A Time-Frequency Technique for Blind Separation and Localization of Pure Delayed Sources

Dimitri Nion, Bart Vandewoestyne, Siegfried Vanaverbeke,
Koen Van Den Abeele, Herbert De Gersem, and Lieven De Lathauwer*

K. U. Leuven Campus Kortrijk, Group Science, Engineering and Technology,
Etienne Sabbelaan 53, 8500 Kortrijk, Belgium
{Dimitri.Nion,Bart.Vandewoestyne,Siegfried.Vanaverbeke,Koen.VanDenAbeele,
Herbert.DeGersem,Lieven.DeLathauwer}@kuleuven-kortrijk.be

Abstract. In this paper we address the problem of overdetermined blind separation and localization of several sources, given that an unknown scaled and delayed version of each source contributes to each sensor recording. The separation is performed in the time-frequency domain via an Alternating Least Squares (ALS) algorithm coupled with a Vandermonde structure enforcing strategy across the frequency mode. The latter allows to update the delays and scaling factors of each source with respect to all sensors, up to the ambiguities inherent to the mixing model. After convergence, a reference sensor can be chosen to remove these ambiguities and the Time Difference of Arrival (TDOA) estimates can be exploited to localize the sources individually.

1 Introduction

Assume that N unknown sources $s_n(t)$, $n = 1, \ldots, N$, are simultaneously and isotropically broadcasting in an anechoic propagation environment. The noise-free signal $r_m(t)$ received by the mth sensor, $m = 1, \ldots, M$, is

$$r_m(t) = \sum_{n=1}^{N} a_{mn} s_n(t - \tau_{mn}) = \sum_{n=1}^{N} (a_{mn}\delta(t - \tau_{mn})) \star s_n(t), \tag{1}$$

where $a_{mn} \in \mathbb{R}$ and $\tau_{mn} \in \mathbb{R}^+$ are the attenuation factor and the propagation delay (in seconds), respectively, between the nth source and the mth sensor, $\delta(t)$ is the Dirac impulse function and $\star$ is the linear convolution operator. The separation, time delays estimation and localization of several source signals propagating in an open-space acoustics environment is an important problem in signal

* Research supported by: (1) Research Council K.U.Leuven: GOA-Ambiorics, GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT1/08/023, (2) F.W.O.: (a) projects G.0321.06 and G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization", 2007–2011), (4) EU: ERNSI.

V. Vigneron et al. (Eds.): LVA/ICA 2010, LNCS 6365, pp. 546–554, 2010.
© Springer-Verlag Berlin Heidelberg 2010

processing, finding applications in seismics, biomedicine, sonar, radar and communications. For existing methods to handle this problem, we refer to [4,8,1,5]. In this paper, we will propose a new separation technique that belongs to the class of time-frequency algorithms, as the extended AC-DC algorithm of [8]. However, contrarily to the latter method, our approach is not limited to the case of two sources. Moreover, we do not use the second-order statistics of the observed signals. Instead, we exploit the algebraic structure of the data, by embedding a Vandermonde structure enforcing strategy within an ALS updating scheme.

Notation. The pseudo-inverse of a matrix $\mathbf{Y}$ is denoted by $\mathbf{Y}^\dagger$, its transpose by $\mathbf{Y}^T$ and its Frobenius norm by $\|\mathbf{Y}\|$. The diagonal matrix diag($\mathbf{y}$) holds the entries of $\mathbf{y}$ on its diagonal and diag($\mathbf{Y}$) is the vector consisting of the diagonal entries of the square matrix $\mathbf{Y}$. We will also use a Matlab-type notation for matrix sub-blocks, i.e., $[\mathbf{A}]_{l:m,:}$ represents the matrix built after selection of $m-l+1$ rows of $\mathbf{A}$, from the lth to the mth, and all columns of $\mathbf{A}$. The mode-2 product of a third-order tensor $\mathcal{Y} \in \mathbb{C}^{L\times M\times N}$ by a matrix $\mathbf{B} \in \mathbb{C}^{J\times M}$, denoted $\mathcal{Y} \bullet_2 \mathbf{B}$, is an $(L \times J \times N)$ tensor defined, for all index values, by $(\mathcal{Y} \bullet_2 \mathbf{B})_{ljn} = \sum_{m=1}^{M} y_{lmn} b_{jm}$.

