Tensor decomposition exploiting diversity of propagation velocities; application to localization of icequake events

Francesca Raimondi, Pierre Comon, Olivier J.J. Michel, Souleymen Sahnoun, Agnes Helmstetter

Tensor decomposition exploiting diversity of propagation velocities; application to localization of icequake events[☆]

Francesca Raimondi[1], Pierre Comon[1], Olivier Michel[1], Souleymen Sahnoun[1], Agnes Helmstetter[2]

Abstract

The problem of direction of arrival (DoA) estimation of seismic plane waves impinging on an array of sensors is considered from a new deterministic perspective through tensor decomposition techniques. In addition to temporal and spatial sampling, further information is taken into account, based on the different propagation speed of volume waves (P and S) through solid media. Performances are evaluated through simulated data in terms of the Cramér-Rao bounds and compared to other reference methods such as ESPRIT and MUSIC, in the presence of additive Gaussian circular noise. The proposed approach is then applied to real seismic data recorded at Argentière glacier, occurring at the interface between the ice mass and the underlying bedrock.

Keywords: localization, DoA estimation, antenna array processing, tensor, seismics, elastic waves, propagation, icequakes

1. Introduction

In many cases, the superimposition principle applies in practical problems, provided the nonlinearity domain is not reached (turbulence, saturation, etc). This allows us to model the physical phenomena as linear combinations of simpler ones. In this paper, we are interested in the decomposition of a multivariate function into a sum of functions whose variables separate. In particular, this simplified model is relevant in narrow-band antenna array processing in the far-field, which we consider in the present framework.

In the context of seismic monitoring, seismology aims at studying waves generated by rupture phenomena taking place within the volume of interest (earth, ice, etc.). Although the most interesting events take place at a certain depth - mostly unknown - within the analyzed volume, acquisition systems and sensor arrays are most often situated at a given distance from the surface, given the difficulty of accessing to deeper layers. The main quantity to be measured is ground displacement (in the form of its derivative - velocity - or its second derivative - acceleration), produced by impinging elastic waves. The localization of the sources forms the first requirement of data analysis, in order to prevent damage provoked by seismic events, and to monitor the activity of complex structures such as glaciers or volcanos. In wider terms, direction of arrival (DoA) estimation is a central problem in array signal processing, concerning several areas of engineering including telecommunications, speech, astronomy, seismology, and medical applications. Array processing requires a set of multiple sensors placed at different positions in space, receiving source signals from different directions [1], [2]. The techniques aiming at estimating the direction of arrival of narrowband sources can be grouped in two categories: the algorithms based on exhaustive search (including beamforming and MUSIC [3]) and the algebraic approaches (including root-MUSIC [4] and ESPRIT [5]). Whenever propagation speed is considered as a constant v, traditional array processing only relies on a temporal and spatial sampling of the propagating wavefield. MUSIC (multiple signal classification) algorithm is based on the spectral decomposition of the sample covariance matrix under the spatially white noise assumption (to be presented in Appendix B). This method has the advantage of being *statistically efficient*, unlike beamforming techniques, despite a serious sensitivity to SNR and resolution limitations for correlated or closely spaced sources [2]. Moreover, the algorithm requires the perfect knowledge of the position of each sensor. ESPRIT (estimation of signal parameters via rotational invariance techniques) applies to an array composed of two identical subarrays displaced one from the other according to an unknown translation vector, whereas the calibration of the array is needed otherwise. The concept of signal-subspace processing embodied by MUSIC and ESPRIT can be generalized to the wide-band case [6, 7, 38].

A deterministic approach based on tensor decomposition has been introduced in [8], through the extension of the rotational invariance principle to more than one displacement. It provides the localization of more sources than sensors, with less restrictive requirements for signal stationarity than the afore mentioned statistical methods. The advantage of tensor decompositions lies in the need for shorter data records, since the estimation of statistical quantities from available samples is not a requirement anymore. Furthermore, like ESPRIT, it allows to estimate the impinging signals up to a scale factor, without

[☆]This work has been supported by ERC AdG-2013-320594 grant "DECODA".

