Algebraic Algorithms to Separate Overlapping Secondary Surveillance Radar Replies

Nicolas Petrochilos, *Member, IEEE*, and Alle-Jan van der Veen, *Fellow, IEEE*

***Abstract*—The secondary surveillance radar (SSR) is a transponder system used in air-traffic control (ATC). Due to growing traffic densities, it is increasingly likely that a ground station receives a mixture of responses of various aircraft, partly overlapping in frequency and time. Currently such "collisions" are disregarded, at a loss of system performance and reliability. In this paper, we propose to equip the ground station with an antenna array, and investigate techniques to blindly separate such a mixture based on source waveform properties. At baseband, a received SSR signal consists of a binary sequence with alphabet $\{0, 1\}$, modulated by a complex exponential due to the residual carrier frequency. We present three algebraic algorithms to compute the separating beamformers by taking into account the particular modulation format of the received signal. The Cramér–Rao bound (CRB) is derived, extensive simulations are presented, and an experimental platform has been built to collect measurement data and demonstrate the algorithms.**

***Index Terms*—Array signal processing, blind source separation, deterministic algorithms, secondary radar.**

I. Introduction

SECONDARY surveillance radar (SSR) is essential for air-traffic control (ATC). Unlike the primary radar, the SSR is a communication radar (transponder system) that informs the ATC about the identity and the altitude of aircraft [1]. An SSR ground station uses a rotating scanning beam and transmits interrogating queries, consisting of pulse trains modulated on a carrier at 1030 MHz. Upon receiving a query, an aircraft responds by transmitting an SSR reply signal, a bursty pulse train modulating a carrier at 1090 MHz and containing the requested information. The system was designed in the 1950s, but is currently limited by the fact that all replies nominally use the same carrier frequency, and may overlap in time. This occurs, e.g., if two aircraft are close to each other but at different heights, or when an aircraft responds to a query from a neighboring ground station. If two responses overlap, the receiver cannot decode the message and both are lost [2].

Manuscript received March 29, 2006; revised November 21, 2006. The associate editor coordinating the review of this manuscript and approving it for publications was Dr. Sven Nordebo.

N. Petrochilos was with the Delft University of Technology, 2628 CD Delft, The Netherlands and with the University of Tor Vergata, Roma 00133, Italy. He is now with the University of Reims, 51687 Reims Cedex 2, France, on sabbatical leave at the Department of Electrical Engineering, University of Hawaii at Mānoa, Honolulu, HI 96822 USA (e-mail: petro@ieee.org).

A.-J. van der Veen is with the Department of Electrical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: a.j.vanderveen@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.894248

To bring some relief, a new protocol (mode S, for selective), operating at the same frequencies, is currently being installed [3]. In mode S, aircraft can be individually addressed to give a response, enabling short data communications between the station and the aircraft. This new mode will also assist the Traffic Advisory and Collision Avoidance System (TCAS) by providing automated communication between the aircraft. Nonetheless, also in this protocol, overlaps may occur.

Today, the ground station uses the same rotating antenna for transmission and reception of SSR signals. We may envision the following two extensions: 1) create a distributed system, where the existing radar system is used for transmission of the queries only, but where a network of receivers is placed at various locations and 2) equip each receiver with an array of antennas [4], [5]. This enables multilateral location estimation and facilitates message detection [4], [6], [7]. Indeed, at each receiver base station, the overlapping reply signals can be separated using blind beamforming, and subsequently, for each recovered signal, we can detect the individual symbols and estimate the direction of arrival (DOA) and the time of arrival (TOA). The combined information from several receivers allows multilateral location estimation at the system level. Estimating the beamformers and the parameters at the receiver is the aim of this paper.

Blind source separation can be done based on properties of the array response matrix or on properties of the source signals. The former has as disadvantage that a carefully calibrated array must be used, and that no multipath is tolerated. Therefore, we consider the rich structure of the source signals as a feature for separation. Indeed, SSR replies have a very structured source model: Each sample at the receiver is the product of a binary pulse-amplitude modulation (PAM) symbol taken from the alphabet $\{0, 1\}$, multiplied by a complex exponential (phase progression) due to the residual carrier frequency.

Blind source separation of SSR reply signals was first considered by Comon in [8]. This algorithm considered maximizing a contrast function based on higher order statistics (HOS). However, it was noted by Petrochilos and Comon in [9] and [10] that such HOS methods are unreliable because, for SSR mode S signals, all cumulants of order three, four, and five have a large probability to be zero. Thus, algorithms using cumulants of sixth order or higher need to be used.

In this paper, we present a collection of algorithms which implicitly use sixth- and eighth-order statistics. The proposed algorithms are *algebraic*: Similar to the analytic constant modulus algorithm (ACMA) [11], the beamformers are computed from a batch of data by solving a joint diagonalization problem. Some of these algorithms were first presented at conferences [12], [13]. Here, we present the algorithms in a broader perspec-

tive, and compare them in simulations and using experimental data. Full details can be found in [5]. Follow-up work includes [14], where an alternative time-domain approach has been proposed for the restrictive case of two partly overlapping mode S replies with a sufficiently large time difference of arrival. This paper concentrates on the complementary case where signals are highly overlapping.

This paper is structured as follows. Section II introduces the data model and lists the assumptions and some preliminary material. Section III presents three properties of the source signals, and Section IV uses these to derive three algorithms to find separating beamformers for each of the sources. Section V derives the relevant Cramér–Rao bound (CRB), and Section VI compares the algorithms in simulations. In the course of this work, we have constructed an experimental platform consisting of an array of four antennas. We apply the algorithms to measurement sets collected with the array, and show the results in Section VII.

Notation: $\mathbf{I}$ denotes the identity matrix, and $\mathbf{0}$ and $\mathbf{1}$ are the vectors with all entries being equal to 0 and 1, respectively. We denote by $(\cdot)^*$ the complex conjugation, by $(\cdot)^T$ the matrix transpose, and by $(\cdot)^H$ the matrix conjugate transpose.

$E\{\cdot\}$ denotes the mathematical expectation operator, and Vec is the operator that stacks the columns of a matrix $\mathbf{A}$ into a single vector $\mathbf{a}$. The notation $(\cdot)^\dagger$ refers to the Moore–Penrose inverse (pseudoinverse). The symbol $\odot$ denotes the Schur–Hadamard (elementwise) matrix product, and $\otimes$ the Kronecker product.

II. Data Model and Preliminaries

The SSR communicates via two different protocols: mode A/C and mode S. The A/C mode was initiated during World War II, and is supposed to be soon totally replaced by mode S. In this paper, we make the assumption that *only* mode S replies are present. A combined model containing both mode A/C and mode S replies is more complex, due to the slightly different (incommensurate) pulse lengths.

A. Received Data Model

A mode S reply frame contains either 56 or 112 binary symbols b_n. The bits are encoded in a "Manchester encoding" scheme, where $b_n = 0$ is coded as $\mathbf{b}_n = [0,1]$ and $b_n = 1$ as $\mathbf{b}_n = [1,0]$. The transmitted bit stream $\mathbf{b}$ is a burst (packet) consisting of a preamble $\mathbf{p} = [1,0,1,0,0,0,0,1,0,1,0,0,0,0,0,0]$ followed by the encoded data bits, i.e.,

$$\mathbf{b} = [\mathbf{p}, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{56/112}] \tag{1}$$

with a total length $N \in \{128, 240\}$. The preamble is aimed at facilitating the detection of the start of a packet.

The mode S reply signal emitted by the aircraft transponder is a PAM of $\mathbf{b}$, and has the form

$$b(t) = \sum_{n=0}^{127/239} \mathbf{b}[n] p(t - nT_s) \tag{2}$$

where $\mathbf{b}[n]$ is the nth entry of $\mathbf{b}$, and $p(t)$ is a (nominally) rectangular pulse of width $T_s = 0.5\mu$s, for mode S.

Before being emitted by the antenna, the signal is upconverted to have a center frequency f_e. Nominally, this carrier frequency is $f_c = 1090$ MHz, but the tolerance permitted by the International Civil Aviation Organization (ICAO) is $\pm$ 3 MHz, thus $f_e \neq f_c$. (In future, this tolerance will be reduced to $\pm$ 1 MHz.) Due to this possible carrier frequency mismatch, a residual frequency $f = f_e - f_c$ will remain after downconversion by f_c to baseband. This residual carrier adds a significant phase rotation $\exp(j2\pi ft)$ to the transmitted symbols.

At the receiver, each antenna signal is downconverted and sampled at rate $1/T_s$. Not assuming temporal synchronization within a symbol period, the received baseband signal $s[n] = s(nT_s)$ is described (up to a complex gain factor) as

$$s[n] = b[n] \exp(j2\pi n f T_s) = b[n]\phi^n \tag{3}$$

where $\phi = \exp(j2\pi f T_s)$ is the phase shift due to the residual carrier frequency over a sampling period. This signal is actually multiplied by an unknown complex gain factor representing the effect of the channel, the receiver amplifier, and an initial phase offset. This will be taken into account in Section II-B.

B. Problem Statement

We extend the single source model (3) to the reception of a mixture of d independent SSR source signals impinging on an m-element antenna array. The baseband antenna signals are sampled at rate $1/T_s$ and stacked in vectors $\mathbf{x}[n]$ (size m, where the ith entry corresponds to the ith antenna signal). After collecting N samples, the observation model is

$$\mathbf{X} = \mathbf{M}\mathbf{S} + \mathbf{E} \tag{4}$$

where $\mathbf{X} = [\mathbf{x}[1], \ldots, \mathbf{x}[N]]$ is the $m \times N$ received signal matrix, $\mathbf{M}$ is an unknown $m \times d$ mixing matrix which includes the antenna responses, path coefficients, initial phase offsets, the array signatures, and the complex gains of the sources, $\mathbf{S} = [\mathbf{s}[1], \ldots, \mathbf{s}[N]]$ is the $d \times N$ source matrix, where $\mathbf{s}[n] = [s_1[n], \ldots, s_d[n]]^T$ is a stacking of the d source signals, and $\mathbf{E}$ is the $m \times N$ noise matrix. We assume that the sources have unit amplitude and that the noise is temporally and spatially white.

Note that each source i transmits data in packets of finite length ℓ_i ($\ell_i = 128$ or 240 bits) and with an arbitrary starting time. Each source can thus have an arbitrary time offset τ_i (positive or negative integer) with respect to the start of the observation interval. To consider this effect, we apply zero-padding and truncation operations to the data time series $\mathbf{b}[n]$ in (1) to obtain a similar sequence $\tilde{\mathbf{b}}[n], 1 \leq n \leq N$, where

$$\tilde{\mathbf{b}}[n] = \begin{cases} 0, & n - \tau < 1 \quad \text{or} \quad n - \tau > \ell_i \\ \mathbf{b}[n-\tau], & 1 \leq n - \tau \leq \ell_i \end{cases}.$$

The source matrix $\mathbf{S}$ has structure

$$\mathbf{S} = \mathbf{F} \odot \mathbf{B} \tag{5}$$

where $\odot$ is the Schur–Hadamard (pointwise multiplication) operator, and

$$\mathbf{F} \stackrel{\text{def}}{=} \begin{bmatrix} \phi_1 & \phi_1^2 & \ldots & \phi_1^T \\ \vdots & \vdots & & \vdots \\ \phi_d & \phi_d^2 & \ldots & \phi_d^T \end{bmatrix}$$

$$\mathbf{B} \stackrel{\text{def}}{=} \begin{bmatrix} \tilde{\mathbf{b}}_1[1] & \tilde{\mathbf{b}}_1[2] & \dots & \tilde{\mathbf{b}}_1[N] \\ \vdots & \vdots & & \vdots \\ \tilde{\mathbf{b}}_d[1] & \tilde{\mathbf{b}}_d[2] & \dots & \tilde{\mathbf{b}}_d[N] \end{bmatrix}.$$

Although the algorithms which we will propose will not be based on a calibrated array (i.e., a parametric structure for $\mathbf{M}$), it is sometimes interesting to make the following assumptions: 1) it allows comparison to source separation algorithms which are based on direction finding, such as MUSIC and ESPRIT and 2) in some cases, the DOAs are also of interest, e.g., for localizing the aircraft (although this is the task of the primary radar), or for assigning each retrieved message to the corresponding aircraft (since the blind source separation algorithms retrieve the messages in an arbitrary ordering). Thus, if the array is a calibrated uniform linear array (ULA) with half-wavelength antenna spacing, and if the multipath is negligible, we can further write $\mathbf{M} = \mathbf{AG}$ where $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_d)]$ is the $m \times d$ steering matrix, θ_i is the direction of incidence of the ith source with respect to the ULA boresight, $\mathbf{a}(\theta)$ is the array steering vector defined as

$$\mathbf{a}(\theta) = [1, \exp(j\pi\cos(\theta)), \dots, \exp(j\pi(m-1)\cos(\theta))]^T \quad (6)$$

and $\mathbf{G}$ is a $d \times d$ diagonal matrix containing the angle-dependent antenna response, propagation gain, and initial phase offset of each source.

