Deterministic blind beamforming methods
Alle-Jan van der Veen

Abstract—Deterministic blind beamforming algorithms try to separate superpositions of source signals impinging on a phased antenna array, by using deterministic properties of the signals or the channels such as their constant modulus or directions-of-arrival. Unlike optimal or adaptive methods, the algebraic methods discussed in this review act on a fixed block of data and give closed-form expressions for beamformers by focusing on algebraic structures. This typically leads to subspace estimation and generalized eigenvalue problems.

I. INTRODUCTION

In the context of array signal processing, beamforming is concerned with the reconstruction of source signals from the outputs of an sensor array. This can be done either by coherently adding the contributions of the desired source, or by nulling out the interfering sources. The latter is an instance of the more general problem of source separation.

Classically, beamforming requires knowledge of a look direction, which is the direction of the desired source. Blind beamforming tries to copy sources without this information, relying instead on various structural properties of the problem.

The first blind beamforming techniques proposed were based on direction finding. The direction of each incoming wavefront is estimated, at the same time producing a beamformer to recover the signal from that direction. This requires at least that the antenna array is calibrated. If a source comes in via several directions (coherent multipath), then direction finding is more complicated. Depending on the situation, we also need to consider delay spread. Thus, the applicability of these techniques is much dependent on the channel conditions and in general requires a small number of well defined propagation paths per source.

More recently, new types of blind beamformers have been proposed that are not based on specific channel models, but instead exploit properties of the signals. A striking example is the constant modulus algorithm (CMA), which separates sources on the fact that their baseband representation has a constant amplitude, such as is the case for FM or phase modulated signals. A prime advantage is that these beamformers are not dependent on channel properties or array calibration. For man-made signals, such as encountered in wireless communications, signal properties are often well known and accurate, leading to robust algorithms. Several other properties are available, for example cyclostationary caused by the banded nature of digital communication signals or introduced by small differences in carrier frequencies. Ultimately, sources can be separated based on their statistical independence alone.

The paper is a summary of a review centered around algebraic techniques for deterministic blind beamforming [1]. We consider two classes of algorithms: those that are based on channel properties, and others based on signal properties. Despite the fact that these properties are widely differing, the resulting algorithms show a remarkable homogeneity. All are subspace-based techniques, and end with a generalized eigenvalue problem: the beamformers are found as the eigenvectors of a simultaneous diagonalization problem in which several matrices can be diagonalized by the same (eigenvector) matrix. The message of the paper is that joint diagonalization is the fundamental problem for source separation.

II. DATA MODELS

A. Instantaneous mixtures

Assume that $d$ source signals $s_1(t), \ldots, s_d(t)$ are transmitted from $d$ independent sources at different locations. If the delay spread is small, then we will receive a simple linear combination of these signals:

$$x(t) = a_1 s_1(t) + \cdots + a_d s_d(t)$$

where $x(t)$ is a stack of the output of the $M$ antennas. Suppose we collect a batch of $N$ samples, then

$$X = AS, \quad A = [a_1 \cdots a_d],$$

where $X = [x(0) \cdots x(N-1)]$ and $S = [s(0) \cdots s(N-1)]$. The resulting $[X = AS]$ model is called an instan-
taneous multi-input multi-output model, or I-MIMO for short. It is a generic linear model for source separation, valid when the delay spread of the dominant rays is much smaller than the inverse bandwidth of the signals.

The objective of beamforming for source separation is to construct a left-inverse $W$ of $A$, such that $WA = I$ and hence $WX = S$; see figure 1(a). This will recover the source signals from the observed mixture. It immediately follows that in this scenario it is necessary to have $d \leq M$ to ensure interference-free reception, i.e., not more sources than sensors. If we know already (part of) $S$, e.g., because of training, then $W = SX^+$, where $X^+$ denotes the Moore-Penrose pseudo-inverse of $X$. Blind beamforming is to find $W$ with knowledge only of $X$.

