An Adaptive CM Array Preconditioner for Blind Multi-User Separation

Stanislaw Gorlow, João Paulo C. L. da Costa, and Martin Haardt

Abstract—The family of constant-modulus algorithms is widely used in wireless communication systems and in radar. The classical constant-modulus adaptive (CMA) algorithm, however, fails to lock onto a single mode when used in conjunction with an antenna array. Instead, it equalizes the entire spatial spectrum. In this paper, we describe in full detail our recently proposed approach for the separation of multiple users in a radio system with frequency reuse, such as a cellular network, making use of the CMA algorithm. Based on the observation that the differential filter weights resemble a superposition of the array steering vectors, we cast the original task to a direction-of-arrival estimation problem. With rigorous theoretical analysis of the array response based on the discrete-space Fourier transform we elaborate a solution that solves the problem by finding the roots of a polynomial equation. We provide a numerical example to demonstrate the validity of the approach under high-SNR conditions. In addition, we propose a more general preprocessor for the CMA array which allows the modulated signals to differ in amplitude. As a byproduct, the preprocessor yields a low-cost estimate of the number of concurrent users, i.e. the model order, by simply counting the roots with the strongest response.

Index Terms—Adaptive and array signal processing, constant modulus modulation, blind multi-user separation, discrete-space Fourier transform, polynomial roots, cellular frequency reuse.

I. Introduction

A constant-modulus adaptive (CMA) algorithm represents a class of adaptive digital filtering algorithms that are sensitive to amplitude modulation and insensitive to angle modulation. Most commonly, it is realized as a transversal finite impulse response (FIR) digital filter based on gradient descent. In its original formulation the CMA algorithm compensates for the frequency-selective multipath and interference on signals that have a constant envelope. Such radio signals can be generated using frequency modulation or quadrature phase-shift keying (QPSK) [2]. Despite its close resemblance to the least mean squares (LMS) filter, the main feature of the CMA filter is that it does not require a reference signal to operate. For a topic review, the reader is referred to [3].

In [2] it is argued that the CMA filter is not only useful to counter multipath induced fading, but also to reduce narrow-band or sinusoidal interference if the interferer's power is significantly below the carrier power. In that case, the CMA filter acts as a notch filter. Nevertheless, in order to deal with a strong interferer, extra measures, such as prefiltering, are

necessary. In [4], based on the model of two sinusoids at the input and nonoverlapping spectra, two different solutions can be found for the filter: one to suppress the interferer and another to capture the signal of interest. There are cases, however, in which the interferer is mistakenly captured as the signal of interest.

With regard to the convergence of the CMA array, we find works such as [5] or [6]. Both rely on the capture effect of the CMA filter, i.e. the assumption that the algorithm will lock onto the strongest CM signal with a particular polarization or direction of arrival. In [7], however, it is mentioned that even in the case of two sources that are separated by an angle of $90°$ the array's capture effect may fail, as the filter may switch arbitrarily from one source to the other. The problem of capturing one particular source in the presence of multiple CM sources is further acknowledged in [8]. Without providing any evidence, it is inferred that the problem can be solved in multiple stages. For this to be possible, the CMA array would need to lock in each stage, which is not guaranteed.

Various application areas for CM algorithms can be found in the literature. In [9], a CM algorithm serves as the basis for semi-blind adaptive beamforming in an array-based communication system. In [10], a CM algorithm is employed to estimate the carrier-frequency offset in an optical system. In [11], a CM algorithm is used to design the probing waveform for a MIMO radar. In this paper, we present a complete derivation of our recent approach to the blind multi-user separation problem based on the CMA array [1] and extend it by a more general adaptive preprocessor. The weights of the preprocessor are the coefficients of a polynomial in a single variable z with the constant term equal to one. The number of roots of the polynomial that lie near the unit circle in the z-plane represent the model order, while the arguments of the complex roots correspond to the angles of the sources. From the angles one can construct a response matrix and *condition* the CMA array in such a way that it is likely to focus on the user of interest and to reject all the competing users.

The rest of the paper is organized as follows. In Section II, we define the data model of an antenna array and formulate the multi-user separation problem. In Section III, we review the classical literature on the CMA algorithm and conclude the section with more remarks concerning the capture effect. Section IV defines the discrete-space Fourier transform that provides the mathematical basis for our approach. In Section V, we elaborate our extension of the CMA algorithm for the antenna array. In Section VI, we present a more practical solution in the form of an adaptive preprocessor for the CMA array. Conclusions are drawn in Section VII.

S. Gorlow is with Dolby Sweden, Stockholm, Stockholm County, 113 30 Sweden e-mail: szgorl@dolby.com. Parts of this work have been published in [1].

J. P. C. L. da Costa is with the University of Brasília. This work was partially funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under grant number 303905/2014-0.

M. Haardt is with the Ilmenau University of Technology.

II. DATA MODEL AND PROBLEM STATEMENT

In this section, we present the data model for a frequency-reuse cellular network and state the blind multi-user separation problem which we seek to solve using the constant-modulus adaptive algorithm. We consider a scenario in which multiple sources (users) broadcast signals at the same frequency and at the same time at distinct locations from potentially different cells. The CM signals are received by a base station equipped with (structurally) identical antennas arranged as an array. By combining the antenna outputs, the objective is to separate the CM signals with minimal interference.

A. Data Model

Consider the linear data model

$$\mathbf{X} = \mathbf{A}\,\mathbf{S}^{\mathsf{H}} + \mathbf{N}, \tag{1}$$

where $\mathbf{S}^{\mathsf{H}}$ is the Hermitian transpose of the signal matrix $\mathbf{S} \in \mathbb{C}^{N \times D}$,

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_D \end{bmatrix} \tag{2}$$

with the dth column belonging to the dth source signal, $\mathbf{A} \in \mathbb{C}^{M \times D}$ is the array response matrix, the columns of which are vectors on the array manifold associated with a direction of arrival (DOA) or a spatial frequency,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_D \end{bmatrix}, \tag{3}$$

$\mathbf{N} \in \mathbb{C}^{M \times N}$ is a (zero-mean) Gaussian noise matrix, and $\mathbf{X} \in \mathbb{C}^{M \times N}$ carries N temporal snapshots collected by M antennas ($N \gg M$). On the assumption that the antenna elements are arranged as a uniform linear array (ULA), $\mathbf{A}$ has the structure of a Vandermonde matrix, i.e.

$$\mathbf{a}_d = \begin{bmatrix} 1 & z_d & z_d^2 & \cdots & z_d^{M-1} \end{bmatrix}^{\mathsf{T}} \quad \text{with } z_d = \mathrm{e}^{\mathrm{j}2\pi\xi_d}, \tag{4}$$

where e is Euler's number, j is the imaginary unit, and ξ_d is the spatial frequency,

$$\xi_d = \frac{\Delta}{\lambda}\,\sin\theta_d, \tag{5}$$

where Δ is the spacing between the elements of the antenna array, λ is the carrier wavelength, and θ_d is the inclination angle of the dth wavefront. We assume a stationary propagation environment in which the channel has a short delay spread, so that temporal equalization can be omitted.

B. Problem Statement

In the noiseless case, the problem at hand can be formulated as a structured matrix factorization problem:

$$\mathbf{X} = \hat{\mathbf{A}}\,\hat{\mathbf{S}}^{\mathsf{H}}$$
$$\text{s.t. } \operatorname{rank}\hat{\mathbf{A}} = \operatorname{rank}\hat{\mathbf{S}} = D \text{ and } |\hat{s}_{nd}| = 1. \tag{6}$$

The task is to find the factors $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$ for a given $\mathbf{X}$, with $\hat{\mathbf{S}}$ satisfying the constant modulus property. In the beamforming context, the problem can also be formulated as:

$$\mathbf{W}^{\mathsf{H}}\mathbf{X} = \hat{\mathbf{S}}^{\mathsf{H}}$$
$$\text{s.t. } \operatorname{rank}\hat{\mathbf{S}} = D \text{ and } |\hat{s}_{nd}| = 1. \tag{7}$$

Once $\mathbf{W}^{\mathsf{H}}$ is known, an estimate of the array response matrix $\hat{\mathbf{A}}$ is given by the Moore–Penrose pseudoinverse of $\mathbf{W}^{\mathsf{H}}$. As a consequence, $\mathbf{W}^{\mathsf{H}}$ would provide the least-squares estimate for $\mathbf{S}$ under the constant modulus constraint. The task of the blind beamformer is to compute the proper weight vectors from the measured data only, without detailed knowledge of the signals and the channel.

III. CONSTANT-MODULUS ADAPTIVE ALGORITHM

In this section, we give a review of the available literature related to the constant-modulus adaptive algorithm. We start from the single-channel transversal filter and end the section with the multi-channel spatial filter: the CMA array.

A. Single-Channel Transversal Filter

The original (single-channel) CMA filter is discussed at full length in [2]. In the following, we give a summary of how to design it. First, consider a sampled quadrature signal $x(n)$ to be frequency modulated and multipath distorted. This signal, which is analytic and hence complex, passes though a tapped delay line with adjustable complex weights, which represents the transversal FIR filter with filter coefficients $\mathbf{w}$. Then, the complex filter output can be written as

$$y(n) = \mathbf{w}^{\mathsf{H}}(n)\,\mathbf{x}(n), \tag{8}$$

where

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-N+1) \end{bmatrix}^{\mathsf{T}} \tag{9}$$

is the data in the delay line at sampling instant n and

$$\mathbf{w}(n) = \begin{bmatrix} w_0(n) & w_1(n) & \cdots & w_{N-1}(n) \end{bmatrix}^{\mathsf{T}} \tag{10}$$

is the vector of N adjustable filter coefficients. Conveniently, we assume that the coefficients are tweaked for each n.