Without considering this structure, $\mathbf{M}$ may also reflect the imperfections of the array such as calibration errors, antenna coupling effects, or inaccuracies in the position of the elements, and propagation effects such as short-delay multipath (scattering in the vicinity of the receiver array). As mentioned, for the purpose of source separation, we do not consider this structure and only assume the matrix $\mathbf{M}$ to be left-invertible (this implies $m \geq d$).

Our aim is to compute beamformers $\mathbf{w}_i, i = 1, \dots, d$, such that $\mathbf{w}_i^H \mathbf{x}[n] = \hat{s}_i[n]$ is an estimate of the ith SSR signal. In this blind source separation context, we can only try to ensure that each $\hat{s}_i[n]$ looks like an SSR signal (i.e., that certain properties are satisfied), and that the collection of signal estimates is independent.

C. Preprocessing

In our application, $\mathbf{M}$ is typically tall and full column rank, but not square. To simplify our algorithms, we assume that first a (standard) preprocessing is applied on $\mathbf{X}$ to reduce its row dimension from m to d. This is done by computing a singular value decomposition (SVD) of $\mathbf{X}$

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary and $\mathbf{\Sigma}$ is diagonal containing the singular values in decreasing order. The number of signals d is detected from $\mathbf{\Sigma}$ using standard rank detection tests, e.g., based on likelihood ratios [15] or information theoretic criteria such as Akaike information criterion (AIC) and minimum description length (MDL) [16]–[18]. Let $\hat{\mathbf{\Sigma}}$ be a diagonal matrix containing the d largest singular values, and $\hat{\mathbf{U}}$ be an $m \times d$-dimensional matrix containing the corresponding columns of $\mathbf{U}$, and define

$$\mathbf{X}' \stackrel{\text{def}}{=} \hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{U}}^H\mathbf{X}.$$

Then, according to the model

$$\mathbf{X}' = (\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{U}}^H\mathbf{M})\mathbf{S} + (\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{U}}^H\mathbf{E}) \stackrel{\text{def}}{=} \mathbf{M}'\mathbf{S} + \mathbf{E}'.$$

This is the same model as we had before, except that $\mathbf{X}'$ is $d \times N$ and $\mathbf{M}'$ is $d \times d$ and invertible. In the algorithms, we assume that this preprocessing has been done, and we drop the primes from the notation. The computational complexity (number of multiplications) is of order $m^2 N$.

The reduction in the number of rows in $\mathbf{X}$ and $\mathbf{M}$ is necessary to avoid the existence of nullspace beamformers $\mathbf{w}_0$ such that $\mathbf{w}_0^H\mathbf{M} = 0$. Indeed, such beamformers could be added to a valid separating beamformer $\mathbf{w}_i$ without changing the output signal, and only changing the output noise. Hence, they would destroy the uniqueness of the solution, and complicate the estimation algorithms.

The data covariance whitening implied by premultiplying with $\hat{\mathbf{\Sigma}}^{-1}$ is not as essential, but has been applied in similar algorithms because it causes the beamformers to converge asymptotically in N to the Wiener beamformer [19]. Wiener beamformers are attractive because they optimize the output signal-to-interference-and-noise ratio (SINR). Note that after the prewhitening step, the noise $\mathbf{E}'$ is no longer spatially white.

D. Joint Diagonalization Problem

The algorithms to be proposed in the next sections lead to joint diagonalization problems. To avoid being repetitious, that problem is presented here in a more general setting.

Let $\mathbf{A}$ be a square matrix. Its eigenvalue decomposition (if it exists, i.e., if $\mathbf{A}$ is regular) is a factorization $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues λ_i, and $\mathbf{V}$ is an invertible matrix containing the eigenvectors $\mathbf{v}_i$ as its columns. These are the solutions to the equation $\mathbf{Av} = \mathbf{v}\lambda$. Numerically, the eigenvalue problem is often replaced by the Schur decomposition, which is the factorization [20]

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{Q}^H$$

where $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper triangular. The diagonal entries of $\mathbf{R}$ are the eigenvalues. It can be obtained by introducing the QR factorization of $\mathbf{V}$, with the advantage that the Schur decomposition always exists whereas the eigenvalue decomposition does not.

Similarly, the generalized eigenvalue problem (or matrix pencil problem) for a pair of square matrices $(\mathbf{A}, \mathbf{B})$ is to find solutions to $\mathbf{Av} = \mathbf{Bv}\lambda$ or the factorization $\mathbf{AV} = \mathbf{BV}\mathbf{\Lambda}$. If $\mathbf{B}$ is invertible, this is the same as computing the eigenvalue decomposition of $\mathbf{B}^{-1}\mathbf{A}$. It is convenient to write this as a joint decomposition. By introducing a matrix $\mathbf{F}$ and $\mathbf{W} = \mathbf{V}^{-1}$ and two diagonal matrices $\mathbf{\Lambda}_A, \mathbf{\Lambda}_B$, we obtain

$$\mathbf{A} = \mathbf{F}\mathbf{\Lambda}_A\mathbf{W}^H \quad \mathbf{B} = \mathbf{F}\mathbf{\Lambda}_B\mathbf{W}^H. \quad (7)$$

The generalized eigenvalues are $\mathbf{\Lambda} = \mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_A$, but the added generality allows to handle cases where $\mathbf{A}$ and/or $\mathbf{B}$ are singular, or some diagonal entries of $\mathbf{\Lambda}_A$ and/or $\mathbf{\Lambda}_B$ are zero. The corresponding generalized Schur decomposition is

$$\mathbf{A} = \mathbf{Q}\mathbf{R}_A\mathbf{Z}^H \quad \mathbf{B} = \mathbf{Q}\mathbf{R}_B\mathbf{Z}^H \quad (8)$$

where $\mathbf{Q}$ and $\mathbf{Z}$ are unitary and $\mathbf{R}_A$ and $\mathbf{R}_B$ are upper triangular. This decomposition always exists and can be computed iteratively, e.g., using the QZ algorithm, Jacobi iterations, etc. [20]. The diagonal entries of $\mathbf{R}_A$ and $\mathbf{R}_B$ are those of $\mathbf{\Lambda}_A$ and $\mathbf{\Lambda}_B$.

The joint diagonalization problem is a further generalization of the aforementioned to more than two matrices

$$\mathbf{A}_i = \mathbf{F}\mathbf{\Lambda}_i\mathbf{W}^H, \qquad i = 1, \ldots, K.$$

For $K > 2$, the problem is overdetermined, hence solutions, in general, do not exist. However, for a set of matrices that on the basis of model assumptions is expected to admit this factorization, we can try to compute the matrix pair $(\mathbf{F}, \mathbf{W})$ that best diagonalizes the set of given matrices, usually in some least squares sense. It can be computed using generalizations of the QZ algorithm, Jacobi iterations [21], [22], [11], [23], [24], alternating least squares [25], [26], or subspace fitting techniques [27].

The problem can be further generalized by considering rectangular matrices $\mathbf{A}_i$, as in this paper. In particular, the alternating least squares algorithms are readily generalized to handle this situation. For K matrices of size $m \times n$, the computational complexity is of order $K(m^2n + mn^2)$. Alternatively, assuming $m > n$ (tall matrices) and $\mathbf{F}$ to be of full column rank n, the problem for rectangular matrices can be reduced to a joint diagonalization of square matrices as follows. Construct $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_K]$ and use an SVD to estimate the common column space, which is equal to the column space of $\mathbf{F}$. Let $\mathbf{U}$ be a matrix containing an orthogonal basis (n columns), and define $\mathbf{A}'_i = \mathbf{U}^H\mathbf{A}_i$. Then, $\mathbf{A}'_i = \mathbf{F}'\mathbf{\Lambda}_i\mathbf{W}^H$, with $\mathbf{F}' = \mathbf{U}^H\mathbf{F}$. The $\mathbf{A}'_i$ is square $(n \times n)$ and satisfies a joint diagonalization model, and the standard algorithms apply.

III. SSR Signal Properties

The model of a single SSR reply signal (3) gives rise to several algebraic properties that will be used for blind source separation in Section IV.

A. Encoding Properties

The Manchester encoding of the SSR signals gives rise to an interesting temporal correlation property which is deterministic and independent of the actual transmitted data. If we multiply a sample from the first phase of the Manchester symbol by a sample from the second phase, the result will always be equal to zero. More generally, if we are unsynchronized to the beginning of a Manchester symbol, we can multiply three consecutive T_s-spaced samples and observe that the result is always zero (see Fig. 1). A similar property holds for a single baseband signal $s[n]$ at the receiver, independent of a fractional sampling offset and of the residual carrier frequency.

Property III.1: Independent of the transmitted data, a received mode S reply signal $s[n]$ of the form (3) obeys

$$s[n-1]s[n]s[n+1] = 0 \qquad \forall n \in \mathbb{N}. \tag{9}$$

This property will be used to design a receiver algorithm to separate multiple SSR signals.

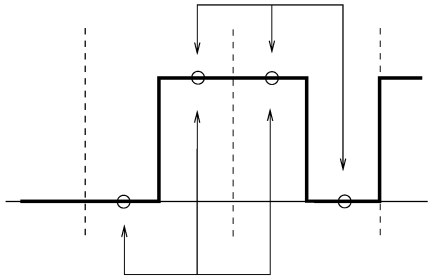

Fig. 1. Manchester encoding property: The cross product of three consecutive T_s-spaced samples is always equal to zero.

B. Zero/Constant Modulus Property

Due to the residual frequency, the received signal samples $s[n]$ are not binary as transmitted, but lie either on the unit circle, or are equal to zero. Moreover, if two subsequent received samples are nonzero, then these samples are related by a factor $\phi = \exp(2\pi j f T_s)$. These observations lead to the following two properties.

Property III.2: Static Property: $s[n]$ is a zero-constant modulus (ZCM) source if

$$s[n] = 0 \quad \text{or} \quad |s[n]| = 1 \qquad \forall n \in \mathbb{N}.$$

This is equivalent to

$$s[n]s^*[n]s[n] = s[n]. \tag{10}$$

Property III.3: Dynamic Property: For any integer k

$$s[n]s^*[n-k] = 0 \quad \text{or} \quad s[n]s^*[n-k] = \phi^k \qquad \forall n \in \mathbb{N}. \tag{11}$$

This is equivalent to

$$s[n]s^*[n-k]s[n]s^*[n-k] = \phi^k s[n]s^*[n-k] \qquad \forall n \in \mathbb{N}. \tag{12}$$

Note that the mentioned properties also hold for the rows of the source matrix $\mathbf{S}$, as defined in (5).