If each source is received from only a single direction (no multipath), then the columns of $A$ can be described by the array response vector $a(\theta)$. E.g., for a uniform linear array and a single source,

$$x(t) = \begin{bmatrix} 1 \\ \theta \\ \vdots \\ \theta^{M-1} \end{bmatrix} s(t) = a(\theta) s(t), \quad \theta = e^{j2\pi \Delta s \sin(\alpha)}$$

where $\alpha$ is the direction of the source and $\Delta$ is the spacing between the elements of the array (in wavelengths). Without multipath, the columns of $A$ lie on the array manifold $\{ a(\theta) : |\theta| = 1 \}$.

B. Convolutiove mixtures

An often-used parametric channel model that is valid for wideband sources is

$$x(t) = \sum_{i=1}^{r} a(\theta_i) \beta g(t - \tau_i) \ast s(t) = h(t) \ast s(t).$$

Here, it is assumed that the source is digital (more precisely, a dirac-pulse sequence), linearly modulated by a pulse shape function $g(t)$. The channel is supposed to be a simple multipath propagation channel, consisting of $r$ distinct paths, each parametrized by a direction $\theta_i$, a relative path delay $\tau_i$, and a complex amplitude (fading) $\beta_i$.

Suppose that the pulse shape function $g(t)$ has support (length) $L$ and that we sample at a rate $P$. We can then define the temporal signature vector

$$g(\tau) := \begin{bmatrix} g(0 - \tau) \\ g(\frac{P}{2} - \tau) \\ \vdots \\ g(L - \frac{P}{2} - \tau) \end{bmatrix}.$$ 

It is thus seen that $h(t) = \sum_i a(\theta_i) \beta_i g(t - \tau_i)$ has structure: let $g_i = g(\tau_i), a_i = a(\theta_i)$, then

$$h := \begin{bmatrix} h(0) \\ h(\frac{P}{2}) \\ \vdots \\ h(L - \frac{P}{2}) \end{bmatrix} = [g_1 \otimes a_1, \ldots, g_r \otimes a_r] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_r \end{bmatrix}.$$ 

The combined vector $g(\tau) \otimes a(\theta)$ is the space-time response vector ($\otimes$ denotes a Kronecker product.)

After collecting data samples during $N$ symbol periods, the convolutive model $x(t) = h(t) \ast s(t)$ can be written in matrix form as $X = HS_L$, where

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(N - 1) \\ x(\frac{P}{2}) & x(1 + \frac{P}{2}) & \cdots \\ \vdots & \vdots & \ddots \\ x(L - \frac{P}{2}) & \cdots & \cdots & x(L - 1) \end{bmatrix}$$

$$H = \begin{bmatrix} h(0) & \cdots & h(L - 1) \\ h(\frac{P}{2}) \\ \vdots \\ h(L - \frac{P}{2}) \end{bmatrix}$$

$$S_L = \begin{bmatrix} s_0 & s_1 & \cdots & s_{N-2} & s_{N-1} \\ s_{-1} & s_0 & \cdots & s_{N-2} & s_{N-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{-L+1} & \cdots & \cdots & s_{N-2} & s_{N-L} \end{bmatrix}.$$
III. PRINCIPLES OF BLIND BEAMFORMING

A summary of the data model developed so far is

\[
\begin{align*}
\text{I-MIMO:} & \quad X = AS, \quad A = [a(\theta_1), \ldots, a(\theta_d)] \\
\text{FIR:} & \quad X = HS, \quad h = \text{vec}(H) \\
& \quad = [g_1 \otimes a_1, \ldots, g_L \otimes a_L]B.
\end{align*}
\]

The first part of these model equations is generally valid for LTI channels, whereas the second part is a consequence of the adopted multiray model.

Based on this model, the received data matrix \( X \) has several structural properties. In several combinations, these are often strong enough to allow to find the factors \( A \) (or \( H \)) and \( S \) (or \( S \)), from knowledge of \( X \) alone. A number of properties are discussed below.

A. Toeplitz structure

The fixed baud rate of communication signals, along with time invariance, result in the fact that \( X \) has a factorization in which \( S \) is block Toeplitz. This is a strong property, and allows e.g., the blind equalization of unknown channels carrying unknown digital signals with equal baud rates [2-5]. It cannot be used for source separation, but it is useful for reducing \( X = HS \) to \( X = AS \).

B. Signal modulation structure

The signal modulation structure relates to the instantaneous amplitude of the modulated signal and includes the symbol constellation. Some typical modulation structures are listed below.