The objective is to restore the output $y(n)$ to a form, which on average has a constant instantaneous modulus. It is achieved by choosing the coefficients $\mathbf{w}$ in such a way that the following cost or performance function is minimized:

$$J = \frac{1}{4}\,\mathrm{E}\left\{\left[|y(n)|^2 - 1\right]^2\right\}, \tag{11}$$

where E denotes expectation. Note that in (11) the modulus of the unknown transfer signal is normalized to unity. Hence, J is a differentiable positive measure of the average amount by which the output $y(n)$ deviates from unit modulus, see Fig. 1. Also note the inherent phase ambiguity in the cost.

With the cost function and the filter structure given, (11) is minimized using stochastic gradient descent. In particular, $\mathbf{w}$ is updated according to the recurrence relation

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \gamma\,\nabla_{\mathbf{w}}J(n), \tag{12}$$

where γ is the step size (adaptation constant) and $\nabla_{\mathbf{w}}$ is the gradient with respect to the filter coefficients. The stochastic gradient,

$$\nabla_{\mathbf{w}}J(n) = \mathrm{E}\left\{\left[|y(n)|^2 - 1\right]\mathbf{x}(n)\,y^*(n)\right\}, \tag{13}$$
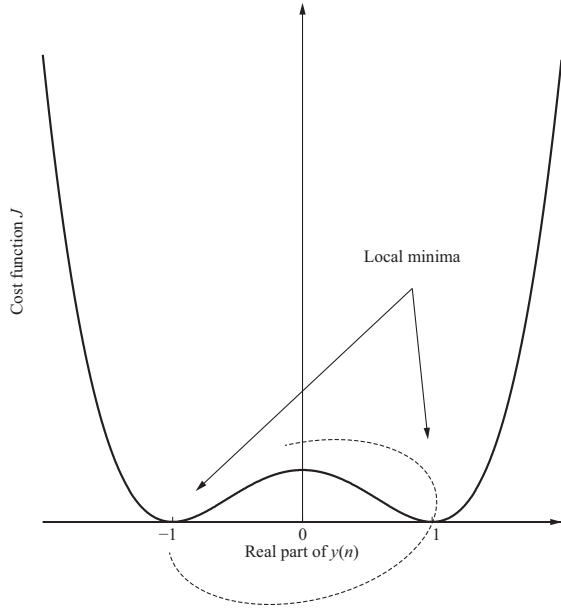
Fig. 1. J as a function of $\Re\{y(n)\}$ for $\Im\{y(n)\} = 0$. Any local minimum on the unit circle around the origin satisfies $J = 0$. Positive deviations from unit modulus are penalized more than negative deviations.

can be approximated by the instantaneous gradient leaving out the expectation. The update rule from (12) then reads:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \gamma \left[|y(n)|^2 - 1\right] \mathbf{x}(n)\, y^*(n). \quad (14)$$

By defining

$$\left[1 - |y(n)|^2\right] y(n) \stackrel{\text{def}}{=} e(n), \quad (15)$$

the update rule (14) becomes

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \gamma\, \mathbf{x}(n)\, e^*(n), \quad (16)$$

which appears to be formally identical with the LMS filter.

Please note that $y(n) = 0$ is also a stationary point (a local maximum) for the cost function J, see (13). However, due to (8) and since $\mathbf{x}(n) \neq \mathbf{0}\ \forall n$, this point can only be reached iff $\mathbf{w}(n) = \mathbf{0}$. This is avoided by setting $\mathbf{w} \neq \mathbf{0}$ at initialization.

B. Multi-Channel Transversal Filter

In [12], a multi-channel extension of the above CMA filter is presented in the form of a polarization diversity combiner. By means of polarization beam steering, the combiner rejects additive interference and compensates for channel-induced polarization rotation at the same time. The authors provide an example in which a two-channel processor with one channel for the vertical and the other for the horizontal polarization successfully separates an FM signal from a QPSK signal, both transmitted over orthogonal channels that exhibit frequency-selective multipath and cross talk. It should be noted that the two signals have an identical nominal carrier frequency and the same power. And although the result is obtained in the high-SNR regime, i.e., at an SNR of 20 dB at the receiver feeds, it looks promising. According to the authors, the multi-channel processor can be employed in combination with any other orthogonal scheme. Still, it is not quite clear what such a scheme would look like for more than two transmitter feeds.

Two possible implementations of a two-channel processor are discussed below.

1) One-Stage Processor: The one-stage processor has one pair of weight vectors, $\mathbf{w}_1$ and $\mathbf{w}_0$, for a CM signal transmitted on either the horizontal or the vertical feed. Each tapped delay line on the receiver side is initialized in such a way as to have an all-pass response for the desired signal or to fully attenuate the cross polarization by setting all weights to zero. The output from the two filters is summed to yield the desired CM signal by which the weight vectors are adjusted. Hence, if $x_h(n)$ is the signal of interest and $x_v(n)$ is the interferer, the estimate is formed by

$$y_h(n) = \mathbf{w}_1^{\mathsf{H}}(n)\, \mathbf{x}_h(n) + \mathbf{w}_0^{\mathsf{H}}(n)\, \mathbf{x}_v(n), \quad (17)$$

where

$$\begin{aligned}
\mathbf{w}_1(n_0) &= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^{\mathsf{T}}, \\
\mathbf{w}_0(n_0) &= \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}^{\mathsf{T}},
\end{aligned} \quad (18)$$

and

$$\begin{aligned}
\mathbf{x}_h(n) &= \begin{bmatrix} x_h(n) & \cdots & x_h(n-N+1) \end{bmatrix}^{\mathsf{T}}, \\
\mathbf{x}_v(n) &= \begin{bmatrix} x_v(n) & \cdots & x_v(n-N+1) \end{bmatrix}^{\mathsf{T}}.
\end{aligned} \quad (19)$$

By stacking the weight vectors to a single vector

$$\mathbf{w}_h(n) = \begin{bmatrix} \mathbf{w}_1(n) \\ \mathbf{w}_0(n) \end{bmatrix} \quad (20)$$

and by doing the same for the signal vectors,

$$\mathbf{x}(n) = \begin{bmatrix} \mathbf{x}_h(n) \\ \mathbf{x}_v(n) \end{bmatrix}, \quad (21)$$

we see that (17) is equivalent to

$$y_h(n) = \mathbf{w}_h^{\mathsf{H}}(n)\, \mathbf{x}(n), \quad (22)$$

which again is identical with the single-channel filter (8). As a result, we may state that space diversity is not exploited in that scheme. The interference rejection performance is hence subject to the degree of orthogonality of the transmitter feeds, the mechanical alignment of the transmitter with the receiver feeds, the channel-induced polarization rotation, but also the SIR at the receiver. As a rule of thumb, the stronger the CM signal and the weaker the cross talk, the better the outcome. With the above definitions, the processor is updated as per

$$\begin{aligned}
\begin{bmatrix} \mathbf{w}_h(n+1) \\ \mathbf{w}_v(n+1) \end{bmatrix} &= \begin{bmatrix} \mathbf{w}_h(n) \\ \mathbf{w}_v(n) \end{bmatrix} - \gamma \left\{ \begin{bmatrix} |y_h(n)|^2 \\ |y_v(n)|^2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \\
&\quad \cdot \mathbf{x}(n) \begin{bmatrix} y_h^*(n) \\ y_v^*(n) \end{bmatrix}
\end{aligned} \quad (23)$$

with

$$\mathbf{w}_v(n) = \begin{bmatrix} \mathbf{w}_0(n) \\ \mathbf{w}_1(n) \end{bmatrix} \quad (24)$$

and

$$y_v(n) = \mathbf{w}_v^{\mathsf{H}}(n)\, \mathbf{x}(n), \quad (25)$$

respectively.

2) Two-Stage Processor: The above single-stage processor can be realized in two stages with an adaptive signal canceller (ASC) in between. The approach has shown to be more robust to alignment imperfections and polarization crosstalk between

two equally powered signals. The first stage filters out the FM signal, followed by the QPSK signal in the second stage. This order is preferred because the CMA filter has a more negative effect on the QPSK signal in the presence of an interferer. As a direct consequence, the SIR of the QPSK signal improves after removing the FM signal from the receiver inputs using the LMS filter. The error $e(n)$ in that case is defined as

$$e(n) \stackrel{\text{def}}{=} y(n) - \underbrace{\mathbf{u}^{\mathsf{H}}(n)\,\mathbf{x}(n)}_{\hat{y}(n)}, \qquad (26)$$

where $\mathbf{u}$ denotes the LMS filter coefficients. The input to the second stage is given by

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_h(n) \\ \hat{x}_v(n) \end{bmatrix} = \begin{bmatrix} x_h(n) \\ x_v(n) \end{bmatrix} - \begin{bmatrix} \mathbf{u}_1^{\mathsf{H}}(n)\,\mathbf{x}_h(n) \\ \mathbf{u}_0^{\mathsf{H}}(n)\,\mathbf{x}_v(n) \end{bmatrix}, \qquad (27)$$

and hence

$$y_v(n) = \mathbf{w}_v^{\mathsf{H}}(n)\,\hat{\mathbf{x}}(n). \qquad (28)$$

Theoretically, this scheme can be extended to any number of stages. To put it into practice, all that is needed is one CMA filter and another LMS filter combined in recursion [12].