IV. Separation Algorithms

The properties presented in Section III are used to derive three algorithms: Algebraic Zero-Constant Modulus Algorithm (AZCMA), Manchester Decoding Algorithm (MDA), and Multishift Zero-Constant Modulus Algorithm (MS-ZCMA).[1]

A. AZCMA

The following algorithm was originally derived by Van der Veen and Tol in [12] and is included here for reference. The algorithm is derived for noise-free data (but will, of course, be applied to noisy data). We consider Property III.2 and substitute $\mathbf{w}^H\mathbf{x}[n] = s[n]$. This shows that $\mathbf{w}$ is a beamformer which returns a ZCM signal if and only if

$$\mathbf{w}^H\mathbf{x}[n]\mathbf{x}[n]^H\mathbf{w}\mathbf{w}^H\mathbf{x}[n] = \mathbf{w}^H\mathbf{x}[n] \qquad \forall n = 1, \ldots, N.$$

[1]Matlab implementations of these algorithms are available from the authors upon request.

Using properties of Kronecker products, we can separate the unknown $\mathbf{w}$ from the known $\mathbf{x}[n]$s. Note that the left-hand side contains only third-order terms of the entries of $\mathbf{w}$, whereas the right-hand side only has first-order terms. This imbalance is overcome by defining

$$\alpha = \|\mathbf{w}\|^2 = \mathbf{w}^H\mathbf{w}$$

(which is constant for each $\mathbf{w}$) and multiplying the right-hand side by $1 = \frac{1}{\alpha}\mathbf{w}^H\mathbf{w}$. This gives

$$\begin{aligned}\mathbf{w}^H\mathbf{x}[n]\mathbf{x}[n]^H\mathbf{w}\mathbf{w}^H\mathbf{x}[n] &= \frac{1}{\alpha}\mathbf{w}^H\mathbf{w}\mathbf{w}^H\mathbf{x}[n] \qquad \forall n \\ \Leftrightarrow (\mathbf{x}^*[n]\otimes\mathbf{x}[n]\otimes\mathbf{x}^*[n])^H(\mathbf{w}^*\otimes\mathbf{w}\otimes\mathbf{w}^*) &= \\ \frac{1}{\alpha}\mathrm{vec}(\mathbf{I}_d\otimes\mathbf{x}^*[n])^H(\mathbf{w}^*\otimes\mathbf{w}\otimes\mathbf{w}^*) &\qquad \forall n\end{aligned}$$

where $\otimes$ is the Kronecker product. Define matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ with rows $(\mathbf{x}^*[n]\otimes\mathbf{x}[n]\otimes\mathbf{x}^*[n])^H$ and $\mathrm{vec}(\mathbf{I}_d\otimes\mathbf{x}^*[n])^H$, respectively. Then, the ZCM separation problem is seen to be equivalent to finding all solutions $(\alpha,\mathbf{y}), \alpha\neq 0$ to

$$\alpha\mathbf{P}_1\mathbf{y} = \mathbf{P}_2\mathbf{y}, \qquad \text{where} \quad \mathbf{y} = \mathbf{w}^*\otimes\mathbf{w}\otimes\mathbf{w}^*. \tag{13}$$

To ensure an overdetermined system of equations, we require $\mathbf{P}_1$ and $\mathbf{P}_2$ to be "tall," i.e., $N \geq d^3$. Equation (13) is then a rectangular matrix pencil problem of the form $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$. The pencil is "singular," i.e., $\mathbf{P}_1$ and $\mathbf{P}_2$ are not full rank, because the structure of the rows of $\mathbf{P}_1$ implies that some of its columns are repeated. This causes additional nullspace solutions that need to be avoided. Similarly, $\mathbf{y} = \mathbf{w}^*\otimes\mathbf{w}\otimes\mathbf{w}^*$ has repeated entries, and we want our solutions to satisfy this structure. Because it is known which entries are repeated, it is straightforward to remove the duplicate entries in both $\mathbf{y}$ and the corresponding columns of $\mathbf{P}_1$ and $\mathbf{P}_2$ by defining a selection matrix $\mathbf{J}$ of size $d^3\times(1/2)d^2(d+1)$, such that

$$\mathbf{y} = \mathbf{J}\mathbf{y}' \tag{14}$$

where $\mathbf{y}'$ generically has no repeated entries. Set

$$\mathbf{P}_1' = \mathbf{P}_1\mathbf{J} \quad \mathbf{P}_2' = \mathbf{P}_2\mathbf{J} \tag{15}$$

then, generically $\mathbf{P}_1'$ has no repeated columns and has full column rank.[2] At this point, the pencil problem is replaced by

$$\alpha\mathbf{P}_1'\mathbf{y}' = \mathbf{P}_2'\mathbf{y}' \tag{16}$$

where we will assume from now on that $\mathbf{P}_1'$ has full column rank. Note that by construction $\mathbf{P}_2'$ has only d nonzero columns. Hence, there are *at most* d nonzero solutions $\{\alpha_i\}$ to (16). On the other hand, for d SSR signals, we know that there are precisely d beamformers $\mathbf{w}_i$, so that there are *at least* d nonzero solutions. It follows that the pair $(\mathbf{P}_2',\mathbf{P}_1')$ has precisely d generalized eigenvalues, necessarily equal to $\alpha_i = \|\mathbf{w}_i\|^2$; the other eigenvalues are 0. The corresponding eigenvectors $\mathbf{y}_i'$ are transformed to $\mathbf{y}_i = \mathbf{J}\mathbf{y}_i'$ to add back the repeated entries.

[2]Specific situations still lead to $\mathbf{P}_1'$ being singular, e.g., if two signals are purely constant modulus, or if two sources share exactly the same frequency. Such pencils can be analyzed using more advanced techniques (see, e.g., [28]), but for the sake of simplicity, we will not consider them here. If such a situation arises, the algorithm will fail.

At this point, there are two cases. If there are no repeated nonzero eigenvalues, then the $\mathbf{y}_i$ are (up to an arbitrary scaling) equal to $\mathbf{y}_i = \mathbf{w}_i^*\otimes\mathbf{w}_i\otimes\mathbf{w}_i^*$, from which $\mathbf{w}_i$ is immediately obtained, up to scaling. The correct scaling of $\mathbf{w}_i$ follows from the corresponding eigenvalue α_i. Alternatively, if some of the eigenvalues are repeated, then the corresponding eigenvectors form an arbitrary basis of a subspace which contains the vectors we are looking for. We need to find the correct linear combinations such that the Kronecker structure holds: This is a joint diagonalization problem as shown later.

In fact, the prewhitening step as described in Section II-C has led to a data matrix $\mathbf{X}'$ with orthonormal rows. In this case, $\alpha_i = \mathbf{w}_i^H\mathbf{w}_i = \mathbf{s}_i^H\mathbf{s}_i$, so that there are repeated eigenvalues whenever two signals have an equal number of nonzero entries, i.e., nearly always. We propose to avoid the detection of equal eigenvalues and apply the joint diagonalization step to the full collection of d eigenvectors of the pencil (16), as follows.

Each eigenvector $\mathbf{y}_i$ of size d^3 is a linear combination of the solutions, or $\mathbf{y}_i = \sum_{j=1}^d \lambda_{ij}\mathbf{w}_j^*\otimes\mathbf{w}_j\otimes\mathbf{w}_j^*$, for $i = 1,\ldots,d$. If we reshape a single vector $\lambda_{ij}(\mathbf{w}_j^*\otimes\mathbf{w}_j\otimes\mathbf{w}_j^*)$ into a $d^2\times d$ matrix, we obtain the rank-1 matrix $(\mathbf{w}_j\otimes\mathbf{w}_j^*)\lambda_{ij}\mathbf{w}_j^H$. Similarly, reshaping $\mathbf{y}_i$ into a $d^2\times d$ matrix $\mathbf{Y}_i$, we obtain

$$\mathbf{Y}_i = \mathbf{F}\mathbf{\Lambda}_i\mathbf{W}^H \tag{17}$$

where $\mathbf{W} = [\mathbf{w}_1,\ldots,\mathbf{w}_d], \mathbf{F} = [\mathbf{w}_1\otimes\mathbf{w}_1^*,\ldots,\mathbf{w}_d\otimes\mathbf{w}_d^*]$, and $\mathbf{\Lambda}_i$ is a diagonal matrix containing the coefficients λ_{ij}.

Ignoring the structure of $\mathbf{F}$, the problem to obtain $\mathbf{W}$ and the $\{\mathbf{\Lambda}_i\}$ from (17) is recognized as a joint diagonalization problem, generalized to rectangular matrices. Thus, the algorithms mentioned in Section II-D can be applied.

With noise, we follow the same algorithm. With $N\geq d^3$, the pencil problem (16) is overdetermined. The usual reduction to a pencil with square matrices $(\mathbf{P}_1'^H\mathbf{P}_2', \mathbf{P}_1'^H\mathbf{P}_1')$ amounts to a projection of the column span of $\mathbf{P}_2'$ onto that of $\mathbf{P}_1'$, and is a form of noise mitigation. The square pencil has d^3 eigenvalues, of which we keep the d largest (the others are close to 0). The corresponding eigenvectors $\{\mathbf{y}_i'\}$ are used in the joint diagonalization step. This by itself is also an overdetermined problem, hence provides additional noise mitigation.

The computational cost of the algorithm is determined by the generalized eigenvalue decomposition of the pair $(\mathbf{P}_1'^H\mathbf{P}_2', \mathbf{P}_1'^H\mathbf{P}_1')$. The algorithm is summarized in Fig. 2. The complexity is of order Nd^6 multiplications.

B. MDA

The Manchester encoding Property III.1 can be used to design a receiver algorithm to separate multiple SSR signals. Indeed, if we consider a beamformer $\mathbf{w}$ such that $\hat{s}_i[n] = \mathbf{w}^H\mathbf{x}[n]$ satisfies (9), we obtain

$$[\mathbf{x}[n+1]\otimes\mathbf{x}[n]\otimes\mathbf{x}[n-1]]^H(\mathbf{w}\otimes\mathbf{w}\otimes\mathbf{w}) = 0 \tag{18}$$

for $n = 2,\ldots,N-2$. To collect these conditions, define the matrix $\mathbf{P}: (N-2)\times d^3$ as the stack of rows $[\mathbf{x}[n+1]\otimes\mathbf{x}[n]\otimes\mathbf{x}[n-1]]^H$ for $n = 2,\ldots,N-1$, so that

$$\mathbf{P}\mathbf{w}^\diamond = 0 \quad \mathbf{w}^\diamond \stackrel{\text{def}}{=} \mathbf{w}\otimes\mathbf{w}\otimes\mathbf{w}. \tag{19}$$

Given a data matrix $\mathbf{X} = [\mathbf{x}[1], \mathbf{x}[2], \cdots, \mathbf{x}[N]]$, compute separating beamformers $\mathbf{W}$:

1) Detect number of signals d and prewhiten (Section II-C) (Nm^2)
2) Construct $\mathbf{P}'_1$ with rows $(\mathbf{x}^*[n] \otimes \mathbf{x}[n] \otimes \mathbf{x}^*[n])^H \mathbf{J}$ and $\mathbf{P}'_2$ with rows $\text{vec}(\mathbf{I}_d \otimes \mathbf{x}^*[n])^H \mathbf{J}$, $n = 1, \cdots, N$ (Nd^3)
3) Compute $\mathbf{P}'^H_1 \mathbf{P}'_2$ and $\mathbf{P}'^H_1 \mathbf{P}'_1$ (Nd^6)
4) Solve the matrix pencil $\alpha \mathbf{P}'^H_1 \mathbf{P}'_1 \mathbf{y}' = \mathbf{P}'^H_1 \mathbf{P}'_2 \mathbf{y}'$ (d^9)
   Let $\{\mathbf{y}'_i\}$ be the eigenvectors corresponding to the d largest eigenvalues
5) $\mathbf{y}_i = \mathbf{J}\mathbf{y}'_i$, $i = 1, \cdots, d$ (d^7)
   Reshape $\mathbf{y}_i$ into matrices $\mathbf{Y}_i$ (size $d^2 \times d$)
6) Solve $\mathbf{Y}_i = \mathbf{F}\mathbf{\Lambda}_i \mathbf{W}^H$ $(i = 1, \cdots, d)$ for $\mathbf{W}$ (Section II-D) $\frac{(d^6)}{Nd^6}$

Fig. 2. Summary of AZCMA (in brackets, the order of complexity of each step).