Email address: francesca.raimondi@gipsa-lab.grenoble-inp.fr (Francesca Raimondi)

[1]Gipsa-Lab, CNRS, BP 46, 38402 Saint Martin d'Heres Cedex, France

[2]ISTerre, CNRS, OSUG-C, 38400 Saint Martin d'Heres Cedex, France

resorting to a spatial matched filter. The tensor model puts forward parsimony and separability [9]:

1. Parsimony expresses a function f as a finite sum of simpler constituents:
$$g = \sum_{r=1}^{R} \varsigma_r h_r \tag{1}$$

2. Separability decouples a function h that depends on multiple factors into a product of simpler constituents ϕ_d, $d = 1, \ldots, D$, each one depending only on one factor $\boldsymbol{x}_d$:
$$h(\boldsymbol{x}_1, ..., \boldsymbol{x}_D) = \prod_{d=1}^{D} \phi_d(\boldsymbol{x}_d)$$

In the field of array processing for source separation and DoA estimation, R refers to the number of sources impinging on an array, and D to the tensor order, i.e. the dimension of multilinearity within the model:

$$g(\boldsymbol{x}_1, ..., \boldsymbol{x}_D) = \sum_{r=1}^{R} \varsigma_r \prod_{d=1}^{D} \phi_{rd}(\boldsymbol{x}_d)$$

Tensor decomposition derives from the need to solve the *inverse problem*, i.e., the identification of factors ϕ_{rd} based on noisy measurements of f: as it will be hereinafter discussed, the direction of arrival can be extracted after the resolution of this problem. For this purpose, the most common tensor decomposition for *three-way* arrays is the polyadic decomposition: the measurements are stored in a multidimensional array and decomposed into a sum of *rank one* terms [10, 8]. A *decomposable* three-way tensor can be defined by a vector triplet [3]:

$$D_{lmk} = a_l b_m c_k \text{ or equivalently } \mathcal{D} = \boldsymbol{a} \otimes \boldsymbol{b} \otimes \boldsymbol{c}$$

Any order-3 tensor admits a decomposition into a sum of decomposable tensors:

$$\mathcal{M} = \sum_{r=1}^{R} \varsigma_r \, \mathcal{D}(r) \tag{2}$$

where coefficients ς_r can always be chosen to be real positive, and decomposable tensors $\mathcal{D}(r)$ to have unit norm, *i.e.* for L^p norms, $\|\mathcal{D}\| = \|\boldsymbol{a}\| \|\boldsymbol{b}\| \|\boldsymbol{c}\| = 1$. The minimal value of R such that this decomposition holds is called the tensor rank: if R is not too large, the corresponding decomposition is unique and deserves to be referred to as Canonical Polyadic (CP); other terminologies include *rank decomposition* or *Candecomp/Parafac*. Hence, decomposable tensors have a rank equal to 1, by definition [10]. Now in terms of coordinates, tensor $\mathcal{M}$ is represented by a $L \times M \times K$ three-way array, which consequently decomposes as

$$M_{lmk} = \sum_{r=1}^{R} \varsigma_r A_{lr} B_{mr} C_{kr} \tag{3}$$

where the three factor matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ have unit norm columns. This is equivalent to the general R-term trilinear model

$$\mathcal{M} = \sum_{r=1}^{R} \varsigma_r \boldsymbol{a}_r \otimes \boldsymbol{b}_r \otimes \boldsymbol{c}_r \tag{4}$$

where each array $\boldsymbol{a}_r \otimes \boldsymbol{b}_r \otimes \boldsymbol{c}_r$ is a rank-1 array.

This paper is aimed at exploiting *another type of diversity*, in addition to spatial and temporal sampling traditionally employed in array processing (cf. Section 3 for a detailed explanation of the concept of diversity): the propagation speed diversity [4] of volume waves through solids, namely pressure (P) and shear (S) waves. Current array processing methods like [3, 5] only focus on information conveyed by a single volume wave, like the P wave, whereas the contents delivered by the other is somehow wasted. Our approach intends to exploit this information as a whole, whereas translation invariance used in [8] is no longer necessary.

This paper is organized as follows. Section 2 presents the physical model and the assumptions. Section 3.1 makes a synthesis of the main narrowband 2D preexisting algorithms. Our deterministic method, exploiting the propagation speed diversity of volume waves, is presented in Section 3. Section 4 is devoted to the description of the employed algorithms for tensor decomposition and DoA estimation. Simulated data and real seismic data are treated in Sections 5 and 6 respectively, where results are compared with traditional methods detailed in Section 3.1 and with a statistical performance study.