- **Constant modulus.** In many wireless applications, the transmitted waveform has a constant modulus (CM). This occurs e.g., in FM modulation, or in phase modulation as in GSM. So-called constant modulus algorithms can separate arbitrary linear superpositions of such signals, by finding out which linear combinations of the antenna outputs give back signal that have the CM property. This property is robust and can be used for blind equalization as well [6-8].

- **Finite alphabet.** Another important structure in digital communication signals is their finite alphabet (FA). The modulated signal is a linear or nonlinear map of an underlying finite alphabet, e.g., \( \{+1, -1\} \) for signals with a BPSK constellation. As with the constant-modulus property, it is possible to separate arbitrary linear combinations of FA signals in a more or less unique way [9-11].

- **Distributional properties and independence.** More in general, if the source distribution is known and not Gaussian, separation is possible by restoring the distribution functions at the output of the beamformer, e.g., using maximum-likelihood techniques. Even if the distributions are not known, we can restore distributional properties expressing the independence of sources. This is a vast area of research with many directions. Algebraic methods are possible by using higher-order stochastic moments and functions thereof, such as cumulants; see e.g., [12,13].

C. Temporal and spectral structure

The temporal structure relates to \( s(t) \) as well, but now with regard to its temporal properties. These can include knowledge of its pulse shape function and, in the case of CDMA signals, knowledge of the source codes, but also certain statistical properties for sources that are temporally non-white.

- **Temporally non-white and independence.** If the sources are independent and temporally non-white, separation is possible by using the fact that the cross-covariance and cross-cumulants of the signals at the output of the beamformer should be zero for all time lags. This allows to separate sources, but in this form cannot be used to equalize them. Often, already the second-order conditions are sufficient to find the beamformer; an example of an algebraic technique is [14].

- **Cyclostationarity.** Many signals exhibit cyclo-stationary properties, i.e., their cyclic autocorrelation function \( R_d^\alpha(\tau) = E(x(t)x(t+\tau) + j2\pi\alpha\tau) \) is wide-sense stationary and has spectral lines at selective lags \( \tau \) and frequencies \( \alpha \). This is typically caused by periodicities such as the symbol rate in banded communication signals, or residual carrier frequencies. If two sources have spectral peaks for different \((\alpha, \tau)\), then they can be separated based on this [15]. It is usually required that these parameters are known, although they can be estimated in specific cases.

For digital communication signals, a straightforward way in which the cyclostationarity property can be expressed is by oversampling the antenna outputs. The samples obtained during one symbol period presumably give independent linear combinations of the same transmitted bits, just as antennas give independent linear combinations from sampling in space. This fact was noted first in [2] and has stirred a lot of interest since; see e.g. [3-5]. Although initially called a second-order technique, the Toeplitz structure is a deterministic rather than a stochastic property.
D. Parametric properties

Parametric properties relate to the multipath model that we have derived, and extensions of this. It makes sense to use such models if the number of parameters is much smaller than, e.g., the number of coefficients in an unstructured FIR model.

The spatial manifold. In the I-MIMO model, each column of \( A \) is a linear combination of array response vectors \( \{a(\theta_i)\} \), each of which is on the array manifold. If the array manifold is known, e.g., by calibration or from structural considerations, then we can try to fit the column span of \( X \) (hence \( A \)) to the appropriate linear combinations. This will work if the number of rays is not large and if the calibration data is reliable. For this purpose, various direction finding techniques have been proposed, notably ESPRIT [16].

The temporal manifold. Similarly, \( \mathbf{h} \) is a linear combination of vectors of the form \( \{g(\tau_i) \otimes a(\theta_i)\} \), where \( g(\tau) \) is the temporal manifold function, the sampled response to an incoming pulse \( g(t-\tau) \). If the specular multipath model holds true and the number of rays is not large, the received signal is constructed from several delays of \( g(t) \), hence can be viewed as superpositions of a number of vectors \( g(\tau) \). The temporal manifold is usually known to a good accuracy. With knowledge of both the spatial and temporal manifold, we can also attempt to do a joint estimation of all angles and delays [17–19].