Given a data matrix $\mathbf{X} = [\mathbf{x}[1], \mathbf{x}[2], \cdots, \mathbf{x}[N]]$, compute separating beamformers $\mathbf{W}$:

1) Detect number of signals d and prewhiten (Section II-C) (Nm^2)
2) Construct $\mathbf{P}$ with rows $(\mathbf{x}[n+1] \otimes \mathbf{x}[n] \otimes \mathbf{x}[n-1])^H$, $n = 1, \cdots, N$ (Nd^3)
3) Compute the SVD of $\mathbf{P}$ (size $d^3 \times d$) and determine a basis $\{\mathbf{y}_i\}$ for the kernel (Nd^6)
4) Reshape each $\mathbf{y}_i$ into a matrix $\mathbf{Y}_i$ (size $d^2 \times d$), $i = 1, \cdots d$
5) Solve $\mathbf{Y}_i = \mathbf{F}\mathbf{\Lambda}_i \mathbf{W}^H$ $(i = 1, \cdots, d)$ for $\mathbf{W}$ (Section II-D) $\frac{(d^6)}{Nd^6}$

Fig. 3. Summary of MDA (in brackets, the order of complexity of each step).

Unlike the case in the previous algorithm, it is not necessary to reduce the dimension of $\mathbf{w}^\diamond$ as the rows of the matrix $\mathbf{P}$ do not have redundant entries. On the other hand, operations similar to (14) and (15) may improve the estimation and help force some structure in the solutions $\mathbf{y}_i$.

The rest of the algorithm is quite similar to the one in Section IV-A. For d sources, there are d linearly independent separating beamformers $\mathbf{w}_i, i = 1, \ldots, d$, as shown in Proposition IV.1. Thus, we have d linearly independent vectors $\mathbf{w}_i^\diamond$ that belong to the kernel of $\mathbf{P}$. If the kernel is d-dimensional, then the subspace spanned by $\{\mathbf{w}_i^\diamond, i = 1, \ldots, d\}$ is exactly equal to the kernel of $\mathbf{P}$, and a basis $\{\mathbf{y}_i\}$ for the kernel must be a linear combination of the $\mathbf{w}_i^\diamond$. Thus, the algorithm is to estimate an arbitrary basis $\{\mathbf{y}_i\}$ for the kernel using the SVD of $\mathbf{P}$, find linear combinations of the basis vectors to map them to the structured vectors $\{\mathbf{w}_i^\diamond\}$, and then, estimate the corresponding $\mathbf{w}_i$ from each vector $\mathbf{w}_i^\diamond$.

The key step is to find the linear combinations of the basis vectors. This is again a joint diagonalization problem in three dimensions, similar to the case discussed in Section IV-A, and can be solved for the $\mathbf{w}_i$s.

The algorithm is summarized in Fig. 3. The computational cost of the algorithm is determined by the estimation of the kernel of $\mathbf{P}$ $((N-2) \times d^3)$. The complexity is of order Nd^6.

Proposition IV.1: Assume that $\mathbf{M}$ is square and invertible, the sources are statistically stationary and temporally totally overlapping, and that there is no noise. Then, for large number of samples N, the matrix $\mathbf{P}$ will almost surely have rank $(d^3 - d)$; equivalently, its kernel will almost surely be of dimension d.

The proof is given in Appendix VIII. The proposition implies that for sufficiently large N there are no other solutions than $\{\mathbf{w}_i^\diamond, i = 1, \ldots, d\}$, so that the problem is identifiable. Experience with similar algorithms indicates that $\mathbf{P}$ is already of maximal rank once it is tall, i.e., $N \geq d^3$ is sufficient in practice [29], [30]. This is because it is very unlikely that a random square or tall matrix has a kernel unless there is a structural reason for it. The proposition showed there is no structural reason.

A limitation of the algorithm is that, for completely or almost completely nonoverlapping SSR replies, there are additional vectors in the kernel. Indeed, if $\mathbf{w}_1$ and $\mathbf{w}_2$ are corresponding beamformers, then vectors of the form $\mathbf{w}_1 \otimes \mathbf{w}_1 \otimes \mathbf{w}_2$ are in the kernel, because they correspond to conditions $s_1[n+1] \otimes s_1[n] \otimes s_2[n-1] = 0$, which is always satisfied for nonoverlapping source signals. The additional vectors in the kernel will break the assumption on which the algorithm is based (i.e., any vector in the kernel is a linear combination of the $\mathbf{w}_i^\diamond, i = 1, \ldots, d$), and without further corrections the algorithm will show poor performance in this situation.

C. MS-ZCMA

Let $\mathbf{w}$ be a d-dimensional beamforming vector to recover $s[n], \mathbf{w}^H \mathbf{x}[n] = \hat{s}[n]$. Using properties of the Kronecker product, (12) from the ZCM Property III.3 becomes

$$\begin{aligned}(\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau] \otimes \mathbf{x}[n] \\ \otimes \mathbf{x}^*[n-\tau])^T (\mathbf{w}^* \otimes \mathbf{w} \otimes \mathbf{w}^* \otimes \mathbf{w}) \\ = \phi^\tau (\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau])^T (\mathbf{w}^* \otimes \mathbf{w}). \end{aligned} \tag{20}$$

Let $\mathbf{w}^{\oslash 4}$ be an $(d(d+1)/2)^2$-dimensional vector that contains only the nonredundant elements of the Kronecker product $(\mathbf{w}^* \otimes \mathbf{w} \otimes \mathbf{w}^* \otimes \mathbf{w})$. We define by $\mathbf{J}'$ the $d^4 \times (d(d+1)/2)^2$ matrix such that $(\mathbf{w}^* \otimes \mathbf{w} \otimes \mathbf{w}^* \otimes \mathbf{w}) = \mathbf{J}' \mathbf{w}^{\oslash 4}$. We also define

$$\begin{aligned} \mathbf{w}^{\oslash 2} &= \mathbf{w}^* \otimes \mathbf{w} \\ \mathbf{p}^{(1)}_{\tau,n} &= (\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau] \otimes \mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau])^T \mathbf{J}' \\ \mathbf{p}^{(2)}_{\tau,n} &= (\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau])^T. \end{aligned}$$

Then, (20) can be written as

$$\mathbf{p}^{(1)}_{\tau,n} \mathbf{w}^{\oslash 4} = \phi^\tau \mathbf{p}^{(2)}_{\tau,n} \mathbf{w}^{\oslash 2}. \tag{21}$$

Stacking the rows $\mathbf{p}^{(1)}_{\tau,n}$ and $\mathbf{p}^{(2)}_{\tau,n}, n = \tau, \ldots, N$, into matrices $\mathbf{P}^{(1)}_\tau$ and $\mathbf{P}^{(2)}_\tau$, respectively, we obtain

$$\mathbf{P}^{(1)}_\tau \mathbf{w}^{\oslash 4} = \phi^\tau \mathbf{P}^{(2)}_\tau \mathbf{w}^{\oslash 2} \tag{22}$$

Given a data matrix $\mathbf{X} = [\mathbf{x}[1], \mathbf{x}[2], \cdots, \mathbf{x}[N]]$, compute separating beamformers $\mathbf{W}$:

1) Detect number of signals d and prewhiten (Section II-C) (Nm^2)
2) For each delay τ (L times)
   (a) Construct $\mathbf{P}_\tau^{(1)}$ with rows $(\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau] \otimes \mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau])^T \mathbf{J}'$ and $\mathbf{P}_\tau^{(2)}$ with rows $(\mathbf{x}[n] \otimes \mathbf{x}^*[n-\tau])^T$, $n = 1, \cdots, N$ (Nd^4)
   (b) Estimate $\mathbf{U}_\tau$ by subspace intersection of $\mathbf{P}_\tau^{(1)}$ and $\mathbf{P}_\tau^{(2)}$ (Nd^8)
   (c) Compute $\mathbf{B}_\tau^{(2)}$ from $\mathbf{P}_\tau^{(2)}$ and $\mathbf{U}_\tau$ (Nd^4)
3) Construct $\mathbf{Z}$ as in eqn. (27) and compute the SVD of $\mathbf{Z}$. Let $\{\mathbf{y}_i\}$ be the d eigenvectors corresponding to its kernel (Ld^6)
4) Reshape eacn $\mathbf{y}_i$ into a matrix $\mathbf{Y}_i$ (size $d \times d$) Solve $\mathbf{Y}_i = \mathbf{W}\mathbf{\Lambda}_i\mathbf{W}^H$ $(i = 1, \cdots, d)$ for $\mathbf{W}$ (Section II-D) (d^6)

Total: NLd^8

Fig. 4. Summary of MS-ZCMA (in brackets, the order of complexity of each step).

where $\mathbf{P}_\tau^{(1)}$ is $(N-\tau)\times(d(d+1)/2)^2$ and $\mathbf{P}_\tau^{(2)}$ is $(N-\tau)\times d^2$. This equation holds for all $\tau \in \mathbb{N}$ and is somewhat similar to the matrix pencil problem (13) considered before, except that the solution vectors on the left- and right-hand side of the equations are not equal.

Various approaches are possible. As in Section IV-A, we can multiply with $\alpha = \mathbf{w}^{\mathrm{H}}\mathbf{w}$ to reach an equation that involves only $\mathbf{w}^{\oslash 4}$. A second approach is proposed in [12], called AFZA, where the equation is written in terms of a single unknown vector $[\mathbf{w}^{\oslash 4T}\ \mathbf{w}^{\oslash 2T}]^T$. A third approach which seems to work better is outlined as follows.

To solve (22), we first need to find the common column span of $\mathbf{P}_\tau^{(1)}$ and $\mathbf{P}_\tau^{(2)}$. Let $\mathbf{U}_\tau$ be a matrix whose columns form an orthonormal basis of this subspace, and let $\mathbf{U}_\tau^{(1)}$ and $\mathbf{U}_\tau^{(2)}$ be the orthogonal complements of $\mathbf{U}_\tau$ over $\mathbf{P}_\tau^{(1)}$ and $\mathbf{P}_\tau^{(2)}$. Then, we can compute the decomposition into "common" and "not common" subspaces as

$$\mathbf{P}_\tau^{(1)} = \begin{bmatrix}\mathbf{U}_\tau & \mathbf{U}_\tau^{(1)}\end{bmatrix}\begin{bmatrix}\mathbf{A}_\tau^{(1)} \\ \mathbf{B}_\tau^{(1)}\end{bmatrix}$$

$$\mathbf{P}_\tau^{(2)} = \begin{bmatrix}\mathbf{U}_\tau & \mathbf{U}_\tau^{(2)}\end{bmatrix}\begin{bmatrix}\mathbf{A}_\tau^{(2)} \\ \mathbf{B}_\tau^{(2)}\end{bmatrix}$$

where the $\mathbf{B}_\tau^{(i)}, i \in \{1,2\}$ are of full row rank. Inserting these two equations into (22), we obtain

$$\begin{bmatrix}\mathbf{U}_\tau & \mathbf{U}_\tau^{(1)} & \mathbf{U}_\tau^{(2)}\end{bmatrix}\begin{bmatrix}\mathbf{A}_\tau^{(1)}\mathbf{w}^{\oslash 4} - \phi^\tau \mathbf{A}_\tau^{(2)}\mathbf{w}^{\oslash 2} \\ \hline \mathbf{B}_\tau^{(1)}\mathbf{w}^{\oslash 4} \\ \hline -\phi^\tau \mathbf{B}_\tau^{(2)}\mathbf{w}^{\oslash 2}\end{bmatrix} = \mathbf{0} \quad (23)$$

where the first matrix compound has full column rank by definition. Thus, we have

$$\mathbf{A}_\tau^{(1)}\mathbf{w}^{\oslash 4} - \phi^\tau \mathbf{A}_\tau^{(2)}\mathbf{w}^{\oslash 2} = \mathbf{0} \quad (24)$$

$$\mathbf{B}_\tau^{(1)}\mathbf{w}^{\oslash 4} = \mathbf{0} \quad (25)$$

$$\mathbf{B}_\tau^{(2)}\mathbf{w}^{\oslash 2} = \mathbf{0}. \quad (26)$$

Since it is complicated to work with the three equations simultaneously, we propose in our algorithm to use only (26). This equation holds for any τ, and we can obtain several similar conditions by taking a range of L different $\tau \in \mathbb{N}$. Stacking the matrices $\mathbf{B}_\tau^{(2)}$ in a single matrix $\mathbf{Z}$, we obtain

$$\mathbf{Z}\mathbf{w}^{\oslash 2} \stackrel{\text{def}}{=} \begin{bmatrix}\mathbf{B}_0^{(2)} \\ \mathbf{B}_1^{(2)} \\ \vdots \\ \mathbf{B}_L^{(2)}\end{bmatrix}\mathbf{w}^{\oslash 2} = \mathbf{0}. \quad (27)$$

For d SSR sources, there are d linearly independent beamformers $\mathbf{w}_i, i = 1, \ldots, d$, and these correspond to d independent solutions: nonzero vectors in the kernel of $\mathbf{Z}$. Note that $\mathbf{Z}$ has d^2 columns. As with the preceding algorithms, for a sufficient number of time-lags L, the matrix becomes very tall and will not have other vectors in the kernel.