C. Multi-Channel Spatial Filter

A multi-channel spatial (CMA) filter is easily derived from (8), (16), and (15). Changing the independent time variable n in (9) to the channel index m indicating an antenna element, we instantly obtain the formulation of a CMA array that can be put into effect in an adaptive antenna framework [7]. One would then refer to the CMA filter as a "beamformer" that steers the beam in the direction of the CM signal and rejects any additive interference. The array's weights thus determine the "beam pattern". The adaptation of the weights rests upon the fact that interference causes fluctuations to a CM signal's amplitude. The CMA array is meant, or at least expected, to counteract such fluctuations [2].

In the presence of several CM signals, the CMA array can also be implemented as a multi-stage processor, as shown in the previous subsection, to improve the SIR for the follow-up stage. In addition, its convergence performance can considerably be improved by adaptation of the step size γ in reference to the recursive least squares (RLS) filter [7], [8], [13]:

$$\gamma_n = \frac{\mathbf{P}(n-1)}{\alpha + \mathbf{x}^{\mathsf{H}}(n)\,\mathbf{P}(n-1)\,\mathbf{x}(n)} \qquad (29)$$

with

$$\mathbf{P}(n) = \left[1 - \frac{\mathbf{P}(n-1)\,\mathbf{x}(n)\,\mathbf{x}^{\mathsf{H}}(n)}{\alpha + \mathbf{x}^{\mathsf{H}}(n)\,\mathbf{P}(n-1)\,\mathbf{x}(n)} \right] \\ \cdot \frac{1}{\alpha}\,\mathbf{P}(n-1), \qquad (30)$$

where α is the forgetting factor, usually chosen between 0.98 and 1, and $\mathbf{P}(n)$ denotes the inverse of the weighted sample correlation matrix for $\mathbf{x}(n)$. The latter is initialized as

$$\mathbf{R}_{\mathbf{xx}}(n_0) = \sigma^2\,\mathbf{I}_M, \qquad (31)$$

where $\mathbf{I}_M$ is an M-by-M identity matrix and σ has a positive value. Note that in (29) γ_n is no longer a scalar but rather a normalized inverse of $\mathbf{R}_{\mathbf{xx}}(n)$. That being the case, γ_n has the function of orthogonalizing the modes of $\mathbf{x}(n)$, which in turn speeds up the convergence.

D. Further Remarks

In [4], in the attempt to predict the algorithm's behavior, the authors derive conditions that allow to control the algorithm's general tendency. They conclude that the CMA filter's ability to capture the desired signal is subject to a combination of the following factors: the signals' amplitudes at the input and the output, i.e. the SIRs, and the initial filter response. It should be pointed out, however, that most of the analysis applies only to the special case where the two time series are, in a broader sense, *orthogonal* (see polarization). Furthermore, the authors also mention that according to their convergence analysis it is not sufficient to initialize the filter with an all-pass response when the amplitude of the signal of interest is near unity. The amplitude of the interferer has to be considered as well.

Especially from a spatial filtering perspective, where space diversity and frequency reuse play a major role, it is hard to see from the cost function why the phased array should steer the beam in the direction of one CM source but drive to zero the contribution from all other CM sources. Firstly, all QPSK signals impinging on the array could be considered as equally powered due to the far-field assumption. Secondly, since no reference is given in (11), mutual independence of the set of CM sources is not exploited and so is irrelevant. A complete suppression of the interfering CM sources is realistic only if nulls are put explicitly in their directions by means of linear constraints, or by computation of the pseudoinverse of $\mathbf{A}$. In both cases, the angles of the incoming wavefronts have to be known a priori. This, however, is an unrealistic scenario.

In fact, our experiments suggest that the CMA array more likely processes the CM sources as a whole, in proportion to the sources' amplitudes. Therefore, it does not seem to show any preference towards one particular CM source, neither for the strongest. The totality of all CM sources is treated as one single source with several multipath components, in the case of a channel with a long delay spread, or as several but fully correlated sources, if the delay spread is short. For this, it is impossible for the CMA array to distinguish several users in a frequency-reuse radio system that uses, e.g., QPSK. What it does instead is to adjust the filter weights in such a way that the filtered sum of all CM signals exhibits unit modulus. By transmitting two signals on different polarizations, the SIR is improved due to crosstalk reduction, given that the feeds are mechanically aligned. For all the reasons adduced above, we draw the conclusion that there is no capture effect that could be assigned to the CMA filter in a general sense.

In summary it can be stated that the CMA array initialized with an all-pass response does not capture one particular CM signal within a mixture, but rather whitens the mixture. That means that the CMA array equalizes the channel inclusive of the radiation pattern of a phased array. This basic explanation is furthermore consistent with every observation or statement made by other authors. Fig. 2 illustrates the case where three QPSK sources send data from different angles and at different power levels. The corresponding radiation pattern is shown in
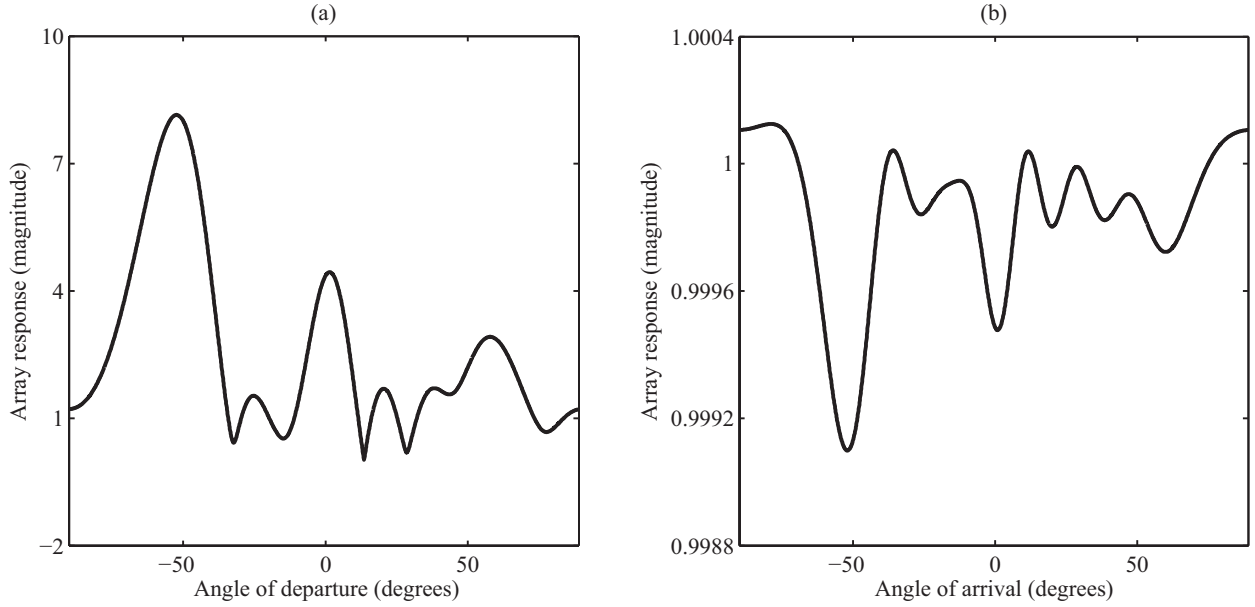
Fig. 2. Illustrative example of (a) the radiation pattern of three QPSK sources and (b) the response of the CMA array (baseband).

(a). The response of the CMA array after 8,000 iterations is shown in (b), averaged over 100 simulations. It is worth to be pointed out that the algorithm did not converge. The average output modulus is $E[|y(n)|] = 1.084$. The CMA array acts as a soft equalizer of the radiation pattern without locking onto any mode. The array consists of 8 elements and the SNR is 20 dB with respect to the strongest mode.

IV. DISCRETE-SPACE FOURIER TRANSFORM

Before solving the multi-user separation problem using the output from the original single-stage CMA array, let us introduce the discrete-space Fourier transform. The latter, together with the z-transform, is made use of in the following Section V, so as to have a formal basis for the pursued approach and to justify the results.

A. Definition

The discrete-space Fourier transform (DSFT) of a series of real (or complex) numbers is a Fourier series that is periodic w.r.t. the angular spatial frequency variable μ, also known as the angular wave number. The DSFT is defined as:

$$x(m) \circlestcirc\!\!- X\left(e^{j\mu}\right) = \sum_{m=-\infty}^{\infty} x(m) e^{-j\mu m}. \tag{32}$$

The corresponding inverse is given by

$$X\left(e^{j\mu}\right) -\!\!\circ x(m) = \frac{1}{2\pi} \int_{2\pi} X\left(e^{j\mu}\right) e^{j\mu m} d\mu. \tag{33}$$

B. Transform Pairs and Properties

Although many more transform pairs may be found in the vast amount of literature on the Fourier transform, let us seek out the two that are most relevant to our problem:

TABLE I
MOST RELEVANT TRANSFORM PAIRS

Space series $x(m)$	$\circ\!\!-$	Fourier transform $X\left(e^{j\mu}\right)$
$1 \ (-\infty < m < \infty)$		$2\pi \sum_{\kappa=-\infty}^{\infty} \delta(\mu - 2\pi\kappa)$
$x(m) = \begin{cases} 1 & \text{for } 0 \leqslant m < M \\ 0 & \text{otherwise} \end{cases}$		$e^{-j\mu \frac{M-1}{2}} \dfrac{\sin\left(M\frac{\mu}{2}\right)}{\sin\left(\frac{\mu}{2}\right)}$

Note that the Fourier transform of the second pair is a variant of the Dirichlet kernel for a sum over all nonnegative m. The corresponding space series is equivalent in notation to

$$x(m) = \text{rect}\left[\frac{m - \frac{M-1}{2}}{M - 1}\right], \tag{34}$$

where rect $\cdot$ denotes the rectangular function. In Table I, $\delta(\mu)$ stands for the Dirac delta function. To complete this section, let us recall two main operations in the space domain and the effect they have on the frequency domain.

