high-order QAM constellations.

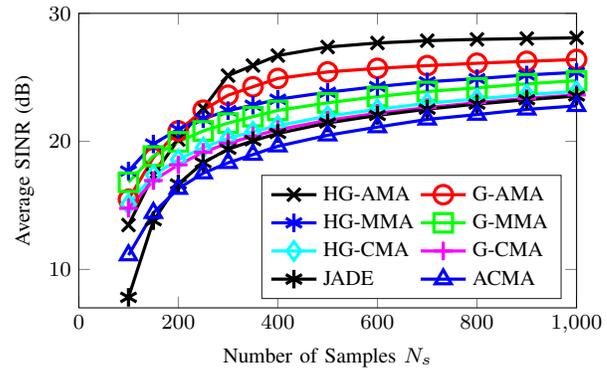
Note also, that the flat behaviour observed at high SNR is related to the ‘occasional ill’ convergence of the algorithms due probably to bad initializations in some of the simulation runs (we recall that we used only few iterations of MMA for the initialization). Fortunately, this convergence problems are ‘rare’ and in most cases the algorithms converge properly as illustrated by their good averaged performance results.

IX. CONCLUSION

In this paper, fundamental problems with the physical layer for MIMO systems are addressed. The targeted problems include channel estimation and blind demixing. Mainly, the problem focussed here is to design algorithms for high-order QAM signals without using pilot symbols. Four new iterative batch BSS algorithms are presented; two of them dealing with the MM criterion namely G-MMA and HG-MMA and the



(a) 64-QAM



(b) 256-QAM

Fig. 10: Average SINR of batch BSS algorithms vs. N_s for $N_t = 5$, $N_r = 7$, $\text{SNR} = 30\text{dB}$ and $N_{\text{Sweeps}} = 8$.

other two dealing with the AM criterion namely G-AMA and HG-AMA. The proposed algorithms are designed using a pre-whitening operation to reduce the complexity of optimization problem, followed by a recursive separation method of unitary Givens and J-unitary hyperbolic rotations for the minimization of MM/AM criteria. Instead of using complex matrices, a real transformation is considered where a special structure of the separation matrix in the whitened domain is suggested and maintained throughout all transformations.

The proposed algorithms are mainly designed for the blind demixing of MIMO systems involving high-order QAM signals. Simulation results demonstrate their favorable performance as compared to the state of the art algorithms dealing with the CM criterion such as G-CMA, HG-CMA and ACMA. It is noticed that the G-MMA and G-AMA are cheaper and more suitable for large number of samples but in the case of small number of samples the HG-MMA and HG-AMA should be used. Moreover, out of all the currently available batch BSS algorithms and the presented ones, the alphabet matched algorithm designed by combining Givens and hyperbolic rotations (HG-AMA) is the most efficient one for high-order QAM signals such as 64-QAM and 256-QAM.

REFERENCES

- [1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Elsevier Sci., 2010.

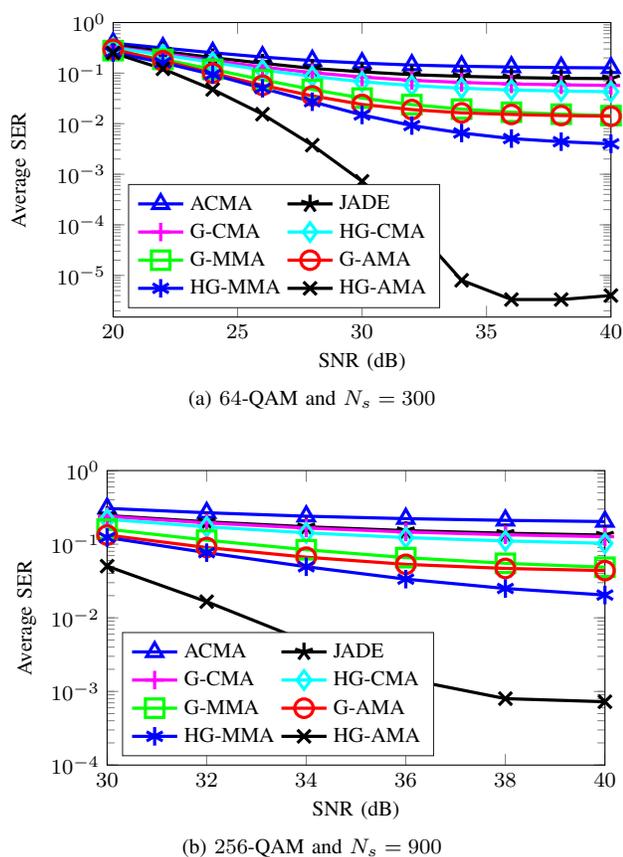


Fig. 11: Average SER of batch BSS algorithms vs. SNR for $N_t = 5$, $N_r = 7$ and $N_{\text{Sweeps}} = 8$.