2 Time-Frequency Formulation

Let F_s denote the sampling frequency and the $N_s \times 1$ vectors $\mathbf{r}_m$ and $\mathbf{s}_n$ denote the discrete-time versions of $r_m(t)$ and $s_n(t)$, respectively. Consider a partition of these vectors into P (possibly overlapping) frames of F samples each and compute the Discrete Fourier Transform (DFT) of each frame, to get a collection of PF time-frequency samples $\underline{\mathrm{r}}_m(p,f)$, $p = 1,\ldots,P$, $f = 1,\ldots,F$, for each sensor. From the Fourier transform shift-theorem, the time-frequency discrete version of (1) can be written as

$$\underline{\mathrm{r}}_m(p,f) \simeq \sum_{n=1}^{N} a_{mn}\omega^{(f-1)D_{mn}}\underline{\mathrm{s}}_n(p,f), \quad f = 1,\ldots,F, \tag{2}$$

where $\omega = \exp(-2j\pi/F)$ and $D_{mn} = F_s\tau_{mn}$ is the Time Of Arrival (TOA), in number of samples, between source n and sensor m. Note that the approximation (2) is exact only for periodic signals $s_n(t)$, or equivalently, if the time-convolution is circular. This approximation is satisfactory if F is significantly larger than the maximum delay [6]. To limit the circularity effect, a spectral smoothing approach is commonly used. In practice, we will compute the DFT of consecutive overlapping windowed frames (a Hanning window will be used). The time-frequency model (2) can be written as

$$\underline{\mathrm{R}}(f) = \mathbf{H}(f)\cdot\underline{\mathrm{S}}(f), \quad f = 1,\ldots,F, \tag{3}$$

where $[\underline{\mathrm{R}}(f)]_{m,p} \stackrel{\text{def}}{=} \underline{\mathrm{r}}_m(p,f)$, $[\underline{\mathrm{S}}(f)]_{n,p} \stackrel{\text{def}}{=} \underline{\mathrm{s}}_n(p,f)$, $[\mathbf{H}(f)]_{m,n} \stackrel{\text{def}}{=} a_{mn}\omega^{(f-1)D_{mn}}$. Let the third-order tensor $\mathcal{H} \in \mathbb{C}^{M\times N\times F}$ be defined as $[\mathcal{H}]_{m,n,f} \stackrel{\text{def}}{=} [\mathbf{H}(f)]_{m,n}$. For any sensor-source pair (m,n), the vector

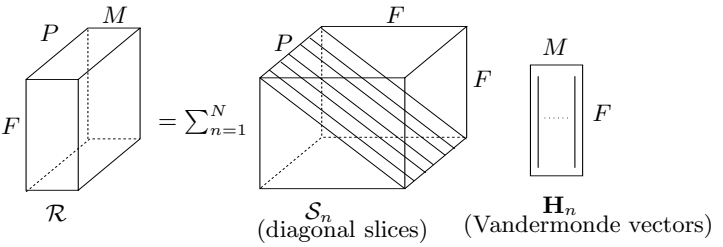


Fig. 1. Tensor view of the problem

$$\mathbf{h}_{mn} \overset{\text{def}}{=} [\mathcal{H}]_{m,n,1:F} = [a_{mn}, a_{mn}\omega^{D_{mn}}, \ldots, a_{mn}\omega^{(F-1)D_{mn}}]^T, \tag{4}$$

is a Vandermonde vector. This specific structure will be enforced on its estimate $\hat{\mathbf{h}}_{mn}$ at each step of the iterative algorithm proposed in Section 4. In the following, we will work under the following assumptions:

(A1) $P \geq N$ and $M \geq N$, i.e., we work in the overdetermined case,

(A2) $\mathbf{H}(f)$ and $\underline{\mathrm{S}}(f)$ are rank-N, for $f = 1, \ldots, F$, which is generically satisfied in practice.