1) Shift in Frequency: The first property is also referred to as frequency modulation by a (real) offset μ_0. The respective Fourier transform pair is:

$$e^{j\mu_0 m} x(m) \circ\!\!- X\left[e^{j(\mu-\mu_0)}\right]. \tag{35}$$

2) Multiplication in Space: The other property states that a multiplication of two sequences in the space domain results in a periodic convolution in the frequency domain. Thus, the corresponding transform pair is:

$$w(m) x(m) \circ\!\!- \frac{1}{2\pi} \int_{-\pi}^{\pi} W\left(e^{j\nu}\right) X\left[e^{j(\mu-\nu)}\right] d\nu. \tag{36}$$

In the following subsections, we will use the transform pairs

and properties from above to derive an analytic expression of the response of a beamformer such as the CMA array.

C. Transform of a Finite-Length Sequence

The expression in (32) applies to an infinite series $x(m)$. In order to analytically evaluate the DSFT of a finite-length data sequence, we apply a rectangular window $w(m)$ of length M to the input sequence $x(m)$, resulting in

$$X\left(e^{j\mu}\right) = \sum_{m=-\infty}^{\infty} w(m)\,x(m)\,e^{-j\mu m} \qquad (37)$$

$$= \sum_{m=0}^{M-1} x(m)\,e^{-j\mu m} \qquad (38)$$

with $w(m)$ as in (34). Using the transform pair from (36), we can see that the DSFT $X\left(e^{j\mu}\right)$ of a finite-length sequence as in (38) is formally identical with

$$\widetilde{X}\left(e^{j\mu}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\nu \frac{M-1}{2}} \frac{\sin\left(M\frac{\nu}{2}\right)}{\sin\left(\frac{\nu}{2}\right)} \\ \cdot X\left[e^{j(\mu-\nu)}\right] d\nu. \qquad (39)$$

In other words, to compute the DSFT of a *finite* series is the same as to convolve the transform of an *infinite* series with the causal Dirichlet kernel, see Table I. Likewise, (39) is a Fourier series approximation of degree $M-1$, see [14], [15].

D. Fourier Series and Array Response

Consider the case where we seek to compute the DSFT of the array steering vector consisting of the elements $a_m = e^{j\mu_0 m}$, $m = 0, 1, \ldots, M-1$, which can be written as $a(m) = e^{j\mu_0 m} x(m)$ with $x(m) = 1$. Resorting to the first correspondence from Table I and due to (35), once can easily show that, for $-\infty \leqslant m \leqslant \infty$,

$$a(m) \; \circ\!\!-\!\!-\; A\left(e^{j\mu}\right) = 2\pi \sum_{\kappa=-\infty}^{\infty} \delta(\mu - \mu_0 - 2\pi\kappa). \qquad (40)$$

In further consequence, for $m = 0, 1, \ldots, M-1$,

$$A\left(e^{j\mu}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j(\mu-\nu)\frac{M-1}{2}} \frac{\sin\left(M\frac{\mu-\nu}{2}\right)}{\sin\left(\frac{\mu-\nu}{2}\right)} \\ \cdot 2\pi\,\delta(\nu - \mu_0)\,d\nu \\ = e^{-j(\mu-\mu_0)\frac{M-1}{2}} \frac{\sin\left(M\frac{\mu-\mu_0}{2}\right)}{\sin\left(\frac{\mu-\mu_0}{2}\right)}. \qquad (41)$$

To derive (41), we use the sifting property of the Dirac delta function. From (41) we hence conclude that the DSFT of the array steering vector is given by the (causal) Dirichlet kernel shifted by μ_0 from zero. It has a global maximum at μ_0 with a value of

$$A_{\max}\left(e^{j\mu}\right) = \lim_{\mu \to \mu_0} A\left(e^{j\mu}\right) = M, \qquad (42)$$

see Fig. 3.

The relation to the array response is established by taking a closer look at its definition. It is very common to write the
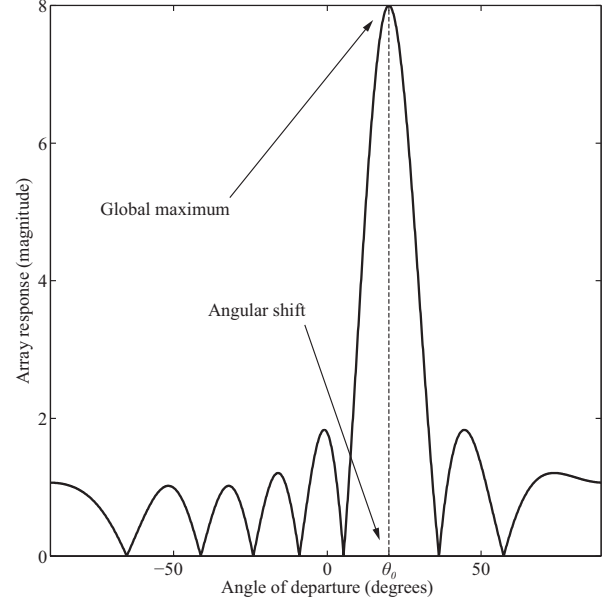

Fig. 3. Magnitude response of a steering vector ($M = 8$, $\Delta = \lambda/2$, $\theta_0 = 20°$).

output as in (8). Hence, the array or beam response $b(\mu)$ can be formulated as

$$b(\mu) = \mathbf{w}^H\,\mathbf{a}(\mu), \qquad (43)$$

which is a function of the direction and the weights. Now, if we plug the elements of the steering vector $a_m(\mu)$ into (43), it reads

$$b(\mu) = \sum_{m=0}^{M-1} w_m^*\,e^{j\mu m}. \qquad (44)$$

Comparing (44) with (38), we see that

$$b(\mu) \equiv W^*\left(e^{j\mu}\right). \qquad (45)$$

The equation above shows the identity between the response of a ULA, which is described by the steering vector, and the complex conjugate of the DSFT of the weights. Thus, one can conclude that the magnitude responses are identical, whereas the phase responses are mirrored about the μ-axis. From this, one can deduce that the beam pattern, which is the magnitude squared of the beam response, is identical with the DSFT of the weights, when taking the magnitude squared of it, too. A final remark is that (42) is equivalent to evaluating the beam response at μ_0 with the weights adjusted according to $w_m = e^{j\mu_0 m}$, which is the same as computing the squared ℓ_2-norm of $\mathbf{a}$, since

$$\|\mathbf{a}\|_2^2 = \mathbf{a}^H\,\mathbf{a} = \sum_{m=0}^{M-1} |a_m|^2 = M. \qquad (46)$$

The squared ℓ_2-norm of the steering vector is thus given by the number of the array elements M. Also note that the DSFT derived in this section refers to an array response matrix that has the Vandermonde structure. This further implies that the first array element is the reference element, which is lying in the origin of the space domain.

E. Transform of a Sum of Steering Vectors

We close this part with a brief analysis of the special case where the space sequence is generated by superposition of a number of distinct array vectors, i.e.,

$$a(m) = \sum_{d=1}^{D} \mathbf{a}_{d_m} = \sum_{d=1}^{D} e^{j\mu_d m}, \tag{47}$$

for $m = 0, 1, \ldots, M - 1$. Since the DSFT is linear, it follows that the DSFT of a sum of finite space sequences is equal to the sum of DSFTs of the sequences alone. And therefore,

$$A\left(e^{j\mu}\right) = \sum_{d=1}^{D} e^{j(\mu-\mu_d)\frac{M-1}{2}} \frac{\sin\left(M\frac{\mu-\mu_d}{2}\right)}{\sin\left(\frac{\mu-\mu_d}{2}\right)} \tag{48}$$

with

$$A\left(e^{j\mu}\right)\Big|_{\mu=\mu_i} \overset{(42)}{=} M$$
$$+ \sum_{\substack{d=1 \\ d \neq i}}^{D} \underbrace{e^{j(\mu_i-\mu_d)\frac{M-1}{2}}}_{\text{phase term}} \frac{\sin\left(M\frac{\mu_i-\mu_d}{2}\right)}{\sin\left(\frac{\mu_i-\mu_d}{2}\right)}, \tag{49}$$

where, in general, $A(e^{j\mu_i}) \neq M$, because of intermodulation products. However, if the difference $|\mu_i - \mu_d|$ is equal to $\frac{2\pi k}{M-1}$, $k \in \mathbb{N}_{>0}$, the phase term is ± 1, and (49) is purely real. Then, $A(e^{j\mu_i})$ has a fixed value given by

$$A\left(e^{j\mu}\right)\Big|_{\mu_i = \mu_d \pm \frac{2\pi k}{M-1}} = M + D - 1 = \frac{\|\mathbf{a}\|_2^2}{D}, \tag{50}$$

see Appendix. This can be achieved for at most

$$D_{\max} = M - 1 \tag{51}$$

different steering vectors, which corresponds to the degree of the Fourier series approximation. If the difference is smaller, there may be a noticeable phase shift from 0 around μ_i, and so the magnitude will deviate more from the value in (50). If however the difference is greater, the phase term vanishes at the sight of a relatively small value of the Dirichlet kernel as compared to the maximum of a single vector M. The impact factor of the phase term can be assessed by

$$\text{IF}_i \overset{\text{def}}{=} \frac{\left|A\left(e^{j\mu_i}\right)\right| - \frac{\|\mathbf{a}\|_2^2}{D}}{\frac{\|\mathbf{a}\|_2^2}{D}} = D\frac{\left|A(e^{j\mu_i})\right|}{\|\mathbf{a}\|_2^2} - 1 \tag{52}$$

with $\left|A(e^{j\mu_i})\right| = |b(\mu_i)|$, where $b(\mu_i) = \mathbf{a}^H \mathbf{a}(\mu_i) = \mathbf{a}^H \mathbf{a}_i$. It should be noted that in the case where the space sequence is a sum of steering vectors, as in (47), $\|\mathbf{a}\|_2^2 \neq D \cdot M$. Fig. 4 illustrates the case where the phase relation holds. Also note that there is a mismatch between the angles of departure and the location of the two local maxima.