2. Physical modeling and assumptions

Matrix algebra notations and main assumptions are hereby introduced:

$\boldsymbol{v}$	vector $\boldsymbol{v}$
v_i	element of $\boldsymbol{v}$
$\boldsymbol{v}^T\boldsymbol{u}$	scalar product between real vectors $\boldsymbol{v}$ and $\boldsymbol{u}$
$\boldsymbol{v} \otimes \boldsymbol{u}$	outer (tensor) product of $\boldsymbol{v}$ and $\boldsymbol{u}$
$\boldsymbol{A}$	matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	i-th column of $\boldsymbol{A}$
A_{ij}	element of $\boldsymbol{A}$
$\boldsymbol{A}^T$	Transpose of $\boldsymbol{A}$
$\boldsymbol{A}^*$	Complex conjugate of $\boldsymbol{A}$
$\boldsymbol{A}^H$	Conjugate transpose of $\boldsymbol{A}$
$\boldsymbol{A}^\dagger$	Moore-Penrose pseudoinverse of $\boldsymbol{A}$
$\boldsymbol{A} \boxtimes \boldsymbol{B}$	Kronecker product between $\boldsymbol{A}$ and $\boldsymbol{B}$
$\boldsymbol{A} \odot \boldsymbol{B}$	Katri-Rao (column-wise Kronecker) product
$\boldsymbol{A} \boxdot \boldsymbol{B}$	Hadamard (element-wise) product
$\|\cdot\|_F$	Frobenius norm
$\mathcal{T}$	three-way tensor $\mathcal{T}$
T_{ijk}	element of $\mathcal{T}$

[3] Once bases in every linear space are fixed, tensors are defined by their array of coordinates. See [10] for details.

[4] Since the focus is on narrow-band processing, the distinction between group and phase propagation velocities is irrelevant.

A1 *Far-field approximation*: the distance of the source from the receiving array is much greater than the array aperture. This is equivalent to assume a planar wavefront at the sensor level. Moreover, sensors and sources are considered point-like, as their size is negligible with respect to the source-to-sensor distance. Because of the rotational indeterminacy around the axis defined by an Uniformly spaced Linear Array (ULA), we do not restrict the generality of this configuration by assuming that sensors and sources are coplanar (we can thus work in a 2-D coordinate systems).

A2 *Narrow-band in base-band*: Signals of interest are the product of a varying amplitude (complex envelope) and a high-frequency signal (cf. Appendix A). We assume that the spectral supports of both parts do not overlap (this is sometimes referred to as the Bedrosian condition). [5] Under this condition, one can work in base-band with the complex envelope of the low-pass signal. For this type of signal, a time delay of the original signal is equivalent to a phase shift of the complex envelope.

A3 *Homogeneous and isotropic medium at the antenna level*: Ray-paths can be approximated by straight lines.

A4 *Dissipation* at the antenna scale is excluded, as the array dimension is negligible with respect to dissipation characteristic length.

A5 The impulse responses of *particle motion* is the only source of variability between impinging P and S waves.

In the context of beamforming applied to seismic events originating at depth, one has to deal with a double arrival of elastic volume waves: P waves and S waves. They have different propagation velocities and a variable frequency content: the former tend to have a higher frequency content than the latter. The P wave arrives first and its particle motion is parallel to the propagation direction, whereas the S wave (the second arrival observed on seismic records) is transversal (its particle motion is perpendicular to the propagation direction). Both are linearly polarized. Linear polarization refers to the fact that more coherent seismic energy is located in one principal direction of particle motion [12]. P waves and S waves have theoretical velocities v_1 and v_2, respectively, given by Ref. [13]:

$$\begin{cases} v_1 = \left(\frac{\lambda+2\mu}{\rho}\right)^{\frac{1}{2}} \\ v_2 = \left(\frac{\mu}{\rho}\right)^{\frac{1}{2}} \end{cases}$$

[5] The Bedrosian theorem [11] is hereby stated: The Hilbert transform of the product of two complex valued functions
$f, g : \mathbb{R} \to L^2(\mathbb{R})$ with non-overlapping Fourier spectra
($F(f) \approx 0$ for $|f| > a$ and $G(f) \approx 0$ for $|f| < a$ where a is a positive constant)
is given by the product of the low-frequency signal f and the Hilbert transform of the high-frequency signal g:

$$H(f)(x)g(x)) = f(x)H(g(x)),\ x \in \mathbb{R}$$

where λ and μ are *Lamé's constants* and ρ is the density of the medium, assumed to be positive constant, thus making v_1 greater than v_2. In particular, the velocity of the S wave ranges from zero up to 70% of the velocity of the P wave. For fluids μ is zero, $v_2 = 0$ and therefore S waves do not propagate.