Thus, the algorithm continues by estimating an arbitrary orthonormal basis $\{\mathbf{y}_i\}_{i=1}^d$ for the kernel of the matrix $\mathbf{Z}$. Similar to the preceding algorithms, each vector $\mathbf{y}_i$ of size d^2 of this basis is a linear combination of the solutions, or $\mathbf{y}_i = \sum_{j=1}^d \lambda_{ij}\mathbf{w}_j^{\oslash 2}$. Reshaping the $\mathbf{y}_i$ into $d \times d$ matrices $\mathbf{Y}_i$, we obtain $\mathbf{Y}_i = \mathbf{W}\mathbf{\Lambda}_i\mathbf{W}^H$, where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_d]$, and the $\mathbf{\Lambda}_i$ are diagonal matrices containing the coefficients λ_{ij}. This is the standard joint diagonalization problem, and can be solved for $\mathbf{W}$ as in Section II-D.

The algorithm is summarized in Fig. 4. The computational cost is dominated by the decomposition (23) for each τ. This corresponds to the cost of a QR factorization of $[\mathbf{P}_\tau^{(1)}\ \mathbf{P}_\tau^{(2)}]$, which is of the order Nd^8; therefore, the complexity is of order LNd^8.

The set of time delays $\{\tau_1, \ldots, \tau_L\}$ can be chosen arbitrarily, as long as the matrix $\mathbf{Z}$ is expected to achieve its maximal rank. To save some computational cost, we could initially take L small and let it grow until the estimate of the kernel of $\mathbf{Z}$ (updated using subspace tracking algorithms) is considered not to change anymore.

V. CRB

The CRB is a lower bound on the variance of unbiased parameter estimates. We consider here the deterministic CRB, which considers the additive noise as stochastic, whereas the model parameters are regarded as deterministic unknown parameters.[3] As usual, we consider the noise variance σ^2 as known, since it can be estimated independently of the other parameters. Further, we consider the case of known signals (matrix $\mathbf{B}$ is discrete-valued with entries $\{0,1\}$ and can be assumed as known for the small perturbations of the parameters under which the CRB is valid). In [32], one can find the CRB for a similar problem, which differs only by the absence of the matrix $\mathbf{B}$.

To obtain results which can be compared to DOA estimation algorithms, we consider a calibrated array and write $\mathbf{M} = \mathbf{A}(\boldsymbol{\theta})\mathbf{G}$, where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_d], \mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_d)]$, and $\mathbf{a}(\theta)$ is the array response vector for a signal from direction θ. The matrix $\mathbf{G}$ is diagonal, and the diagonal entries contain

[3]The bound for this problem was first presented in [31].

the complex gains $g_i = \rho_i e^{j\psi_i}$ of the sources as received by the array. The unknown parameters are collected in a vector $\boldsymbol{\lambda} = \{\theta_i, \rho_i, \psi_i, f_i; 1 \le i \le d\}$.

The CRB on the variance of each parameter is given by the corresponding diagonal entries of the inverse of the Fisher information matrix (FIM) [33].

Proposition V.1: In the case of known signals, the FIM is

$$\mathrm{FIM}_\alpha(\boldsymbol{\lambda}) = \frac{2}{\sigma^2}\begin{bmatrix} \mathbf{I}_{\theta\theta} & \mathbf{I}_{\theta f} & \mathbf{I}_{\theta\rho} & \mathbf{I}_{\theta\phi} \\ \mathbf{I}_{f\theta}^T & \mathbf{I}_{ff} & \mathbf{I}_{f\rho} & \mathbf{I}_{f\phi} \\ \mathbf{I}_{\rho\theta}^T & \mathbf{I}_{\rho f}^T & \mathbf{I}_{\rho\rho} & \mathbf{I}_{\rho\phi} \\ \mathbf{I}_{\phi\theta}^T & \mathbf{I}_{\phi f}^T & \mathbf{I}_{\phi\rho}^T & \mathbf{I}_{\phi\phi} \end{bmatrix} \tag{28}$$

where

$$\begin{aligned}
\mathbf{I}_{\theta\theta} &= \mathrm{Re}\left\{(\mathbf{G}\mathbf{S}\mathbf{S}^H\mathbf{G}^H) \odot (\mathbf{D}_A^H\mathbf{D}_A)^*\right\} \\
\mathbf{I}_{ff} &= \mathrm{Re}\left\{(\mathbf{G}(\mathbf{C}\odot\mathbf{S})(\mathbf{C}\odot\mathbf{S})^H\mathbf{G}^H) \odot \mathbf{R}_A^*\right\} \\
\mathbf{I}_{\theta f} &= \mathrm{Re}\left\{(\mathbf{G}\mathbf{S}(\mathbf{C}\odot\mathbf{S})^H\mathbf{G}^H) \odot (\mathbf{D}_A^H\mathbf{A})^*\right\} \\
\mathbf{I}_{\rho\rho} &= \mathrm{Re}\left\{(\boldsymbol{\Phi}\mathbf{S}\mathbf{S}^H\boldsymbol{\Phi}^H) \odot (\mathbf{A}^H\mathbf{A})^*\right\} \\
\mathbf{I}_{\theta\rho} &= \mathrm{Re}\left\{(\mathbf{G}\mathbf{S}\mathbf{S}^H\boldsymbol{\Phi}^H) \odot (\mathbf{D}_A^H\mathbf{A})^*\right\} \\
\mathbf{I}_{f\rho} &= \mathrm{Re}\{[(\mathbf{G}(\mathbf{C}\odot\mathbf{S}))(\boldsymbol{\Phi}\mathbf{S})^H] \odot (\mathbf{A}^H\mathbf{A})^*\} \\
\mathbf{I}_{\phi\phi} &= \mathrm{Re}\{[(\mathbf{G}\mathbf{S})(\mathbf{G}\mathbf{S})^H] \odot (\mathbf{A}^H\mathbf{A})^*\} \\
\mathbf{I}_{\theta\phi} &= \mathrm{Re}\left\{j[(\mathbf{G}\mathbf{S})(\mathbf{G}\mathbf{S})^H] \odot (\mathbf{D}_A^H\mathbf{A})^*\right\} \\
\mathbf{I}_{f\phi} &= \mathrm{Re}\{j[(\mathbf{G}(\mathbf{C}\odot\mathbf{S}))(\mathbf{G}\mathbf{S})^H] \odot (\mathbf{A}^H\mathbf{A})^*\} \\
\mathbf{I}_{\rho\phi} &= \mathrm{Re}\{j[(\boldsymbol{\Phi}\mathbf{S})(\mathbf{G}\mathbf{S})^H] \odot (\mathbf{A}^H\mathbf{A})^*\}
\end{aligned}$$

where $\mathbf{D}_A = [(\partial\mathbf{a})/(\partial\theta_1)(\theta_1), \ldots, (\partial\mathbf{a})/(\partial\theta_d)(\theta_d)]$, $\boldsymbol{\Phi}$ is a diagonal matrix containing the phases $e^{j\psi_i}$, $\mathbf{R}_A = \mathbf{A}^H\mathbf{A}$, $\mathbf{C} = j2\pi T[1,\ldots,1]^T[0,1,\ldots,N-1]$, and $\mathbf{S} = \mathbf{F} \odot \mathbf{B}$.

The proof is straightforward and presented in [31].

VI. Simulations

To demonstrate the effectiveness of the proposed algorithms (AZCMA, MDA, and MS-ZCMA), we compare them to JADE [21] (an HOS method based on fourth-order statistics), to EF-ICA [34] (am HOS method which forces the statistical independence of the outputs, based on FastICA), to AFZA (one of the other methods proposed in [12]), to ESPRIT [35], and to the CRB.[4]

For the simulations, we have considered a calibrated array of $M = 4$ elements, with an interelement distance of a half wavelength and in absence of multipath. The array setup is chosen to be able to compare to ESPRIT, and to have a simple way to modify the conditioning of the problem by reducing the angle separation between the sources. As source signals, we generated a uniform random sequence of $N = 100$ SSR samples (50 Manchester encoded symbols with two samples per symbol). This is a bit shorter than actual SSR packets, and does not include the training preamble.

Unless specified otherwise, the default simulation parameters are equal powered sources with a signal-to-noise ratio (SNR) of 30 dB per antenna element, DOAs equal to [70°, 110°] measured from array end-fire, no time offsets (completely overlapping data packets), and frequency offsets equal to $[-5\cdot 10^4, 5\cdot 10^4]$ Hz. We will subsequently vary the SNR, the DOA separation, and the time offset, respectively.

Most statistics are based on 10 000 independent Monte Carlo runs. For MS-ZCMA, we choose the following set of $L = 11$ time delays τ_ℓ, which has shown satisfactory results: $\{0, 1, -1, 2, -2, N/10, -N/10, N/4, -N/4, N/2, -N/2\}$.

In this simulation, the computational complexity (number of multiplications) of each of the algorithms is roughly 10^4 flops for AZCMA and MDA, and 10^6 flops for MS-ZCMA.

We show performance first in terms of the failure rate, where a failure is declared if we recover the same source twice, rather than two independent sources. A failure is also declared if the output SINR of any source is below 6 dB. For the cases without failure, we estimate the DOA θ_i of each source via

$$\hat{\theta}_i = \arg\max_\theta \|\mathbf{a}(\theta)^H\hat{\mathbf{a}}_i\|^2, \qquad i = 1, \ldots, d$$

where $\hat{\mathbf{a}}_i$ is the ith column of $\hat{\mathbf{A}} = \mathbf{W}^\dagger$, the pseudoinverse of the matrix $\mathbf{W}$. To judge the quality of the beamformers, we show the root-mean-squared error (RMSE) of the DOA estimates, averaged over both sources. (The two sources have similar RMSEs. Also, the behavior of the frequency estimates is quite similar and is not shown for brevity.) We also present the SINR at the output of the beamformer, which is defined as the ratio of the power of the desired source over the combined power of the interference sources plus the noise power.

In the first simulation, we vary the input SNR over the range [5, 30] dB. Fig. 5(a) presents the corresponding failure rates. We observe that JADE fails nearly always. The reason is that JADE is based on exploiting differences in the fourth-order statistics, whereas, for completely overlapping sources, the cumulants of up to order five are expected to be small or zero [9], [10]: The sources are "pseudo-Gaussian." The EF-ICA includes a nonlinear step and, therefore, implicitly uses also statistics of order above five. It could separate the sources, but the failure rate is high (roughly 0.35). An explanation is that a large number of samples (order few thousand) is needed to exploit the statistical independence. The AZCMA has a high failure rate for SNRs below 15–20 dB. The MDA and AFZA begin to fail for SNRs below 10 dB, while MS-ZCMA and ESPRIT do not fail for SNRs above 5 dB. Fig. 5(b) presents the RMSE of the DOA for the cases without failure. For SNRs above 20 dB, AZCMA and AFZA reach a floor of 0.1° in standard deviation. Note that the MDA, the MS-ZCMA, and the ESPRIT algorithms are quite near the CRB and are consistent. The erratic behavior of JADE is partly due to the very small number of nonfailed runs. In Fig. 5(c), the output SINR is shown as a function of the SNR. The performance of the algorithms are quite similar, except for JADE and EF-ICA. The output SINR is larger than the input SNR by five up to nearly 6 dB, which is consistent with the maximum gain of 6 dB that can be expected with four antennas for a single signal.