3 Model Ambiguities

Let $\mathcal{R} \in \mathbb{C}^{F\times M\times P}$ and $\mathcal{S}_n \in \mathbb{C}^{F\times F\times P}$ denote the third-order tensors defined by $[\mathcal{R}]_{f,m,p} \overset{\text{def}}{=} \underline{\mathrm{r}}_m(p,f)$ and $[\mathcal{S}_n]_{:,:,p} \overset{\text{def}}{=} \mathrm{diag}([\underline{\mathrm{s}}_n(p,1), \underline{\mathrm{s}}_n(p,2), \ldots, \underline{\mathrm{s}}_n(p,F)])$, respectively. Let $\mathbf{H}_n \in \mathbb{C}^{F\times M}$ be the channel Vandermonde matrix associated to the nth source, such that $[\mathbf{H}_n]_{:,m} = \mathbf{h}_{mn}$, i.e., $\mathbf{H}_n$ is a slice of $\mathcal{H}$ obtained by fixing the source index to n. It follows that the time-frequency mixing models (2) and (3) can be written in tensor format (see Fig. 1) as

$$\mathcal{R} = \sum_{n=1}^{N} \mathcal{S}_n \bullet_2 \mathbf{H}_n^T. \tag{5}$$

Even in case of perfect separation, i.e., when the contribution $\mathcal{H}_n \overset{\text{def}}{=} \mathcal{S}_n \bullet_2 \mathbf{H}_n^T$ of each source to the observed tensor $\mathcal{R}$ is perfectly estimated, it is clear that, for any arbitrary non-singular matrices $\mathbf{Z}_n \in \mathbb{C}^{F\times F}$, $n = 1, \ldots, N$, Eq. (5) is equivalent to

$$\mathcal{R} = \sum_{n=1}^{N} (\mathcal{S}_n \bullet_2 \mathbf{Z}_n^{-1}) \bullet_2 (\mathbf{H}_n^T \cdot \mathbf{Z}_n). \tag{6}$$

However, for the tensor $(\mathcal{S}_n \bullet_2 \mathbf{Z}_n^{-1})$ to have the same structure as $\mathcal{S}_n$, i.e., for the P slices of $(\mathcal{S}_n \bullet_2 \mathbf{Z}_n^{-1})$ to be diagonal matrices, the matrix $\mathbf{Z}_n$ has to be diagonal. Moreover, for the Vandermonde structure of $\mathbf{H}_n$ to be preserved, the vector $\mathbf{u}_n \overset{\text{def}}{=} \mathrm{diag}(\mathbf{Z}_n)$ has to be a Vandermonde vector. In other words, if the respective structures of $\mathcal{S}_n$ and $\mathbf{H}_n$ are enforced on their respective estimates in

the computational strategy, the remaining ambiguity on $\hat{\mathbf{H}}_n$ in case of perfect separation is

$$\hat{\mathbf{H}}_n = \mathrm{diag}([\alpha_n, \alpha_n\omega^{\phi_n}, \ldots, \alpha_n\omega^{(F-1)\phi_n}])\mathbf{H}_n, \tag{7}$$

with unknown arbitrary scaling factor α_n and phase factor ϕ_n[1]. This shows that, for a given source n, the coefficients a_{mn} and D_{mn} w.r.t. all sensors can only be recovered up to these ambiguities:

$$\hat{\mathbf{h}}_{mn} = [\tilde{a}_{mn}, \tilde{a}_{mn}\omega^{\tilde{D}_{mn}}, \ldots, \tilde{a}_{mn}\omega^{(F-1)\tilde{D}_{mn}}], \tag{8}$$