Residual carriers. Independent narrow band sources modulated at high frequencies rarely have exactly the same carrier frequency. Consequently, after demodulation, the co-channel sources have unequal residual carriers, with only partially overlapping spectra. If the spectral properties of the sources are known or if we sample sufficiently fast so that we can use stationarity properties of the sources, the residual carriers can be estimated and the sources can be separated, even if the array manifold is unknown. This can be regarded as a special case of cyclostationarity.

IV. Applications

A. Separation of FM radio signals

As an application of blind source separation, consider an area in which two radio towers are present. Conventionally, the two towers have to broadcast at different frequency bands, or else they will interfere with each other. Since the amount of spectrum is limited but there is a pressure to increase the number of programs and their bandwidths (c.f. HDTV), it would be very interesting to reuse the same frequency allocations within the same region. This is very well possible using blind source separation techniques, which essentially make use of the spatial dimension: the fact that the towers are at physically distinct locations. For FM-modulated signals, a suitable property for separation is their constant modulus. Typical interference suppression numbers that can be attained are shown in figure 2, where up to 6 sources are separated by a 6-element antenna array using a constant-modulus algorithm [8]. Some of the sources were separated by as little as 1.5°.

B. Separation of airplane transponder signals

Aircraft transponder signals (secondary surveillance radar (SSR) mode-S reply signals [20]) are in essence binary PAM signals with alphabet \{0,1\}. All aircraft use nominally the same carrier frequency (1090 MHz). A transponder is triggered by a pencil beam of an interrogating ground station. In today’s crowded airspace, it frequently occurs that two (or more) transponders start to broadcast simultaneously, leading to a fatal superposition of the messages when they partially overlap in frequency and time. Since no training is available, this is a good application for blind source separation techniques. This can be done based on their directions (since not much multipath is expected), or based on their modulation format: a “zero/constant modulus” signal [21].
V. EXAMPLES OF ALGORITHMS

A. Angle estimation using ESPRIT

Consider the I-MIMO model $X = AS$. Without multipath, $A = [a(\theta_1) \cdots a(\theta_d)]$. For a uniform linear array, the array manifold vector has the form

$$ a(\theta) = \begin{bmatrix} 1 \\ \theta \\ \vdots \\ \theta^{M-1} \end{bmatrix} $$

Note the shift-invariance property of this vector:

$$ a^{(1)} = \begin{bmatrix} a_1 \\ \vdots \\ a_{M-1} \end{bmatrix}, \quad a^{(2)} = \begin{bmatrix} a_2 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} \theta \\ \theta^2 \\ \vdots \\ \theta^{M-1} \end{bmatrix} $$

so that $a^{(2)} = a^{(1)} \theta$. The well-known ESPRIT algorithm to factor $X = AS$ is based on this property:

1. Find a $d$-dimensional basis $U$ for column span of $X$. Then $A = UT$, for some $d \times d$ invertible $T$.

2. Use the shift-invariance: $A^{(2)} = A^{(1)} \Theta$ \Rightarrow $U^{(2)}T^{-1} = U^{(1)}T^{-1} \Theta$, i.e.

$$ U^{(1)} T^{(2)} = T^{-1} \Theta T, \quad \Theta = \begin{bmatrix} \theta_1 & \cdots & \theta_d \end{bmatrix}. $$

This is an eigenvalue equation: $T$ contains the eigenvectors of $U^{(1)}U^{(2)}$, and $\Theta$ the eigenvalues. The beamformer for constructing $S = WX$ is $W = TU^*$, and the directions can be recovered from the eigenvalues.

B. Joint angle-delay estimation

In certain cases with low delay spread (e.g., CDMA systems), a data model $X = HS$ holds where each column of $H$ has the form $h_i = g(\tau_i) \otimes a(\theta_i)$. It is desired to estimate all delays and directions.

A Fourier transform maps delays into phase rolls:

$$ \tilde{g}_\tau = g_0 \otimes \begin{bmatrix} 1 \\ \phi \\ \vdots \\ \phi^{P-1} \end{bmatrix} = g_0 \otimes f(\phi), \quad \phi = e^{-j\frac{2\pi}{P} \tau}. $$

Since $g_0$ is known, it can be divided out, which is a form of deconvolution. After this transformation, $h_i = f(\phi_i) \otimes a(\theta_i)$. Note that this vector has a double shift-invariance structure: both in $f$ and in $a$.