In a second simulation, we varied the angle separation between the two sources. Note that, with four antennas, the array beamwidth is about 45°. Fig. 6(a) shows again that JADE is not reliable, and also EF-ICA has a very high failure rate. For small angle separations (below 6°), AZCMA and ESPRIT are the first algorithms to break down. Only MDA, MS-ZCMA, and AZFA

[4]We also compared to SOBI [36], but it did not perform well. We do not present the results so as not to clutter the graphs.

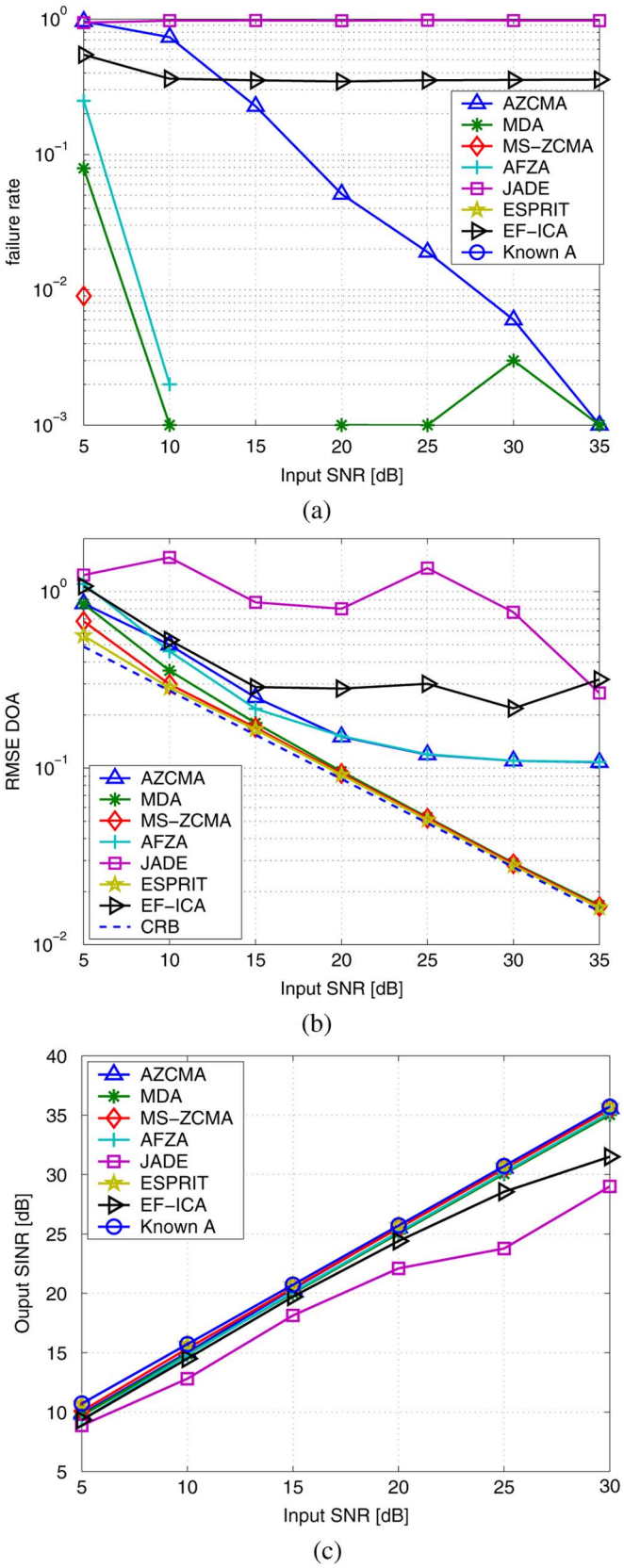

Fig. 5. Algorithm performance for varying SNR: (a) failure rate, (b) RMSE of the DOA estimates, and (c) output SINR.

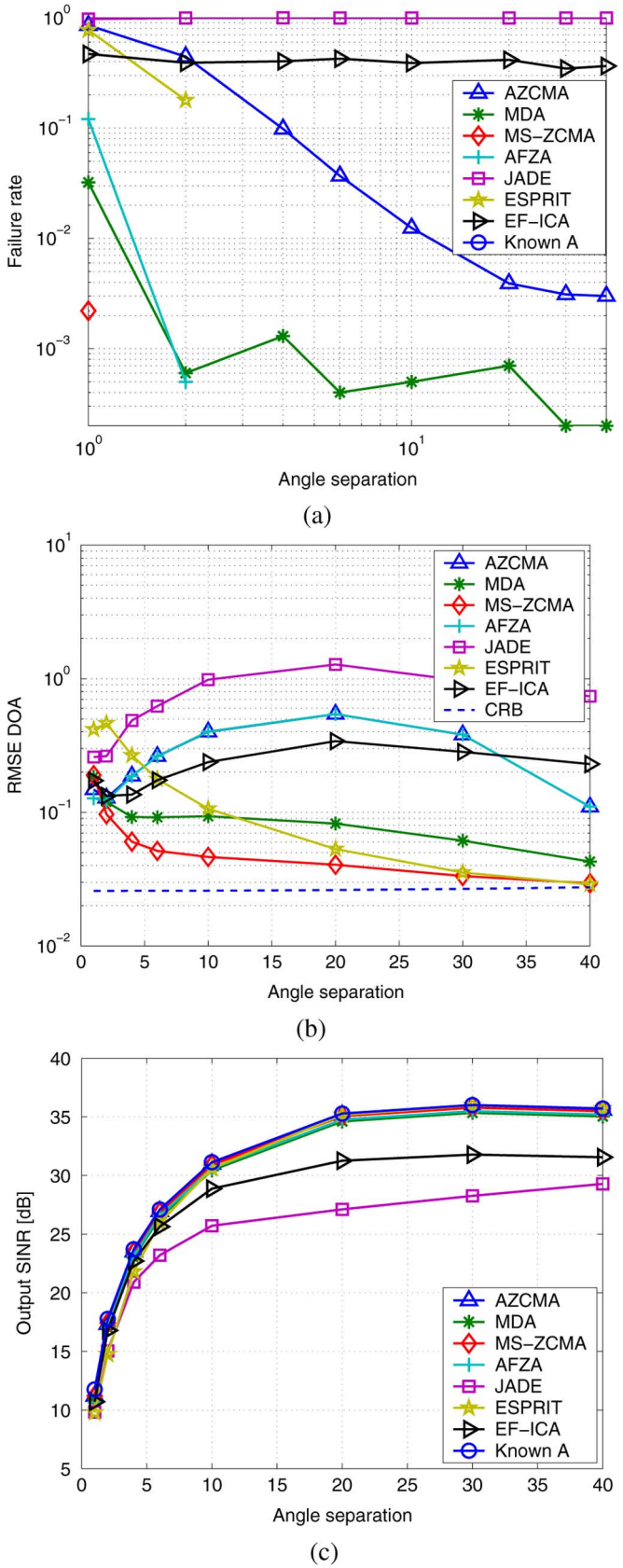

Fig. 6. Algorithm performance for varying DOA separation: (a) failure rate, (b) RMSE of the DOA estimates, and (c) output SINR.

can handle very closely spaced sources. Fig. 6(b) demonstrates that only MDA, MS-ZCMA, and ESPRIT can attain the CRB for large angle separation. For all algorithms, there is room for improvement for small angle separations. Failures are caused

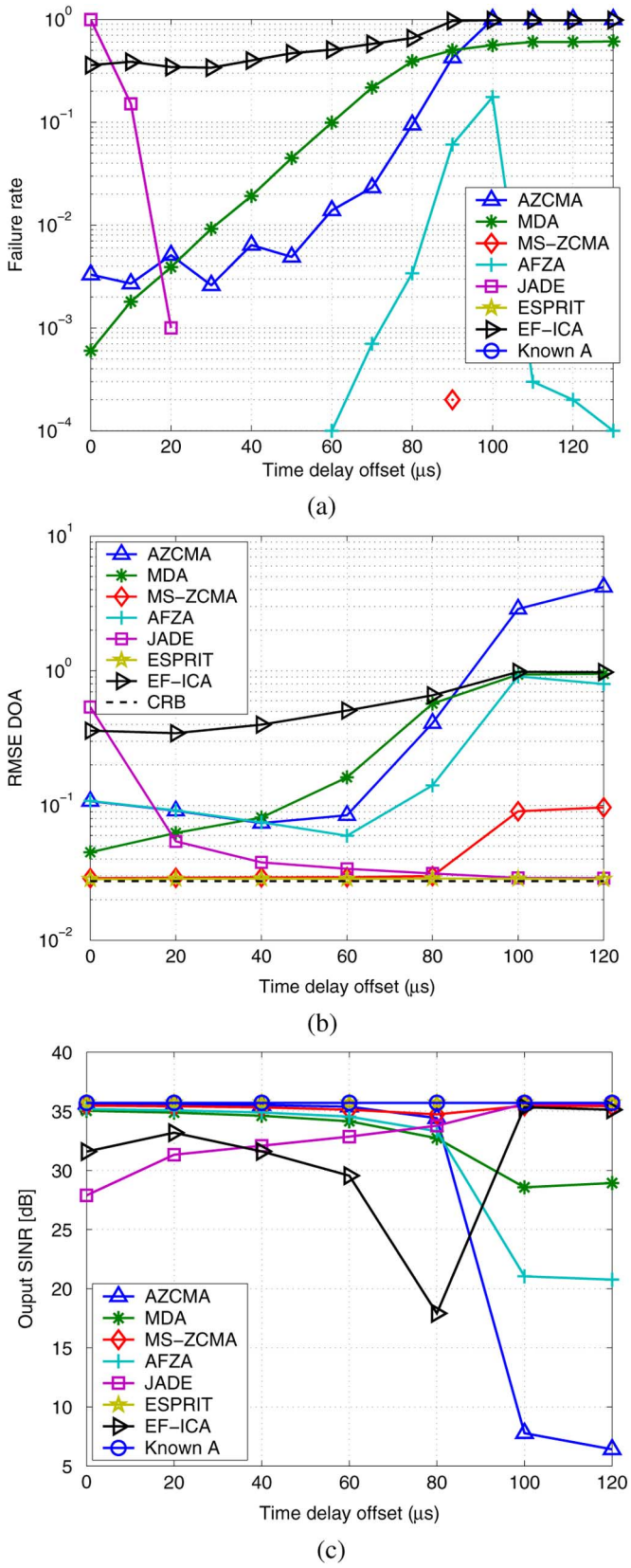

Fig. 7. Algorithm performance for varying time offset between the two data packets: (a) failure rate, (b) RMSE of the DOA estimates, and (c) output SINR.

because, in the initial SVD, a signal singular vector will be replaced by a noise singular vector, which will make it impossible to separate the two closely spaced sources. (The same happens if the number of sources is underestimated.) Note that, in the context of our application, it is rare to have small separations because, typically, the receiver antenna array would be at a different location than the requesting radar beam. In Fig. 6(c), the output SINR of all algorithms are again quite similar, except for JADE and EF-ICA. The SINR tends to a limit of 36 dB for large angle separation, which is equal to the initial input 30 dB plus 6 dB of antenna gain.

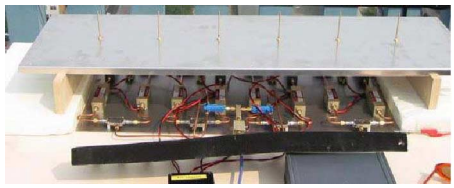

Fig. 8. Experimental system.