V. ROOT CONSTANT-MODULUS ADAPTIVE ARRAY

In this section, we discuss in detail our proposed extension of the CMA array, which we call the root CMA array.

A. Exploratory Experiment

Consider three QPSK sources being placed in the far field of a receiver in a radio system with frequency reuse. Further
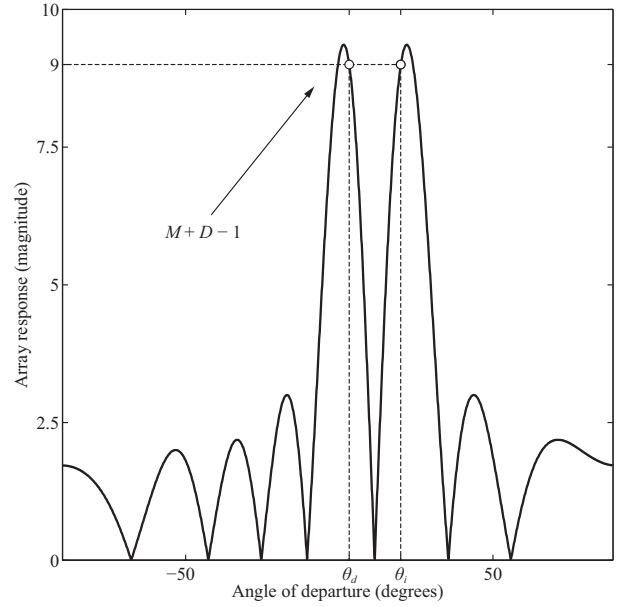

Fig. 4. Magnitude response of two superposing steering vectors for the case where $\Delta\mu = 2\pi/M{-}1$ ($M = 8$, $\Delta = \lambda/2$, $\theta_d = 3.23°$, $\theta_i = 20.0°$).

assume that the point sources are at equal distance, such that the incoming signals have the same amplitude. The angles of inclination of the three plane waves should obey (50), i.e.,

$$\theta_d = \arcsin\left[\sin\theta_i \pm \frac{\lambda}{\Delta(M-1)}\right]. \tag{53}$$

No explicit orthogonality scheme such as polarization should be in use. In Fig. 5, by way of experiment, we show that the flipped and normalized CMA array response is, in good approximation, equal to the sum of the steering vectors associated with the point sources (see also Fig. 2). The array response is flipped by inverting the direction of the gradient, converting (12) into a gradient ascent algorithm. The weight vector $\mathbf{w}(n)$ is normalized by its 2-norm after each update and finally rescaled to

$$\mathbf{v}(n) = \frac{\mathbf{w}(n)}{\|\mathbf{w}(n)\|_2} \sqrt{D^2 + D(M-1)}, \tag{54}$$

so that $\|\mathbf{v}(n)\|_2 = \|\sum_d \mathbf{a}_d\|_2$. This procedure is less sensitive to noise and the choice of γ than the one mentioned in [1].

B. Problem Restatement

On the hypothesis that $\mathbf{v}(n)$ is the sum of steering vectors, the multi-user separation problem is recast as follows. Given $\mathbf{v}(n)$, find the set of steering vectors $\{\hat{\mathbf{a}}_d\}_D$ that satisfy

$$\mathbf{v}^H(n)\hat{\mathbf{a}} \simeq \|\hat{\mathbf{a}}\|_2^2 \qquad \text{with } \hat{\mathbf{a}} = \sum_{d=1}^{D} \hat{\mathbf{a}}_d, \tag{55}$$

or, in terms of the DSFT, find

$$\left\{\hat{\mu}_d \in [-\pi, \pi) \mid \left|V^*\left(e^{j\hat{\mu}_d}\right)\right| \to C\right\} \tag{56}$$

with $C = M + D - 1$, and $d = 1, 2, \ldots, D$. In other words, we seek for the D unique spatial frequencies for the steering
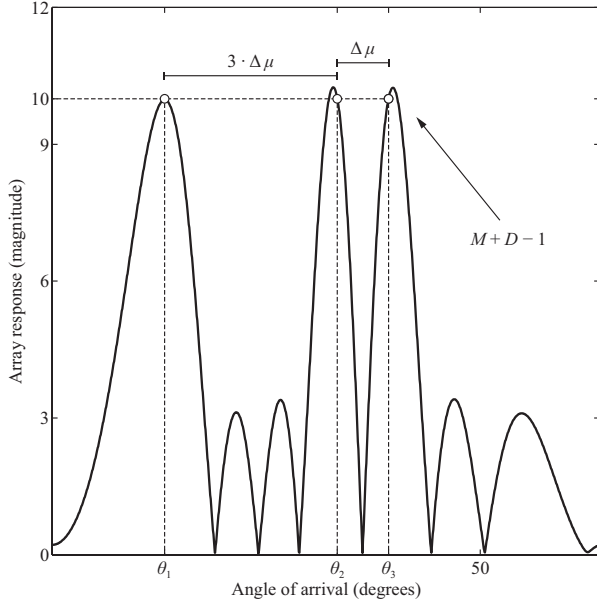
Fig. 5. Inverted and normalized magnitude response of the CMA array for $D = 3$ distinct sources for the case where the phase relation holds ($M = 8$, $\Delta = \lambda/2$, $\theta_1 = -53.2°$, $\theta_2 = 3.23°$, $\theta_3 = 20.0°$, SNR = 20 dB).
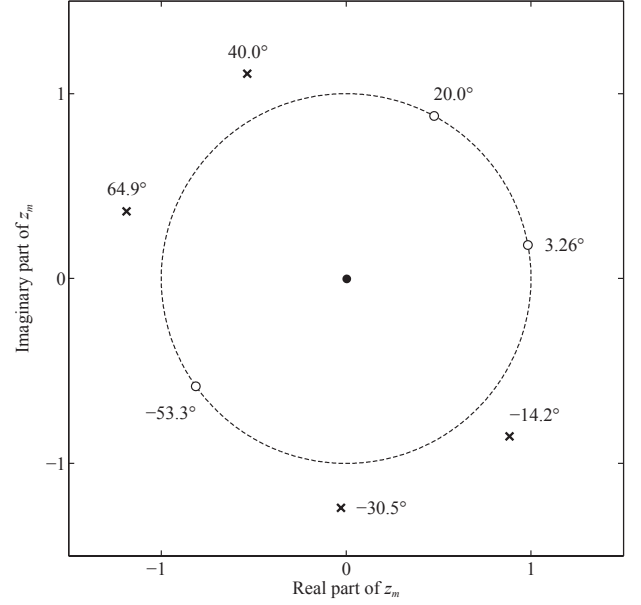

Fig. 6. Roots of $P(z)$ for the response shown in Fig. 5. The solutions are marked with circles. The estimated angles of arrival are also indicated.

vectors, the DSFTs of which in good approximation sum up to the DSFT of $\mathbf{v}(n)$.

One can, of course, go ahead and search for the D maxima of the magnitude response. However, one must be aware that not all maxima match with the underlying modes, cf. Fig. 5. In the next subsection we present a more robust approach.

C. Basic Approach

Substituting $e^{j\mu}$ for z, we postulate that

$$V^*(z) = v^*_{M-1} z^{M-1} + v^*_{M-2} z^{M-2} + \cdots + v^*_0 \qquad (57)$$
$$\simeq C \ (= M + D - 1), \qquad (58)$$

where (57) is tantamount to the z-transform of $\mathbf{v}(n)$. Now if we convert (58) into a polynomial, we obtain

$$P(z) = v^*_{M-1} z^{M-1} + \cdots + \left(v^*_0 - C\right), \qquad (59)$$

which is of degree $M - 1$. For numerical stability, we further normalize (59) by $v^*_{M-1} \neq 0$, resulting in

$$P(z) = \frac{v^*_0 - C}{v^*_{M-1}} + \sum_{m=1}^{M-1} \frac{v^*_m}{v^*_{M-1}} z^m. \qquad (60)$$

The fundamental theorem of algebra (and the factor theorem) states that $P(z)$ has $M - 1$ complex roots including possible multiplicities. Hence, the product representation of (60) is

$$P(z) = \prod_{m=1}^{M-1} (z - z_m), \qquad (61)$$

where z_m are the roots of $P(z)$. In the noise-free case, the sought-after frequencies are the arguments of the D roots that lie closest to the unit circle, see Fig. 6. The array response matrix now can be fully reconstructed with the accuracy of

$\mathbf{v}(n)$, which (for the most part) is subject to the performance of the CMA algorithm.