where $\tilde{a}_{mn} \stackrel{\text{def}}{=} a_{mn}\alpha_n$ and $\tilde{D}_{mn} \stackrel{\text{def}}{=} D_{mn} + \phi_n$. Since the ambiguities $\{\alpha_n, \phi_n\}$ only depend on the source index, this suggests that they can be removed by choosing a reference sensor. Therefore, given the estimates $\tilde{a}_{mn}$ and $\tilde{D}_{mn}$, and a reference sensor, say $\underline{\mathrm{M}}$ (not necessarily the same for each source), one can compute the relative attenuation factor $\underline{\mathrm{a}}_{mn} \stackrel{\text{def}}{=} \frac{\tilde{a}_{mn}}{\tilde{a}_{\underline{\mathrm{M}}_n}} = \frac{a_{mn}}{a_{\underline{\mathrm{M}}_n}}$ and the relative Time Difference Of Arrival (TDOA) $\underline{\mathrm{D}}_{mn} \stackrel{\text{def}}{=} \tilde{D}_{mn} - \tilde{D}_{\underline{\mathrm{M}}_n} = D_{mn} - D_{\underline{\mathrm{M}}_n}$. As illustrated in Section 5, estimation of the relative TDOAs w.r.t. a reference sensor is sufficient to localize the sources.

4 ALS Algorithm with Vandermonde Structure Enforcing

Estimation of $\mathbf{H}(f)$ and $\underline{\mathrm{S}}(f)$, $f = 1, \ldots, F$, can be achieved by solving the following optimization problem

$$\min_{\{\mathbf{H}(f), \underline{\mathrm{S}}(f)\}_{f=1}^{F}} \gamma$$
$$\text{s.t. } \mathbf{h}_{mn} \text{ defined in (4) is a Vandermonde vector, } \forall m, \forall n,$$

where $\gamma \stackrel{\text{def}}{=} \sum_{f=1}^{F} \|\underline{\mathrm{R}}(f) - \mathbf{H}(f)\cdot\underline{\mathrm{S}}(f)\|^2$. In this section, we propose an algorithm that consists of three steps at each iteration. In the first step, given the previous estimates $\hat{\underline{\mathrm{S}}}(f)$, $f = 1, \ldots, F$, the matrices $\hat{\mathbf{H}}(f)$ are updated in the least squares sense:

$$\hat{\mathbf{H}}^{(LS)}(f) = \underline{\mathrm{R}}(f) \cdot \hat{\underline{\mathrm{S}}}(f)^{\dagger}, \quad f = 1, \ldots, F. \tag{9}$$

In the second step, the purpose is to enforce the Vandermonde structure on the MN vectors $\hat{\mathbf{h}}_{mn}^{(LS)} \stackrel{\text{def}}{=} [\hat{\mathcal{H}}^{(LS)}]_{m,n,1:F}$, for $m = 1, \ldots, M$, $n = 1, \ldots, N$, where $[\hat{\mathcal{H}}^{(LS)}]_{m,n,f} \stackrel{\text{def}}{=} [\hat{\mathbf{H}}^{(LS)}(f)]_{m,n}$. Several algorithms have been proposed in the literature for the latter task (see, e.g., [3] and references therein). In practice, we will use the popular periodogram-based technique proposed in [7]. This consists of the computation of the FFT of the zero-padded sequence $\hat{\mathbf{h}}_{mn}^{(LS)}$. For each sensor-source pair (m, n), $\tilde{D}_{mn}$ is then updated as the index l for which the modulus of the FFT takes its maximum value, whereas $\tilde{a}_{mn}$ is updated as the

[1] Of course, the source components are estimated in an arbitrary order since one can arbitrarily permute the N terms of the sum in (5).

Algorithm 1. ALS algorithm with Vandermonde structure enforcing.

STEP 1: Time-frequency computation
Build $\underline{\mathrm{R}}(f) \in \mathbb{C}^{M \times P}$, $f = 1, \ldots, F$ from FFT of P overlapping windowed frames of recorded signals. (Typical parameters: $F = 2048$, Hanning window, 50% overlap).