Finally, we simulate a varying time offset between the arrival of the two data packets. Fig. 7(a) shows that AZCMA and MDA cannot handle well cases of nonoverlapping sources, because this leads to additional nullspace solutions in the matrix pencils. On the other hand, as soon as the sources are not completely overlapping, JADE performs well, since the sources are not pseudo-Gaussian anymore. The EF-ICA performs worse with increasing time offset until it gives 100% failure in case of nonoverlapping sources. The MS-ZCMA and ESPRIT have an acceptable failure rate performance over the range of time-delay offsets. Fig. 7(b) shows that MS-ZCMA, JADE, and ESPRIT have RMSE close to the CRB, with JADE having an error increase for small time offsets and MS-ZCMA for nonoverlapping packets. The output SINR presented in Fig. 7(c) confirms that JADE performs better with increasing time-delay offset, while the effect is opposite for AZCMA, MDA, and AFZA.

In summary, the simulations show that MS-ZCMA is overall the most reliable algorithm, with the exception of cases with completely nonoverlapping packets. The AZCMA is not competitive. The MDA is a viable alternative (in view of its lower complexity) for small delay offsets only. None of the regular blind source separation algorithms is applicable.

VII. Experimental Results

A. Experimental Setup

To test the algorithms on real data, we have developed a four-channel phased array measurement system (see Fig. 8). The array consist of six linearly equispaced antennas. The two endpoint antennas are “dummy” and loaded with a matching impedance, their presence is to make the coupling between antennas uniform. The four central antennas feed different receiving chains, which downconvert the signal from the radio frequency (1090 MHz) to an intermediate frequency (10 MHz), filter around the band of interest and amplify the signal. The four signals are sampled at 50 megasamples per second and digitized by a digital oscilloscope which also stores them for later offline processing. The sampling resolution is 8 bits. Offline, the signal is converted to a complex baseband signal

using a Hilbert transform, digitally low-pass filtered with a cutoff at an equivalent frequency of 10 MHz.

The array has not been calibrated. Measurements indicate that the array response does not follow the ULA response, which may be caused by coupling or multipath, e.g., due to nearby metallic objects. For this reason, it is not possible to evaluate the DOAs, and thus, we do not compare to the ESPRIT algorithm. We limit ourselves to source separation in this section. Moreover, we will not further investigate AZFA and EF-ICA as their performances are not competitive.

B. Experimental Results

The array was installed on the roof of our building at the Delft University of Technology, Delft, The Netherlands (approximately 90 m high, with several large radar dishes and metallic structures nearby), and several data sets were recorded. Most of the recorded data sets have a short time duration of 300 μs, which is sufficient to contain a mode S signal. In these data sets, there were only a few cases of actual overlaps (we are located 50 km away from the nearest major airport, so the aircraft density is relatively low), but with very good SNR (25–40 dB). We have created artificial data sets by randomly mixing pairs of measured signals at different delay offsets, with the advantage that the true delay offsets are known and that we can easily manipulate the delay offsets, noise power, and source powers (i.e., the SNR and SINR). (The array response vectors and frequency offsets cannot be changed.) The mixing is performed simply by adding the two received blocks $\mathbf{X}_1$ and $\mathbf{X}_2$ over the four channels with an additional noise matrix

$$\mathbf{X} = \alpha_1 \mathbf{X}_1 + \alpha_2 \mathbf{X}_2 + \mathbf{E}.$$

This gave in total 700 combinations to establish the algorithm performance. Since we know the mixing and can easily estimate the "true" array response vectors from the unmixed data sets, we know the true mixing matrix $\mathbf{M}$ [see (4)] in each experiment. We use the pseudoinverse of this matrix to obtain a reference performance.

We present the results from two series of experiments, where we compare the performance of the algorithms to the pseudoinverse of the true mixing matrix $\mathbf{M}$.

In a first experiment, for every pair of sources, we add them pairwise without time-delay offset, and vary the SNR by adding complex Gaussian noise.

Fig. 9(a) shows the resulting failure rates as a function of the SNR. Comparing to the simulation (see Fig. 5), we see that AZCMA and JADE have similar performance (high failure rates). The MDA, MS-ZCMA, and "known $\mathbf{M}$" have somewhat worse but still quite similar failure rates for the measured data. Fig. 9(b) shows the output SINR as a function of the SNR. It is consistent with the simulations. The performance of all algorithms except JADE is similar to the case where the true mixing matrix is known.

In a second experiment, we use equipowered sources, add noise to achieve an SNR equal to 20 dB, and only vary the time-delay offset. Fig. 10(a) presents the failure rate as a function of the delay of the second source, and 10(b) the output SINR. Compared to the simulations (Fig. 7), AZCMA now also has a high failure rate for small separations. Again, MDA has

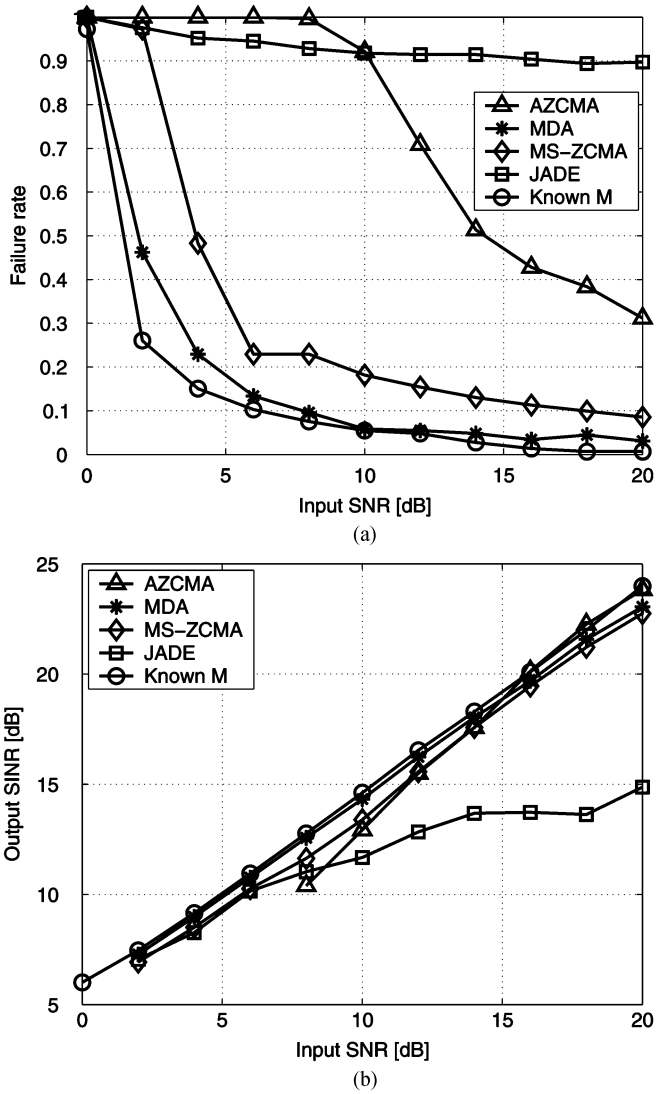

Fig. 9. Performance results using data synthesized from measured single-source data; varying input SNR: (a) failure rate and (b) output SINR.

sufficient performance only for small offsets, and is not reliable for larger offsets. The MS-ZCMA has a good performance over the complete range of offsets and is within 2 dB of the "known mixing matrix" reference line. JADE requires large time offsets for reliable performance, and in that case gets close to the reference line as well.

In summary, we have seen that the algorithms have similar behavior in simulations as in experiments. The performance of MS-ZCMA is close to the case of a known mixing matrix. The MDA is reliable only for small time offsets between the overlapping packets. Conversely, JADE is reliable only for large time offsets. The AZCMA is not competitive.

VIII. Conclusion

We presented three algorithms (AZCMA, MDA, and MS-ZCMA) to separate overlapping SSR replies impinging on an antenna array. Simulations have shown that AZCMA is not reliable, but MDA and MS-ZCMA behave reliably for data packets that are highly overlapping. For small amounts of overlap, only MS-ZCMA is reliable. The proposed algorithms

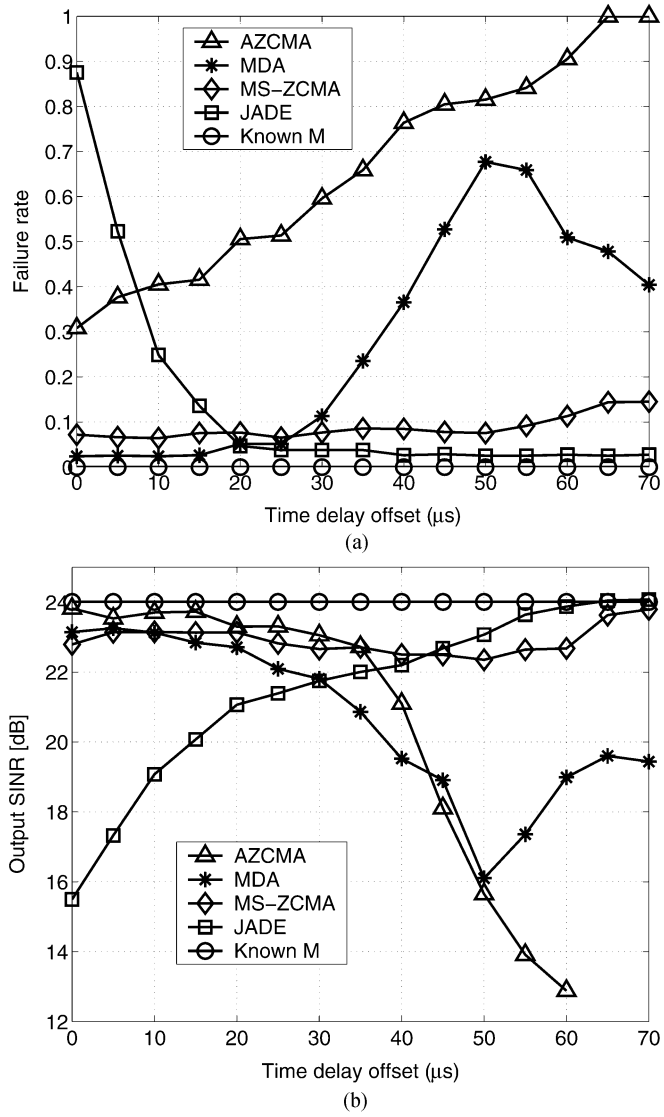

Fig. 10. Performance results using data synthesized from measured single-source data; varying time offsets between the two sources: (a) failure rate and (b) output SINR.

use properties of the sources, hence can work with uncalibrated or nonlinear arrays, which is an advantage over algorithms based on the array manifold structure, such as ESPRIT. For completely overlapping data packets, JADE is not applicable because the fourth-order cumulants are expected to be small; indeed, its performance is poor. For small amounts of overlap, JADE is functional. We developed an experimental platform, and demonstrated that MS-ZCMA performs quite well on real data as well.

In actual implementations, we recommend that, first, a test is made to detect if sources are nonoverlapping or partially overlapping. If so, simple algorithms can be used for estimating the beamformers and separating the sources [14]. If it is detected that sources are significantly overlapping, then MS-ZCMA and MDA can be applied. The complexity of these algorithms is higher, but not prohibitive since the number of simultaneous sources is likely to be small. Together, this makes a good candidate for improving the reception of the next generation of SSR.

APPENDIX
PROOF OF PROPOSITION IV.1

We need to show that, for sufficiently large N, the kernel of $\mathbf{P}$ is of dimension d, and not larger than d. We assume a stationary situation (signals are present for $N \to \infty$), so that we can apply ergodicity. Using the model $\mathbf{x}[k] = \mathbf{M}\mathbf{s}[k]$, we have

$$\mathbf{x}[k+1] \otimes \mathbf{x}[k] \otimes \mathbf{x}[k-1] = (\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M}) \times (\mathbf{s}[k+1] \otimes \mathbf{s}[k] \otimes \mathbf{s}[k-1]). \quad (29)$$

Let $\mathbf{S}_3$ be a matrix with rows $\mathbf{s}^{(3)H} \stackrel{\text{def}}{=} [\mathbf{s}[k+1] \otimes \mathbf{s}[k] \otimes \mathbf{s}[k-1]]^H$, for $k = 2, \ldots, N-1$. Then, $\mathbf{P}$ can be written as

$$\mathbf{P} = \mathbf{S}_3(\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M})^H. \quad (30)$$

Since $\mathbf{M}$ is invertible, $(\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M})$ is invertible as well, so that the rank of $\mathbf{P}$ is equal to the rank of $\mathbf{S}_3$.