1) Analytical Solution for Two Sources: Consider $Q(z)$ to be generated from the roots of interest only, i.e.,

$$Q(z) = \prod_{d=1}^{D} \left(z - e^{j\mu_d}\right). \qquad (62)$$

Further consider the special case where $D = 2$, then

$$Q(z) = z^2 - \left(e^{j\mu_1} + e^{j\mu_2}\right) z + e^{j\mu_1} e^{j\mu_2}$$
$$= z^2 + q_1 z + q_0. \qquad (63)$$

The roots of (63) are given by the quadratic formula

$$z_{1,2} = -\frac{q_1}{2} \pm \sqrt{\frac{q_1^2}{4} - q_0}, \qquad (64)$$

where

$$\begin{aligned} q_1 &= -v_1(n) \\ q_0 &= \Re(q_0) + j\, \Im(q_0) \end{aligned} \qquad (65)$$

with

$$\Re(q_0) = \frac{\Re^2[v_1(n)] - \Im^2[v_1(n)]}{|v_1(n)|^2} \qquad (66)$$
$$\Im(q_0) = \sin \arccos \Re\{q_0\}.$$

For a derivation of (66) see Appendix.

2) Numerical Solution: Roots of polynomial equations in one unknown are numerically approximated by various known methods such as the method by Jenkins and Taub, Laguerre, or Durand and Kerner. More root-finding algorithms can be found in [16]. The function `roots` in MATLAB calculates the eigenvalues of the companion matrix, which in our case

has the following structure:

$$\begin{bmatrix} -\frac{v_{M-2}^*}{v_{M-1}^*} & -\frac{v_{M-3}^*}{v_{M-1}^*} & \cdots & -\frac{v_1^*}{v_{M-1}^*} & -\frac{v_0^*-C}{v_{M-1}^*} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{67}$$

The computation of eigenvalues also plays a central role in the derivation of the algebraic constant-modulus algorithm [17]–[19].

VI. ADAPTIVE CM ARRAY PRECONDITIONER

As previously mentioned, the central problem of the CMA algorithm is its inability to lock onto one single user without proper initialization of the weight vector or preprocessing. A possible solution to this problem is illustrated in the previous subsection. Nonetheless, there are two main limitations linked to that solution. Firstly, the number of users (model order) is assumed to be known a priori. Secondly, the incoming waves are supposed to have the same amplitude. In this section, we present an alternative and more practical solution in the form of an adaptive preprocessor for the CMA array. It tackles the model-order estimation problem, plus it allows the signals to differ in amplitude.

A. Motivation

The empirical evidence, which underlies our original work in [1] and which supports the hypothesis that the response of the CMA array (if inverted) is proportional to the sum of the array steering vectors when the array is initialized, e.g., with an all-pass response, can be interpreted in the following way: After the first iteration, the output is equal to the sum of the constant modulus signals. With every new iteration the array is adjusted in such a way as to minimize the cost, which is a function of the squared magnitude of $y(n)$,

$$\mathrm{E}\left\{|y(n)|^2\right\} \propto \mathbf{w}^{\mathsf{H}}(n)\,\mathbf{A}\,\mathbf{A}^{\mathsf{H}}\,\mathbf{w}(n) \quad \text{with } \{s_i(n)\} \text{ i.i.d.} \tag{68}$$

In general, (68) is positive and real for any $\mathbf{w}(n) \neq \mathbf{0}$, which includes

$$\mathbf{w} = \sum_{d=1}^{D} \mathbf{a}_d \quad \text{with } w_0 = D \tag{69}$$

as some sort of a (spatially) matched filter for the transmitted data. The corresponding polynomial is

$$P(z) = w_{M-1}^* z^{M-1} + \cdots + 1 - M. \tag{70}$$

Since $M > 1$, (70) resembles the transfer function of a linear prediction error filter for a scaled sum of D symbols. Insight and intuition tells us that in a similar manner one can design a more performant predictor that also exploits the phase, see [20].

B. Ordinary Least Squares

Ordinary least squares (OLS) is the most common and the most basic estimator that can be employed to carry out linear

prediction on the CM array. Extensions of the former with or without additional modifications are discussed and evaluated in [20] and later works, such as [21], [22].

Consider the desired or predicted signal to be equal to the signal at the first antenna element, and allow the CM signals to differ in amplitude c, i.e.

$$x_0(n) = \sum_{d=1}^{D} c_d\, s_d(n). \tag{71}$$

The optimum filter coefficients in the least squares sense are obtained by minimizing the sum of squared residuals, which are defined as the differences between $x_0(n)$ and $y(n)$, where $y(n)$ is predicted from the remaining $M - 1$ array signals. In consequence, the OLS solution yields

$$\mathbf{w} = \left(\mathbf{X}_{1,\dots,M-1}\, \mathbf{X}_{1,\dots,M-1}^{\mathsf{H}}\right)^{-1}\mathbf{X}_{1,\dots,M-1}\, \mathbf{x}_0^{\mathsf{H}} \in \mathbb{C}^{M-1}, \tag{72}$$

where

$$\mathbf{x}_0 = \begin{bmatrix} x_0(0) & x_0(1) & \cdots & x_0(N-1) \end{bmatrix} \in \mathbb{C}^{1 \times N} \tag{73}$$

and

$$\mathbf{X}_{1,\dots,M-1} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{M-1} \end{bmatrix} \in \mathbb{C}^{M-1 \times N}, \tag{74}$$

respectively.

C. Adaptive Preprocessor

In reference to OLS, the adaptive preprocessor is designed as an LMS filter with the desired signal as in (71). The LMS filter is initialized with an all-zero response $\mathbf{u} = \mathbf{0}_{M \times 1}$. Then, only the last $M - 1$ weights are updated by

$$\mathbf{u}(n+1) = \begin{bmatrix} 0 \\ u_1(n) + \gamma\, x_1(n)\, e^*(n) \\ \vdots \\ u_{M-1}(n) + \gamma\, x_{M-1}(n)\, e^*(n) \end{bmatrix}, \tag{75}$$

where

$$e(n) = x_0(n) - y(n) \tag{76}$$

and

$$y(n) = \mathbf{u}^{\mathsf{H}}(n)\, \mathbf{x}(n). \tag{77}$$

In this way, we avoid that the LMS filter converges to an all-pass response, which would be the optimum solution but not exactly what we are after. Instead, we force the filter to give us an LMS estimate of $x_0(n)$ using the delayed versions. The hypothesis is that the filter will adjust its coefficients in such a way that separate array responses point in the directions of the users. The step size can be chosen adaptively fulfilling

$$\gamma_n < \frac{2}{\|\mathbf{x}(n)\|_2^2}, \tag{78}$$

which keeps the filter stable. Fig. 7 shows the filter response after 1,000 iterations alongside the radiation pattern. One can observe that the filter response takes the value of one exactly where the modes are. That was to be expected, because $x_0(n)$
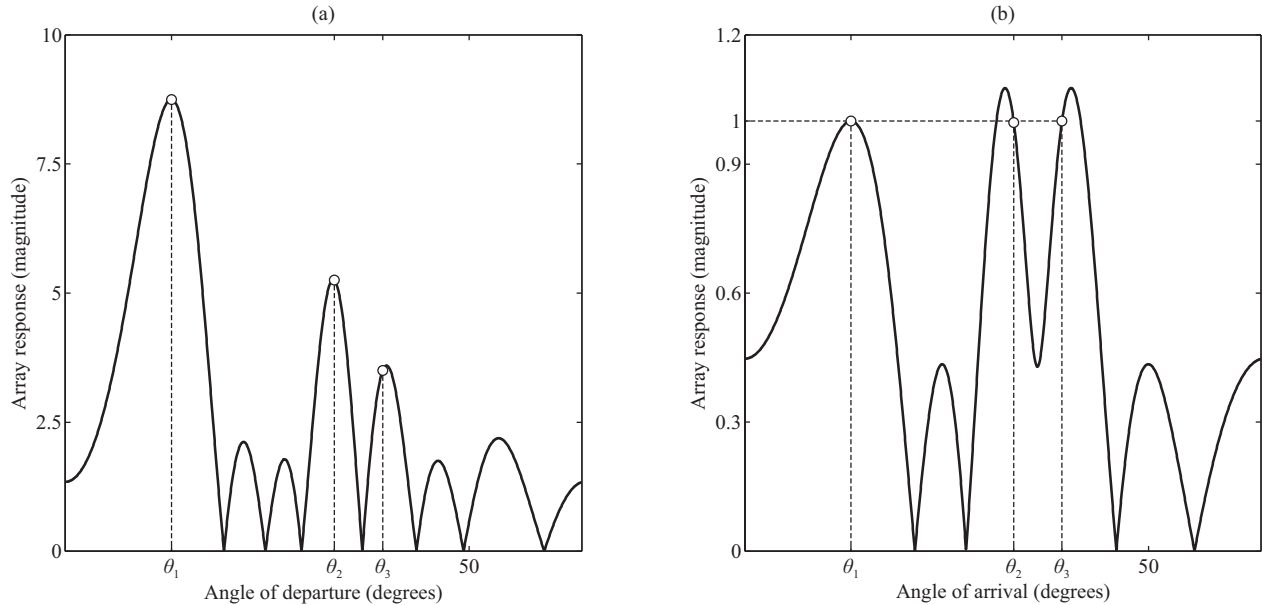
Fig. 7. Three QPSK sources send data from different angles at different power levels. The corresponding radiation pattern is shown in (a). The response of the preprocessor after 1,000 iterations is shown in (b). The array consists of 8 elements with half-wavelength spacing. The angles of interest correspond to those in Fig. 5.
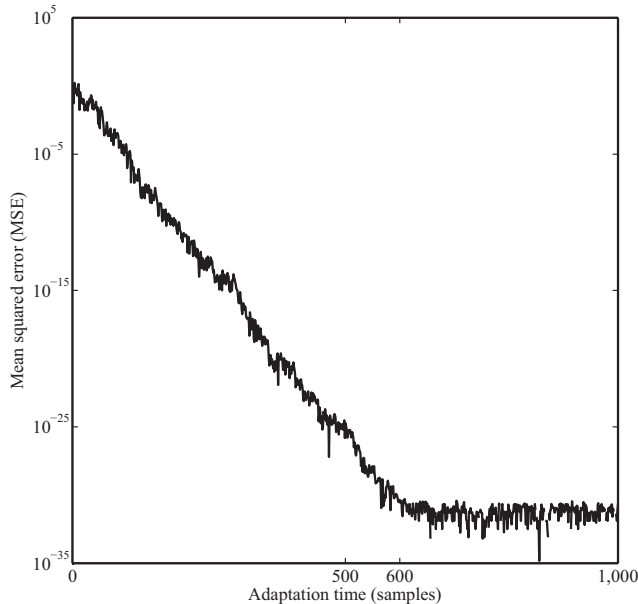