[2] S. Haykin, *Unsupervised Adaptive Filtering: Blind source separation*, ser. Wiley-Interscience publication. Wiley, 2000.

[3] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," in *IEE Proc. F*, vol. 140, pp. 362–370, no. 6, 1993.

[4] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.

[5] C. Simon, P. Loubaton, and C. Jutten, "Separation of a class of convolutive mixtures: a contrast function approach," *Signal Process.*, vol. 81, no. 4, pp. 883–887, Apr. 2001.

[6] M. Castella, S. Rhioui, E. Moreau, and J.-C. Pesquet, "Quadratic higher order criteria for iterative blind separation of a MIMO convolutive mixture of sources," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 218–232, Jan. 2007.

[7] P. Chevalier, "Optimal separation of independent narrow-band sources: Concept and performance," *Signal Process.*, vol. 73, no. 1, pp. 27–47, Jan. 1999.

[8] J. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. on Acoust., Speech and Signal Process.*, vol. 31, no. 2, pp. 459–472, Apr. 1983.

[9] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. of the IEEE*, vol. 86, no. 10, pp. 1927–1950, 1998.

[10] S. Abrar and A. K. Nandi, "An adaptive constant modulus blind equalization algorithm and its stochastic stability analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 55–58, 2010.

[11] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. IEEE ICASSP*, vol. 11, pp. 2523–2526, 1986.

[12] A. J. van der Veen and A. Leshem, "Constant modulus beamforming," in *Robust Adaptive Beamforming*, (J. Li and P. Stoica, eds.). John Wiley & Sons, Inc., ch. 6, pp. 299–351, 2005.

[13] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. on Signal Process.*, vol. 44, pp. 1136–1155, May. 1996.

[14] A. Ikhlef, R. Iferroujene, A. Boudjellal, K. Abed-Meraim, and A. Belouchrani, "Constant modulus algorithms using hyperbolic Givens rotations," *Signal Process.*, vol. 104, pp. 412 – 423, 2014.

[15] K. Oh and Y. Chin, "Modified constant modulus algorithm: blind equalization and carrier phase recovery algorithm," in *Proc. IEEE ICC*, vol. 1, pp. 498–502, Jun. 1995.

[16] P. Sansrimahachai, D. B. Ward, and A. G. Constantinides, "Blind source separation for BLAST," in *Proc. IEEE DSP*, vol. 1, pp. 139–142, 2002.

[17] S. Daumont and D. Le Guennec, "An analytical multimodulus algorithm for blind demodulation in a time-varying MIMO channel context," *Int. J. of digital multimedia broadcasting*, 2010.

[18] 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 *Proc. IEEE ICASSP*, pp. 2155–2159, Apr. 2015.

[19] H. Lüders, A. Minwegen, B. Eschbach, and P. Vary, "An improvement of UMTS LTE 64QAM performance," *European Trans. on Telecom.*, vol. 21, no. 5, pp. 443–448, 2010.

[20] M. M. Khairy, "A novel frequency offset estimation technique for mobile WiMAX," *European Trans. on Telecom.*, vol. 22, no. 1, pp. 45–50, 2011.

[21] T.-H. Li and K. Mbarek, "A blind equalizer for nonstationary discrete-valued signals," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 247–254, Jan. 1997.

[22] S. Barbarossa and A. Scaglione, "Blind equalization using cost function matched to the signal constellation," in *Proc. Asilomar Conf. Signals Syst. Comput.*, vol. 1, pp. 550–554 vol.1, Nov. 1997.

[23] L. He, M. Amin, C. Reed Jr., and R. Malkemes, "A hybrid adaptive blind equalization algorithm for QAM signals in wireless communications," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 2058–2069, Jul. 2004.