Define the sample correlation matrix corresponding to $\mathbf{s}^{(3)}$ as $\hat{\mathbf{R}}_{ss}^{(3)} = (1/N)\mathbf{S}_3^H\mathbf{S}_3$. It has the same rank as $\mathbf{S}_3$ as soon as $N \geq d^3$. We will verify the rank of $\hat{\mathbf{R}}_{ss}^{(3)}$ as N tends to infinity, i.e., when $\hat{\mathbf{R}}_{ss}^{(3)}$ converges to the correlation matrix $\mathbf{R}_{ss}^{(3)}$ of $\mathbf{s}^{(3)}$, under statistical assumptions on the data.

Denote the kth sample of the mth source as $s_m[k] = b_m[k]\phi_m^k$, where $b_m[k]$ is the transmitted symbol (0 or 1 with equal probability) and ϕ_m is the residual phase rotation, random on the unit circle. The i,jth entry of $\hat{\mathbf{R}}_{ss}^{(3)}$ is

$$\left(\hat{\mathbf{R}}_{ss}^{(3)}\right)_{i,j} = \frac{1}{N}\sum_{k=2}^{N-1} b_m[k-1]b_n[k]b_l[k+1] \times b_o[k-1]b_p[k]b_q[k+1] \times \phi_m^{-(k-1)}\phi_n^{-k}\phi_l^{-(k+1)}\phi_o^{k-1}\phi_p^k\phi_q^{k+1}$$

where $i = md^2 + nd + l$ and $j = od^2 + pd + q$, for $m, n, l, o, p, q = 1, \ldots, d$. As $N \to \infty$, $\hat{\mathbf{R}}_{ss}^{(3)}$ converges to $\mathbf{R}_{ss}^{(3)}$.

Before demonstrating that the rank of $\mathbf{R}_{ss}^{(3)}$ is $d^3 - d$, we show it in the case $d = 2$. Reorder the columns and the rows in order to follow the triplets $(m,n,l) = (1,1,1), (1,1,2), (1,2,1), (2,1,1), (1,2,2), (2,1,2), (2,2,1), (2,2,2)$. Then, the matrix converges to a diagonal matrix whose diagonal entries are

$$\frac{1}{16}[0\ 1\ 2\ 1\ 1\ 2\ 1\ 0]$$

which is of rank $6 = 2^3 - 2$. There are precisely $d = 2$ columns equal to zero, which gives the kernel a dimension of two. The rest of the matrix is block-diagonal with nonsingular blocks of size 3×3.

Now, we consider the general case. For triplets that are not equal up to a permutation $\{m,n,l\} \neq \{o,p,q\}$, the residual carrier induces the term $(\hat{\mathbf{R}}_{ss}^{(3)})_{i,j}$ to tend towards zero. By reordering the rows and columns of $\mathbf{R}_{ss}^{(3)}$, we transform it into

a block-diagonal matrix, with the following three kinds of submatrices.

1) The first kind of submatrices are of size 1×1 for triplets of the form (n, n, n), for $n = 1, \ldots, d$, with value

$$\left(\hat{\mathbf{R}}_{ss}^{(3)}\right)_{i,i} = \frac{1}{N}\sum_{k=2}^{N-1} b_n^2[k-1]b_n^2[k]b_n^2[k+1]$$

where $i = n(d^2 + d + 1)$. These elements are equal to zero because of the Manchester encoding Property III.1. There are precisely d submatrices of this kind. They thus contribute d dimensions to the kernel of $\mathbf{R}_{ss}^{(3)}$.

2) The second kind of submatrices are of size 3×3 for triplets of the form (n, n, m) and $n \neq m$. Denote the corresponding submatrices by $\mathbf{R}_2$. Given that any nondiagonal entry of the matrix will contain at least three consecutive samples of b_n, Property III.1 ensures that they are equal to zero. Then, the matrix $\mathbf{R}_2$ converges to a diagonal matrix whose diagonal entries are

$$\mathbf{R}_2 = \frac{1}{16}[1\ 2\ 1]$$

which is full rank. There are precisely $d(d-1)$ submatrices of this kind.

3) The third kind of submatrices are of size 6×6 for triplets of forms $\{m, n, l\}$ and $m \neq n \neq l$. Denote by $\mathbf{R}_3$ the corresponding submatrices. A similar but more tedious analysis reveals that $\mathbf{R}_3$ converges to a certain nondiagonal but full-rank matrix. There are precisely $(d(d-1)(d-2))/(6)$ submatrices of this kind.

Hence, the rank of $\mathbf{R}_{ss}^{(3)}$ is equal to $d^3 - d$. Thus, as the number of samples tends towards infinity, $\mathbf{S}_3$ and $\mathbf{P}_3$ will have rank $(d^3 - d)$. By continuity, this will almost surely be the case for finite but sufficiently large N.

REFERENCES

[1] M. Stevens, *Secondary Surveillance Radar*. Norwood, MA: Artech House, 1988.

[2] G. Galati, S. Bartolini, and L. Menè, "Analysis of SSR signals by super resolution algorithms," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Roma, Italy, Dec. 2004, pp. 166–170.

[3] R. Trim, "Mode S: An introduction and overview," *Electron. Commun. Eng. J.*, vol. 2, pp. 53–59, Apr. 1990.

[4] N. Petrochilos and P. van Genderen, "A new approach to handle SSR replies," in *Conférence Radar*, Brest, France, May 1999, CD-ROM.

[5] N. Petrochilos, "Algorithms for separation of SSR replies" Ph.D. dissertation, Dept. Electr. Eng., Delft Univ. Technol., Delft, The Netherlands, Dec. 2002 [Online]. Available: http://cas.et.tudelft.nl/~nicolas

[6] P. Bezousek, "A passive radar surveillance system VERA for ATC," in *Proc. IRS*, Munich, Germany, 1998, CD-ROM.

[7] G. Galati, P. Magaró, M. Gasbarra, and M. Leonardi, "New signal processing techniques in SSR-Mode S replies multilateration for A-SMGCS," in *Proc. IRS*, Warsaw, Poland, May 19–21, 2004, CD-ROM.

[8] E. Chaumette, P. Comon, and D. Muller, "An ICA-based technique for radiating sources estimation; application to airport surveillance," *Proc. Inst. Electr. Eng.*, vol. 140, pt. F, pp. 395–401, Dec. 1993.

[9] N. Petrochilos and P. Comon, "Separation de signaux ZCM: Application en radar SSR," in *Proc. GRETSI*, Paris, France, Sep. 2003, CD-ROM.

[10] N. Petrochilos and P. Comon, "A zero-cumulant random variable and its applications," *IEEE Signal Process. Mag.*, 2006, accepted for publication.

[11] A. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[12] A. van der Veen and J. Tol, "Separation of zero/constant modulus signals," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Munich, Germany, Apr. 1997, pp. 3445–3448.

[13] N. Petrochilos and A. van der Veen, "Algorithms to separate overlapping secondary surveillance radar replies," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2004, pp. II.49–II.53.

[14] N. Petrochilos, G. Galati, L. Mené, and E. Piracci, "Separation of multiple secondary surveillance radar sources in a real environment by a novel projection algorithm," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Dec. 2005, pp. 125–130.

[15] D. Lawley, "Test of significance for the latent roots of covariance and correlation matrices," *Biometrika*, vol. 43, pp. 128–136, 1954.

[16] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[17] A. Liavas and P. Regalia, "On the behavior of information theoretic criteria for model order selection," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1689–1695, Aug. 2001.

[18] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.

[19] A. van der Veen, "Asymptotic properties of the algebraic constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1796–1807, Aug. 2001.

[20] G. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1989.

[21] J. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Proc. Inst. Electr. Eng. F*, vol. 140, pp. 362–370, Dec. 1993.

[22] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 17, no. 1, pp. 161–164, 1996.

[23] L. D. Lathauwer, "Signal processing based on multilinear algebra," Ph.D. dissertation, ESAT, Katholieke Universiteit Leuven, Leuven, Belgium, Sep. 1997.

[24] K. Abed-Meraim and Y. Hua, "A least-squares approach to joint Schur decomposition," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 1998, pp. 2541–2544.

[25] N. Sidiropoulos, G. Giannakis, and R. Bro, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[26] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1545–1553, Jul. 2002.

[27] A. van der Veen, "Joint diagonalization via subspace fitting techniques," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Salt Lake City, UT, May 2001, pp. 2773–2776.

[28] T. Beelen and P. V. Dooren, "An improved algorithm for the computation of Kronecker's canonical form of a singular pencil," *Lin. Alg. Appl.*, vol. 105, pp. 9–65, 1988.

[29] A. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[30] A. Leshem, N. Petrochilos, and A. van der Veen, "Finite sample identifiability of multiple constant modulus sources," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2314–2319, Sep. 2003.

[31] N. Petrochilos and P. Comon, "ML estimation of SSR signals, identifiability, and Cramér-Rao bounds," in *Proc. EUSIPCO*, Tampere, Finland, Sep. 2000, CD-ROM.

[32] A. Lemma, A. van der Veen, and E. Deprettre, "Analysis of joint angle-frequency estimation using ESPRIT," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1264–1283, May 2003.

[33] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[34] Z. Koldovsky and P. Tichavsky, "Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér-Rao lower bound," in *Proc. IEEE/Signal Process. 13th Workshop Statist. Signal Process.*, Jul. 2005, pp. 1090–1095.

[35] R. Roy and T. Kailath, "ESPRIT—estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 984–995, Jul. 1989.

[36] A. Belouchrani, K. A. Meraim, J. Cardoso, and E. Moulines, "A blind source separation technique based on second order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

Nicolas Petrochilos (S'98–M'01) was born in France in 1970. He received the B.S. degree in theoretical physics from the Ecole Normale Supérieure de Lyon (E.N.S. Lyon), Lyon, France, in 1994, the M.S. degree in signal and image processing from E.N.S.E.A., Cergy-Pontoise, France, in 1996, copromoted by the E.N.S. Lyon, with the distinctions of the jury and being ranked first in the class, the Master of Pedagogy degree in physics from the University J. Fourier, Grenoble, France, in 1997, and the Ph.D. degree in electrical engineering from the Delft University of Technology, Delft, The Netherland, and the University of Nice, Nice, France, in 2002.

Throughout 2003 until 2004, he held a Postdoctoral position at the University of Saint-Etienne, France. In fall 2004, he was a Visiting Professor at the University of Tor Vergata, Rome, Italy. Currently, he is an Assistant Professor at the University of Reims, Reims Cedex, France, on sabbatical leave from the University of Hawaii at Mānoa, Honolulu. His research interests are algebraic methods for array signal processing, and parallel factor analysis (PARAFAC) study for tensor decomposition.

Alle-Jan van der Veen (M'95–SM'03–F'05) was born in The Netherlands in 1966. He received the Ph.D. degree in electrical engineering (*cum laude*) from the Delft University of Technology, Delft, The Netherlands, in 1993.

Throughout 1994, he was a Postdoctoral Scholar at Stanford University, Stanford, CA. Currently, he is a Full Professor in Signal Processing at the Delft University of Technology. His research interests are in the general area of system theory applied to signal processing, and in particular algebraic methods for array signal processing, with applications to wireless communications and radio astronomy.

Dr. van der Veen is the recipient of the 1994 and 1997 IEEE Signal Processing Society (SPS) Young Author paper award. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1998–2001), the Chairman of the IEEE SPS Signal Processing for Communications Technical Committee (2002–2004), and the Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS (2002–2005). Currently, he is the Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and member-at-large of the Board of Governors of IEEE SPS.