Fig. 8. Learning curve of the preprocessor in the noise-free case with $\gamma_n = 2/\|\mathbf{x}(n)\|_2^2 - \epsilon$. The saddle point is reached after 600 iterations.

and $y(n)$ share the same signal amplitudes $\{c_d\}$. Thus, we can use the same approach as before to determine the modes. In Fig. 8 we see the learning curve of the preprocessor for the noise-free case. The algorithm converges after 600 iterations for the largest possible γ according to (78). The mean squared error is virtually zero, i.e., $x_0(n)$ and $y(n)$ are almost identical.

D. Model-Order Selection

The model order is given by the number of roots that lie closest to the unit circle, see Fig. 6. We may consider, e.g.,

the distance $|z_m| - 1$ to detect that number. This is illustrated in Fig. 9. Note that in order to compute the roots correctly, C in (60) must be set to 1. As a more sophisticated alternative, and especially when the data is noisy, the Neyman–Pearson lemma can prove useful [23]. A comparison of more and other schemes is made in [24].

In our experiments, however, we found that when the noise level is high and the spatial frequencies are chosen arbitrarily, it is more reliable to evaluate the real part of the array response at the inclination angles that correspond to the $M-1$ roots (see next section) than to evaluate the distance of the roots from the unit circle. We therefore suggest to consider the D roots with the strongest (real) array response as the signal roots.

E. Direction-of-Arrival Estimation

If the model order D is determined, one can easily derive the directions of arrival from the phase of the corresponding roots. The inclination angles then are given by

$$\theta_d = \arcsin\left(\frac{\lambda \arg z_d}{2\pi \Delta}\right), \tag{79}$$

for $d = 1, 2, \ldots, D$.

F. Array Preconditioning

With the model order D known and the directions of arrival at hand, one can reconstruct the array response matrix $\mathbf{A}$. The pseudoinverse of $\mathbf{A}$, $\mathbf{A}^+$, has the property that the main lobe of each steering vector points in the direction of the respective source. All other sources, i.e., the interferers, are fully suppressed. Hence, one can initialize the weights of the CMA array with the row vector of $\mathbf{A}^+$ that corresponds to the source (user) of interest and subsequently run the CMA algorithm. The analysis of the combined performance of the
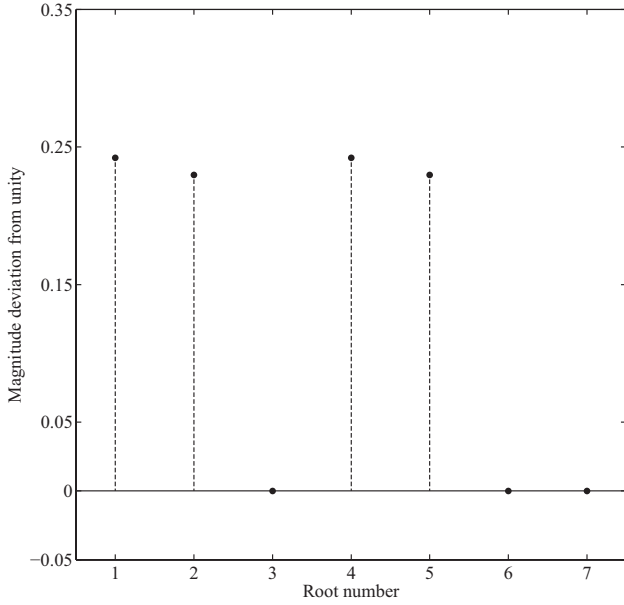
Fig. 9. Deviation of the polynomial roots from unity (noise-free case). As can be seen, three roots lie exactly on the unit circle ($D = 3$).

preconditioner and the CMA array is beyond the scope of this article.

G. Summary

In summary, the preconditioned CMA algorithm consists of the steps that are listed below:

1) Initialize LMS filter with all-zero response;
2) Consider signal at first element as desired response;
3) Update filter weights using gradient descent;
4) Exclude all-pass response;
5) Compute roots of corresponding polynomial;
6) Estimate model order;
7) Choose roots with strongest beam response;
8) Reconstruct array response matrix;
9) Run classical CMA algorithm (optional).

The individual steps are elaborated further in Fig. 10 in the form of pseudo code.

VII. CONCLUSION

We presented a novel approach for the separation of multiple sources in a frequency-reuse radio system making use of the classical CMA array. Our extension of the CMA array makes it applicable when one particular source is to be captured in the presence of multiple CM sources. The cited literature relies on an implicit capture effect of the algorithm, which is hardly sustainable looking at the formulation of the underlying cost function. The CMA array rather acts as a soft spatial equalizer. This fact can be exploited to reconstruct the array response matrix and to find the best fit for the constant modulus signals in, e.g., the least-squares sense immediately after. In addition, we designed an alternative preprocessor for the array based on the LMS filter. The second solution is more practical, because it allows the CM signals to differ in amplitude.

```
 1: procedure ROOTCMA(X, γ)
 2:     u ← [ 0 0 ⋯ 0 ]ᵀ                    all-zero response
 3:     ū ← [ 1 0 ⋯ 0 ]ᵀ                    desired response
 4:     for n ← 1, 2, …, N do
 5:         x ← [ x_{*,n} ]                      M snapshots
 6:         y ← uᴴ x                              estimate
 7:         e ← x₁ − y                               error
 8:         u ← u + γ x e*                          update
 9:         u₁ ← 0                 exclude all-pass response
10:         p ← FLIPUD(u* − ū)    prediction error polynomial
11:         z ← ROOTS(p)                  polynomial roots
12:         for d ← 1, 2, …, M − 1 do
13:             for m ← 0, 1, …, M − 1 do
14:                 a_{md} ← exp (j arg z_d m)     response matrix
15:             end for
16:         end for
17:         b ← REAL(uᴴ A)                   beam response
18:         b ← SORT(b, i)          sort in descending order
19:         D ← FIND(b > ϵ, 1) − 1             model order
20:         A ← [ a_{*,[i₁,i₂,…,i_D]} ]   strongest response matrix
21:         W ← A (Aᵀ A)⁻¹              beam steering matrix
22:         ⋯                              CMA array, etc.
23:     end for
24: end procedure
```

Fig. 10. Root constant-modulus adaptive algorithm (pseudo code).

APPENDIX

DERIVATION OF EQUATION (50)

In any such case where the phase relation $|\mu_i - \mu_d| = \frac{2\pi k}{M-1}$ holds, (49) becomes

$$
\begin{aligned}
A\left(e^{j\mu_i}\right) &= M + \sum_{\substack{d=1 \\ d \neq i}}^{D} e^{j\pi k} \frac{\sin\left(M \frac{\pi k}{M-1}\right)}{\sin\left(\frac{\pi k}{M-1}\right)} \\
&= M + \sum_{\substack{d=1 \\ d \neq i}}^{D} (-1)^k \frac{\sin\left(\frac{\pi k}{M-1} + \pi k\right)}{\sin\left(\frac{\pi k}{M-1}\right)} \\
&= M + \sum_{\substack{d=1 \\ d \neq i}}^{D} (-1)^k (-1)^k \\
&= M + D - 1 \qquad \forall M, k \in \mathbb{N}_{>0}. \qquad (80)
\end{aligned}
$$