[24] A. J. van der Veen, "An adaptive version of the algebraic constant modulus algorithm [blind source separation applications]," in *Proc. IEEE ICASSP*, vol. 4, pp. iv/873–iv/876, Mar. 2005.

[25] A. Boudjellal, K. Abed-Meraim, A. Belouchrani, and P. Ravier, "Adaptive constant modulus algorithm based on complex Givens rotations," in *IEEE Workshop on Stat. Signal Process.*, pp. 165–168, Jun. 2014.

[26] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3508–3519, Dec. 2000.

[27] L. He and M. Amin, "A dual-mode technique for improved blind equalization for QAM signals," *IEEE Signal Process. Lett.*, vol. 10, no. 2, pp. 29–31, Feb. 2003.

[28] A. Beasley and A. Cole-Rhodes, "Blind adaptive equalization for QAM signals using an alphabet-matched algorithm," in *Proc. IEEE GLOBECOM*, pp. 1–5, Nov. 2006.

[29] F. Moazzami and A. Cole-Rhodes, "An adaptive blind equalizer with signal separation for a mimo system transmitting qam signals," in *IEEE Military Commun. Conf., 2008. MILCOM 2008.*, pp. 1–5, Nov. 2008.

[30] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.

[31] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proc. of the IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct. 1998.

[32] S. Abrar and A. K. Nandi, "Blind equalization of square-QAM signals: a multimodulus approach," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1674–1685, Jun. 2010.

[33] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm," *IEEE Trans. Commun.*, vol. 35, no. 9, pp. 877–887, Sep. 1987.

[34] J. Yang, J.-J. Werner, and G. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 997–1015, 2002.

[35] J.-T. Yuan and K.-D. Tsai, "Analysis of the multimodulus blind equalization algorithm in QAM communication systems," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1427–1431, Sep. 2005.

[36] X.-L. Li and W.-J. Zeng, "Performance analysis and adaptive Newton algorithms of multimodulus blind equalization criterion," *Signal Process.*, vol. 89, no. 11, pp. 2263–2273, 2009.

[37] M. Mizuno and J. Okello, "A high throughput pipelined architecture for blind adaptive equalizer with minimum latency," *IEICE Trans. Fund. Elect. Commun. Comp. Sci.*, vol. 86, no. 8, pp. 2011–2019, 2003.

[38] T.-H. Li, "Blind deconvolution of linear systems with multilevel non-stationary inputs," *Ann. Stat.*, vol. 23, no. 2, pp. 690–704, 04 1995.

[39] L. He, M. G. Amin, and C. Reed Jr., "Adaptive equalization techniques for indoor dynamic wireless communication channels," in *Proc. SPIE*, vol. 4395, pp. 28–38, Apr. 2001.

[40] M. Amin, L. He, C. Reed Jr., and R. Malkemes, "A modified constant modulus algorithm for adaptive channel equalization for QAM signals," in *Proc. IEEE SSP Workshop*, pp. 563–566, 2001.

- [41] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. on Mat Anal. Appl.*, vol. 17, no. 1, pp. 161–164, Jan. 1996.
- [42] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins University Press, 1996.
- [43] A. Mesloub, K. Abed-Meraim, and A. Belouchrani, "A new algorithm for complex non-orthogonal joint diagonalization based on Shear and Givens rotations," *IEEE Trans. Signal Process.*, vol. 62, pp. 1913–1925, Apr. 2014.
- [44] A. Labed, T. Chonavel, A. Aïssa-El-Bey, and A. Belouchrani, "Min-norm based alphabet-matching algorithm for adaptive blind equalisation of high-order QAM signals," *Trans. on Emerging Telecomm. Tech.*, vol. 24, no. 6, pp. 552–556, 2013.
- [45] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder–Mead simplex method in low dimensions," *SIAM Journal on optimization*, vol. 9, no. 1, pp. 112–147, 1998.
- [46] A. Ikhlef, "Séparation aveugle de sources dans les systèmes de communication mimo," Ph.D. dissertation, Université de Reims, 2008.
- [47] P. Chevalier, "Optimal separation of independent narrow-band sources: Concept and performance," *Signal process.*, vol. 73, no. 1, pp. 27–47, 1999.