STEP 2: Blind separation
—— Initialization ————
stop=0, $k = 1$, K_{max} (e.g., $K_{max} = 200$) and ϵ (e.g., $\epsilon = 10^{-6}$). Randomly generate unitary matrices $\underline{\hat{\mathrm{S}}}(f) \in \mathbb{C}^{N \times P}$, $f = 1, \ldots, F$. Possibly try several random starting points.
—— Start alternating updates ————
while stop=0
$k = k + 1$
(2.a). $\hat{\mathbf{H}}^{(LS)}(f) = \underline{\mathrm{R}}(f) \cdot \underline{\hat{\mathrm{S}}}(f)^{\dagger}$, $f = 1, \ldots, F$.
(2.b). $\{\hat{\tilde{D}}_{mn}, \hat{\tilde{a}}_{mn}\} \leftarrow \text{periodogram}(\hat{\mathbf{h}}_{mn}^{(LS)})$, $m = 1, \ldots, M$, $n = 1, \ldots, N$.
$\hat{\mathbf{h}}_{mn}^{(VDM)} \leftarrow [\hat{\tilde{a}}_{mn}, \hat{\tilde{a}}_{mn} w^{\hat{\tilde{D}}_{mn}}, \ldots, \hat{\tilde{a}}_{mn} w^{(F-1)\hat{\tilde{D}}_{mn}}]$, $m = 1, \ldots, M$, $n = 1, \ldots, N$.
(2.c). $\underline{\hat{\mathrm{S}}}(f) = \hat{\mathbf{H}}^{(VDM)}(f)^{\dagger} \cdot \underline{\mathrm{R}}(f)$, $f = 1, \ldots, F$.
if $(k = K_{max})$ or $(|\gamma^{(k)} - \gamma^{(k-1)}| \leq \epsilon)$
stop=1;
end
end
If several starting points are used, keep the estimates associated to the smallest final value of γ.
Choose reference sensor $\underline{\mathrm{M}}$ and remove ambiguities: $\underline{\hat{\mathrm{a}}}_{mn} = \frac{\hat{\tilde{a}}_{mn}}{\hat{\tilde{a}}_{\underline{\mathrm{M}}n}}$ and $\underline{\hat{\mathrm{D}}}_{mn} = \hat{\tilde{D}}_{mn} - \hat{\tilde{D}}_{\underline{\mathrm{M}}n}$.

value taken by the real part of the FFT at index l. Each vector $\hat{\mathbf{h}}_{mn}^{(LS)}$ is then substituted by the Vandermonde vector $\hat{\mathbf{h}}_{mn}^{(VDM)}$, built from the estimate of $\tilde{D}_{mn}$ and $\tilde{a}_{mn}$ as in (8). The matrices $\hat{\mathbf{H}}^{(LS)}(f)$, $f = 1, \ldots, F$, are substituted by $\hat{\mathbf{H}}^{(VDM)}(f)$ accordingly. In the last step, the matrices $\underline{\hat{\mathrm{S}}}(f)$ are updated in the least squares sense as

$$\underline{\hat{\mathrm{S}}}(f) = \hat{\mathbf{H}}^{(VDM)}(f)^{\dagger} \cdot \underline{\mathrm{R}}(f), \quad f = 1, \ldots, F. \tag{10}$$

The scaling and phase ambiguities α_n and ϕ_n are removed after convergence, as explained in Section 3. Note that convergence of the resulting algorithm is not guaranteed to be monotonic. Although the least squares updates of $\hat{\mathbf{H}}(f)$ and $\underline{\hat{\mathrm{S}}}(f)$ can only decrease or maintain the current value of γ, this is not guaranteed for the Vandermonde structure enforcing step. However, we observed through numerical experiments that our algorithm converges monotonically in many practical situations. A summary of the proposed technique is given in Algorithm 1.