Furthermore, if we consider that

$$
\begin{aligned}
\|a\|_2^2 &= D^2 + \sum_{m=1}^{M-1} a_m^* a_m \quad \bigg| \quad a_m = \sum_{d=1}^{D} e^{j\mu_d m} \\
&= D^2 + \sum_{m=1}^{M-1} \sum_{d=1}^{D} e^{-j\mu_d m} \sum_{i=1}^{D} e^{j\mu_i m} \\
&= D^2 + \sum_{m=1}^{M-1} \sum_{i=1}^{D} \sum_{d=1}^{D} e^{j(\mu_i - \mu_d) m} \\
&= D^2 + \sum_{m=1}^{M-1} \sum_{i=1}^{D} \sum_{d=i}^{D} \left[e^{j(\mu_i - \mu_d) m} + e^{-j(\mu_i - \mu_d) m} \right]
\end{aligned}
$$

$$= D^2 + \sum_{m=1}^{M-1} \left[D + 2 \sum_{i=1}^{D} \sum_{d=i+1}^{D} \Re \left\{ e^{j(\mu_i - \mu_d)m} \right\} \right], \quad (81)$$

and if we use the phase relation from above, (81) writes

$$\|\mathbf{a}\|_2^2 = D^2 + D(M-1) + 2 \underbrace{\sum_{m=1}^{M-1} \cos\left(\frac{2\pi \, mk}{M-1}\right)}_{= 0 \ \forall \, M, \, k \in \mathbb{N}_{>0}}$$

$$= D^2 + D(M-1) \qquad \forall \, M, k \in \mathbb{N}_{>0}. \quad (82)$$

Comparing (80) with (82), we see that

$$A\left(e^{j\mu_i}\right) = M + D - 1 = \frac{\|\mathbf{a}\|_2^2}{D}. \quad \blacksquare \quad (83)$$

Derivation of Equation (66)

For $D = 2$ and for a given $\mathbf{v}(n)$, we know that

$$e^{j\mu_1} + e^{j\mu_2} = (\cos\mu_1 + \cos\mu_2) + j(\sin\mu_1 + \sin\mu_2)$$
$$= \Re[v_1(n)] + \Im[v_1(n)]. \quad (84)$$

What we seek for is

$$e^{j\mu_1} e^{j\mu_2} = \cos(\mu_1 + \mu_2) + j\sin(\mu_1 + \mu_2)$$
$$= \Re(q_0) + \Im(q_0). \quad (85)$$

Resorting to trigonometric identities, we find that

$$\cos\mu_1 + \cos\mu_2$$
$$= 2\cos\frac{\mu_1 + \mu_2}{2} \cos\frac{\mu_1 - \mu_2}{2}$$
$$= \sqrt{[1 + \cos(\mu_1 + \mu_2)][1 + \cos(\mu_1 - \mu_2)]} \quad (86)$$

and

$$\sin\mu_1 + \sin\mu_2$$
$$= 2\sin\frac{\mu_1 + \mu_2}{2} \cos\frac{\mu_1 - \mu_2}{2}$$
$$= \sqrt{[1 - \cos(\mu_1 + \mu_2)][1 + \cos(\mu_1 - \mu_2)]}, \quad (87)$$

respectively. Dividing (86) by (87), we obtain

$$\frac{\cos\mu_1 + \cos\mu_2}{\sin\mu_1 + \sin\mu_2} = \sqrt{\frac{1 + \cos(\mu_1 + \mu_2)}{1 - \cos(\mu_1 + \mu_2)}}. \quad (88)$$

From (88) it follows that

$$\cos(\mu_1 + \mu_2)$$
$$= \frac{(\cos\mu_1 + \cos\mu_2)^2 - (\sin\mu_1 + \sin\mu_2)^2}{(\cos\mu_1 + \cos\mu_2)^2 + (\sin\mu_1 + \sin\mu_2)^2}. \quad (89)$$

By comparing the right-hand side of (84) with the right-hand side of (89), we immediately see that

$$\cos(\mu_1 + \mu_2) = \frac{\Re^2[v_1(n)] - \Im^2[v_1(n)]}{|v_1(n)|^2}. \quad \blacksquare \quad (90)$$

Acknowledgment

References

[1] S. Gorlow, J. P. C. L. da Costa, and M. Haardt, "A root-CMA algorithm for multi-user separation," in *Proc. WCNPS 2017*, Nov. 2017, pp. 1–5. [Online]. Available: https://doi.org/10.1109/WCNPS.2017.8252936

[2] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 2, pp. 459–472, Apr. 1983. [Online]. Available: https://doi.org/10.1109/TASSP.1983.1164062

[3] C. R. Johnson, Jr., 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. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct. 1998. [Online]. Available: https://doi.org/10.1109/5.720246

[4] J. R. Treichler and M. G. Larimore, "The tone capture properties of CMA-based interference suppressors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 946–958, Aug. 1985. [Online]. Available: https://doi.org/10.1109/TASSP.1985.1164640

[5] J. J. Shynk and R. P. Gooch, "Convergence properties of the multistage CMA adaptive beamformer," in *Proc. ASILOMAR 1993*, vol. 1, Nov. 1993, pp. 622–626. [Online]. Available: https://doi.org/10.1109/ACSSC.1993.342592

[6] A. V. Keerthi, A. Mathur, and J. J. Shynk, "Direction-finding performance of the multistage CM array," in *Proc. ASILOMAR 1994*, vol. 2, Oct./Nov. 1994, pp. 847–852. [Online]. Available: https://doi.org/10.1109/ACSSC.1994.471581

[7] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. ICASSP 1986*, vol. 11, Apr. 1986, pp. 2523–2526. [Online]. Available: https://doi.org/10.1109/ICASSP.1986.1168686

[8] Y. Chen, T. Le-Ngoc, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1452–1456, May 2004. [Online]. Available: https://doi.org/10.1109/TSP.2004.826167

[9] Y. Zhao, B. Ai, Y. Liu, D. Fei, and D. Li, "Semi-blind adaptive beamforming based on constant modulus algorithm for smart antennas," in *Proc. ICCSS 2017*, Jul. 2017, pp. 286–289. [Online]. Available: https://doi.org/10.1109/ICCSS.2017.8091426

[10] M. B. Balogun, O. O. Oyerinde, and F. Takawira, "Efficient constant modulus based carrier frequency offset estimation for CO-OFDM systems," *IEEE Photon. J.*, vol. 9, no. 5, pp. 1–15, Oct. 2017. [Online]. Available: https://doi.org/10.1109/JPHOT.2017.2735902

[11] Z. Cheng, Z. He, S. Zhang, and J. Li, "Constant modulus waveform design for MIMO radar transmit beampattern," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4912–4923, Sep. 2017. [Online]. Available: https://doi.org/10.1109/TSP.2017.2718976

[12] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 420–431, Apr. 1985. [Online]. Available: https://doi.org/10.1109/TASSP.1985.1164567

[13] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, no. 1/2, pp. 149–157, 1950. [Online]. Available: http://www.jstor.org/stable/2332158

[14] T. Grandke, "Interpolation algorithms for discrete Fourier transforms of weighted signals," *IEEE Trans. Instrum. Meas.*, vol. 32, no. 2, pp. 350–355, Jun. 1983. [Online]. Available: https://doi.org/10.1109/TIM.1983.4315077

[15] F. Zhang, Z. Geng, and W. Yuan, "The algorithm of interpolating windowed FFT for harmonic analysis of electric power system," *IEEE Trans. Power Del.*, vol. 16, no. 2, pp. 160–164, Apr. 2001. [Online]. Available: https://doi.org/10.1109/61.915476

[16] J. M. McNamee, *Numerical Methods for Roots of Polynomials – Part I*, 1st ed. Elsevier Science, 2007.

[17] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–431, May 1996. [Online]. Available: https://doi.org/10.1109/78.502327

[18] A.-J. van der Veen, "Asymptotic properties of the algebraic constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1796–1807, Aug. 2001. [Online]. Available: https://doi.org/10.1109/78.934150

[19] ——, "An adaptive version of the algebraic constant modulus algorithm," in *Proc. ICASSP 2005*, vol. 4, Mar. 2005, pp. 873–876. [Online]. Available: https://doi.org/10.1109/ICASSP.2005.1416148

[20] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975–989, Sep. 1982. [Online]. Available: https://doi.org/10.1109/PROC.1982.12428

[21] E. M. Dowling, R. D. DeGroat, D. A. Linebarger, L. L. Scharf, and M. Vis, "Reduced polynomial order linear prediction," *IEEE Signal Process. Lett.*, vol. 3, no. 3, pp. 92–94, Mar. 1996. [Online]. Available: https://doi.org/10.1109/97.481165

[22] A. Lopes, I. S. Bonatti, P. L. D. Peres, R. F. Colares, and C. A. Alves, "A DOA estimator based on linear prediction and total least squares," *JCIS*, vol. 17, no. 1, pp. 71–78, May 2017. [Online]. Available: https://doi.org/10.14209/jcis.2002.14

[23] S. D. Silverstein, "A new use of ESPRIT for detection and model order estimation," *IEEE Signal Process. Lett.*, vol. 1, no. 10, pp. 147–149, Oct. 1994. [Online]. Available: https://doi.org/10.1109/97.329843

[24] J. P. C. L. da Costa, A. Thakre, F. Roemer, and M. Haardt, "Comparison of model order selection techniques for high-resolution parameter estimation algorithms," in *Proc. IWK 2009*, Sep. 2009, pp. 1–6. [Online]. Available: http://www2.tu-ilmenau.de/nt/generic/paper_pdfs/paper_id_061_MOS.pdf

[25] C. B. Papadias and A. J. Paulraj, "A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays," *IEEE Signal Process. Lett.*, vol. 4, no. 6, pp. 178–181, Jun. 1997. [Online]. Available: https://doi.org/10.1109/97.586042

[26] C. Xu, G. Feng, and K. S. Kwak, "A modified constrained constant modulus approach to blind adaptive multiuser detection," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1642–1648, Sep. 2001. [Online]. Available: https://doi.org/10.1109/26.950350