An ESPRIT-like algorithm is now as follows [19]:

1. Find a basis $U$ for the column span of $X$. Then $H = UT$, for some invertible $T$.

2. Fourier transform each column of $U$, divide out $g_0$.

3. Use the double shift invariance to form two pairs of submatrices. This leads to a problem of the form

$$ \begin{cases} E_1 = T^{-1} \Phi T \\ E_2 = T^{-1} \Theta T \end{cases}. $$

4. Find $T$ as the generalized eigenvectors of this joint eigenvalue problem, e.g., by simultaneous diagonalization. $\Phi$ and $\Theta$ contain the eigenvalues, and provide estimates of the delays and directions.

Applications of this algorithm are source localization (since we recover both range and direction), and e.g., initialization of CDMA Rake receivers.

C. Separation based on constant modulus

A source sequence $[s_1, \cdots, s_N]$ represents a constant modulus signal if $|s_k|^2 = 1$, $k = 1, \cdots, N$.

Consider again the instantaneous model $X = AS$. We wish to construct all beamforming vectors $w$ such that the rows of $S$ are recovered. For a candidate $w$, it must hold that $s_k = w^* x_k$. Substitution gives

$$ w^* x_k x_k^* w = 1 \quad (k = 1, \cdots, N). $$

This is an overdetermined system of quadratic equations. A solution can be obtained as follows. Use Kronecker products to write $w^* x_k x_k^* w = [\tilde{x}_k \otimes x_k]^{*} (\tilde{w} \otimes w)$. Thus we can form a matrix $P$ with $N$ rows $[\tilde{x}_k \otimes x_k]^{*}$, and obtain the conditions

$$ P y = 1, \quad y = \tilde{w} \otimes w. $$

This is a linear system of equations, subject to a quadratic constraint. Any solution of the linear system can be written as

$$ y = \alpha_1 y_1 + \alpha_2 y_2 + \cdots + \alpha_d y_d. \quad (1) $$

It remains to find coefficients $\{\alpha_i\}$ such that $y$ satisfies $y = \tilde{w} \otimes w$. For this, we use a connection to rank-1 matrices: $\tilde{w} \otimes w \leftrightarrow \tilde{w}^* w$. In the same way, we can map all vectors in (1) to matrices, and obtain

$$ \tilde{w}^* w^* = \alpha_1 Y_1 + \alpha_2 Y_2 + \cdots + \alpha_d Y_d $$

Further massaging allows to map this problem to [8]

$$ Y_1 = T^* \Lambda_1 T \\ \vdots \\ Y_d = T^* \Lambda_d T. $$

This is again a simultaneous diagonalization problem!
The constant modulus property is quite robust in practice, and can be used without knowledge of the antenna array. An actual experiment on two mobile communication signals and a uniform linear array with 8 elements in suburban terrain indicated that the constant modulus algorithm can obtain 10 dB more interference suppression than ESPRIT [22].

VI. CONCLUSION

This paper has described algebraic methods to deterministic blind beamforming. Even within this limited framework, many properties are available and can be used to blindly separate sources and equalize channels. Column span methods are mostly parametric and try to fit a matrix channel model to the observed data. These methods are applicable if this model is valid to a reasonable accuracy, with a small number of specular rays. The requirement of a model order estimation and the sensitivity to model order mismatch can be considered their Achilles heel. On the other hand, potentially useful side information is obtained, such as delays and angles of multipath rays, which enables source localization. Uncalibrated antenna arrays can be employed if there is sufficient resolution in the delays or residual carrier frequencies.

Row span methods use properties of the signals such as a constant modulus. If these properties are present, they are very powerful and robust, and not dependent on the validity of the channel model or array calibration. The strength and at the same time limitation of deterministic row span methods is that they almost always require the signals to be man-made. More generally applicable signal separation methods are based on stochastic properties, and e.g., force the independence of the outputs of the beamformer, or reconstruct their distributions. Depending on the signal distributions, this typically requires many more samples.

REFERENCES