5 Source Localization

The purpose of this stage is to localize the N sources from the TDOA estimates $\underline{\hat{\mathrm{D}}}_{mn}$, in number of samples, w.r.t. the reference sensor $\underline{\mathrm{M}}$. Let $\mathbf{u}_n = [x_n, y_n]^T$ denote the unknown vector of Cartesian coordinates of the nth source in a bi-dimensional propagation medium[2] and $\tilde{\mathbf{u}}_m = [\tilde{x}_m, \tilde{y}_m]^T$ the vector of known coordinates of the mth sensor. Choose the reference sensor $\underline{\mathrm{M}}$ as the origin of

[2] For simplicity, the localization task is formulated for a 2D medium. It can easily be generalized to the 3D case.

the new system of coordinates, in which the nth source and mth sensor have coordinates $\mathbf{z}_n \stackrel{\text{def}}{=} \mathbf{u}_n - \tilde{\mathbf{u}}_{\underline{\mathrm{M}}}$ and $\tilde{\mathbf{z}}_m \stackrel{\text{def}}{=} \tilde{\mathbf{u}}_m - \tilde{\mathbf{u}}_{\underline{\mathrm{M}}}$, respectively. Let us compute the relative range difference estimates $\hat{\underline{\mathrm{d}}}_{mn} \stackrel{\text{def}}{=} \hat{\underline{\mathrm{D}}}_{mn} v/F_s$, where v denotes the wave velocity in the propagation medium. In case of perfect estimation, $\hat{\underline{\mathrm{d}}}_{mn}$ satisfies

$$\hat{\underline{\mathrm{d}}}_{mn} + \|\mathbf{z}_n\| = \|\tilde{\mathbf{z}}_m - \mathbf{z}_n\|. \tag{11}$$

Squaring both sides of (11) yields

$$\hat{\underline{\mathrm{d}}}_{mn}\|\mathbf{z}_n\| + \tilde{\mathbf{z}}_m^T \mathbf{z}_n = \frac{1}{2}\left(\|\tilde{\mathbf{z}}_m\|^2 - \hat{\underline{\mathrm{d}}}_{mn}^2\right). \tag{12}$$

Considering all sensors except the reference sensor, $m \in \{1, 2, \ldots, M\} - \{\underline{\mathrm{M}}\}$, (12) is equivalent to

$$\begin{bmatrix} \tilde{z}_1(1) & \tilde{z}_1(2) & \hat{\underline{\mathrm{d}}}_{1n} \\ \vdots & \vdots & \vdots \\ \tilde{z}_M(1) & \tilde{z}_M(2) & \hat{\underline{\mathrm{d}}}_{Mn} \end{bmatrix} \begin{bmatrix} z_n(1) \\ z_n(2) \\ \|\mathbf{z}_n\| \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \|\tilde{\mathbf{z}}_1\|^2 - \hat{\underline{\mathrm{d}}}_{1n}^2 \\ \vdots \\ \|\tilde{\mathbf{z}}_M\|^2 - \hat{\underline{\mathrm{d}}}_{Mn}^2 \end{bmatrix},$$

which is compactly written as

$$\mathbf{Z}_n \theta_n = \mathbf{p}_n, \tag{13}$$

where $\mathbf{Z}_n \in \mathbb{R}^{(M-1)\times 3}$ and $\mathbf{p}_n \in \mathbb{R}^{(M-1)\times 1}$ are known. For each source index n, (13) can be solved in the least squares sense. Assuming that $\mathbf{Z}_n$ is of rank three, we get

$$\hat{\theta}_n = \mathbf{Z}_n^{\dagger} \mathbf{p}_n, \tag{14}$$

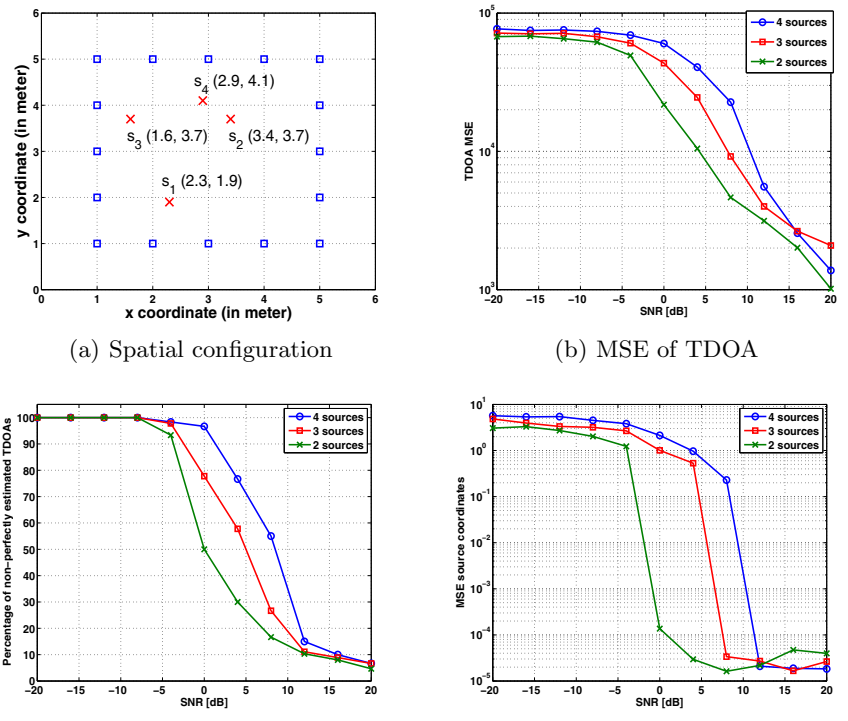
where $\|\mathbf{z}_n\|$ is treated as a variable independent from $z_n(1)$ and $z_n(2)$. A better option is to solve (13) as a constrained minimization problem

$$\begin{aligned} &\min_{\theta_n} \psi \\ &\text{s.t. } \theta_n(3) = \sqrt{\theta_n(1)^2 + \theta_n(2)^2}, \end{aligned} \tag{15}$$

where

$$\psi \stackrel{\text{def}}{=} \|\mathbf{Z}_n \theta_n - \mathbf{p}_n\|^2. \tag{16}$$

We refer to [2,9] and references therein for solutions to this problem. In practice, we will use the Quadratically Constrained Least Squares (QCLS) method of [9]. The localization procedure is repeated independently for each source. Finally, the coordinates of the nth source in the initial Cartesian system are obtained by $\hat{\mathbf{u}}_n = \hat{\mathbf{z}}_n + \tilde{\mathbf{u}}_{\underline{\mathrm{M}}}$, where $\hat{\mathbf{z}}_n = [\hat{\theta}_n(1), \hat{\theta}_n(2)]^T$. Note that the accuracy of the coordinate estimates relies on the accuracy of the TDOA estimates. In practice, it may happen that several TDOA estimates are significantly more accurate than others. This suggests that a subset of $\tilde{M} \geq 3$ reliable estimates among $M-1$ should be used in the localization process. In order to find the $\tilde{M}$ most



(a) Spatial configuration (b) MSE of TDOA

(c) % of non-perfectly estimated TDOAs (d) MSE of source coordinates

Fig. 2. Spatial configuration and results of Monte-Carlo experiments

reliable estimates, one can select $\tilde{M}$ rows of $\mathbf{Z}_n$, then solve (15) with the resulting reduced-size matrix, and repeat the procedure for all possible combinations of $\tilde{M}$ rows chosen among $M-1$. The final estimate of θ_n is the one associated to the smallest value of ψ.

6 Numerical Experiments

Let the noise-free signal $\mathbf{r}_m$ at receiver m be corrupted by Additive White Gaussian Noise (AWGN), $\tilde{\mathbf{r}}_m = \mathbf{r}_m + \sigma_m \mathbf{v}_m$, where the noise vector $\mathbf{v}_m$ is generated from a zero-mean unit-variance Gaussian distribution and σ_m is computed at each receiver to ensure a chosen Signal to Noise Ratio (SNR), $\text{SNR}_m = 10\log_{10}(\|\mathbf{r}_m\|^2/\|\sigma_m\mathbf{v}_m\|^2)$ [dB]. SNR_m is fixed here to the same value for all receivers and is further denoted SNR. We simulate the 2D propagation environment of Fig. 2(a), with $N = 4$ sources and $M = 16$ sensors. The sources consist of 10000 samples of different speech signals, with a sampling frequency $F_s = 16$ kHz. The wave speed is $v = 340$ m.s^{-1}. In this section, we illustrate the performance of our algorithm via Monte-Carlo simulations consisting of 100

independent trials for each value of the SNR. The noise vector $\mathbf{v}_m$ and the scaling factors a_{mn} for $m = 1, \ldots, M$, $n = 1, \ldots, N$, are randomly generated for each trial (a_{mn} is drawn between -3 and 3 with a uniform distribution). All experiments are conducted with the reference sensor located at $(1, 1)$. Fig. 2(b) illustrates the evolution of the Mean Square Error (MSE) ζ of the TDOA estimates, $\zeta = \frac{1}{N(M-1)} \sum_{n=1}^{N} \sum_{m=1}^{M-1} (\hat{\underline{\mathrm{D}}}_{mn} - \underline{\mathrm{D}}_{mn})^2$ for $N = \{2, 3, 4\}$. Fig. 2(c) illustrates the evolution of the percentage of non-perfectly estimated TDOAs ($\hat{\underline{\mathrm{D}}}_{mn} - \underline{\mathrm{D}}_{mn} \neq 0$), computed over the $N(M-1)$ estimates. It can be observed that, for SNR=20 dB, more than 90% of the TDOAs are perfectly estimated, even with $N = 4$ sources. Fig. 2(d) illustrates the evolution of the MSE ρ of the source coordinates, $\rho = \frac{1}{2N} \sum_{n=1}^{N} (x_n - \hat{x}_n)^2 + (y_n - \hat{y}_n)^2$, the latter being computed from the $\tilde{M} = 6$ most reliable sensors. It can be observed that, above a SNR threshold, the value of which depends on the number of sources, the localization of all sources is almost perfect. For instance, with $N = 2$, $\rho \simeq 0$ for SNR ≥ 0 dB, whereas the associated percentage of non-perfectly estimated TDOAs on Fig. 2(c) is 50% for this SNR value. This shows the benefit of searching for the most reliable TDOAs to be used in the localization process.

7 Conclusion

In this paper, we have proposed a novel time-frequency technique to deal with the problem of blind separation and localization of pure delayed sources. The core idea of the separation task is to interleave a Vandermonde structure enforcing strategy on the channel updates across the frequency mode with alternating least squares updates of the source and channel matrices. The localization task relies on the selection of the most reliable subset of TDOA estimates. Monte-Carlo experiments with two, three, and four sources have been conducted to corroborate our findings.

References

1. Chabriel, G., Barrère, J.: An Instantaneous Formulation of Mixtures for Blind Separation of Propagating Waves. IEEE Trans. Sig. Proc. 54(1), 49–58 (2006)
2. Cheung, K.W., So, H.C., Ma, W.K., Chan, Y.T.: A Constrained Least Squares Approach to Mobile Positioning: Algorithms and Optimality. EURASIP J. on Applied Sig. Proc. 2006(ID 20858), 1–23 (2006)
3. Clarkson, I.V.L.: Frequency estimation, phase unwrapping and the nearest lattice point problem. In: ICASSP 1999, pp. 1609–1612 (1999)
4. Emile, B., Comon, P.: Estimation of time delays between unknown colored signals. Signal Proc. 69(1), 93–100 (1998)
5. Omlor, L., Giese, M.: Blind source separation for over-determined delayed mixtures. In: Advances in Neural Information Processing Systems, vol. 19, pp. 1049–1056. MIT Press, Cambridge (2007)
6. Parra, L., Spence, C.: Convolutive blind separation of non-stationary sources. IEEE Trans. on Speech and Audio Processing 8(3), 320–327 (2000)

7. Rife, D.C., Boorstyn, R.R.: Single-tone parameter estimation from discrete-time observations. IEEE Trans. Inform. Theory IT 20(5), 591–598 (1974)
8. Yeredor, A.: Blind Source Separation with Pure Delays Mixture. In: ICA 2001 (2001)
9. Zhou, Y., Lamont, L.: Constrained least squares approach for TDOA localization: a global optimum solution. In: ICASSP 2008. pp. 2577–2580 (2008)