As for the physical model of wave propagation, the following properties are assumed:

P1 Pressure (P) and Shear (S) waves propagate at v_1 and v_2 velocities, respectively, under the approximation of non dispersive medium.

and in order to allow the multilinearity and separability of the final model:

P2 P and S signals are filtered in a narrow-band around f_1 and f_2, respectively, such that $\frac{v_1}{f_1} = \frac{v_2}{f_2} = \lambda$ is constant.

P3 Base-band P and S signals propagating from the same source event and received in two different narrow bands around f_1 and f_2 are proportional: $w^2(t) = \alpha(f_1, f_2)\, w^1(t)$; see Appendix A for more details.

Furthermore, we can summarize the following notations, which will be subsequently used throughout the tensor decomposition phase:

N1 The first sensor composing the array is taken as the origin point for translation in space.

N2 In order to avoid scale indeterminacies in tensor decomposition, the P wave $w^1(t)$ is fixed up to a scale factor as the product of a complex envelope $w(t)$ modulated by a high frequency contribution $m(t)$

$$w^1(t) \approx w(t)m(t)$$

(see Appendix A for more details).

Additional hypotheses or notations are progressively introduced when needed by 2-D traditional algorithms (MUSIC [3] and ESPRIT [5]):

H1 Identical sensor responses (calibration).

H2 Incident signals are uncorrelated to noise.

H3 The number of sources of interest, R, is smaller than the number of sensors L: $R < L$.

H4 Noise spatial coherency is known. Therefore, one can always consider (thanks to spatial prewhitening) that the noise covariance matrix is proportional to identity: $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^H\} = \sigma_n^2 \boldsymbol{I}$, where σ_n may be unknown, after whitening is applied.

H5 Noise is additive and Gaussian complex circular: $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^T\} = \boldsymbol{0}$.

H6 The number of time samples M is greater than the number of sensors L: $M > L$.

3. Tensor-based solution

The advantage of the tensor formalism lies in its ability to restore identifiability of parameters: the impinging signals, their directions of arrival, and the related connection between P and S waves. On the other hand, the multidimensional character of tensor models requires the presence of at least 3 types of diversity. We review below some diversities that may be available in antenna array processing.

1. *Time diversity*: Every signal w is a function of time t. Moreover, it may be stationary or transient. Thus, recorded signal g^1 is given by the sum of R simultaneous sources:

$$g^1(t) = \sum_{r=1}^{R} w_r(t) \iff g^1_m = \sum_{r=1}^{R} B_r(t_m), \tag{5}$$

where $m = 1, \ldots, M$, and the right equation represents a discretization of the left one.

2. *Space diversity*. The basis of traditional array processing consists in performing a spatial sampling, in addition to the temporal diversity. According to assumptions **A**2, **A**1 and **H**1

$$\begin{aligned} g^2(t, \boldsymbol{b}) &= \sum_{r=1}^{R} w_r(t) \exp\left\{\frac{2\pi}{\lambda} \boldsymbol{b}^T \boldsymbol{d}_r\right\} \\ \iff G^2_{ml} &= \sum_{r=1}^{R} B_r(t_m) = \sum_{r=1}^{R} A_r(\boldsymbol{b}_l)\, B_r(t_m) \end{aligned} \tag{6}$$

where $l = 1, \ldots, L$ refers to the sensor cardinality with respect to the reference, λ is the observed narrowband wavelength, $\boldsymbol{b} \in \mathbb{R}^3$ is the recording position within the acquisition system and $\boldsymbol{d}_r$ is the unit vector pointing to the r-th source. [6]

3. *Translational diversity* (refer to [5] for one rotational invariance and to [8] for multiple roto-translations):

$$\begin{aligned} g^3(t, \boldsymbol{b}, \boldsymbol{\delta}) &= \sum_{r=1}^{R} w_r(t) \exp\left\{\frac{2\pi}{\lambda} \boldsymbol{b}^T \boldsymbol{d}_r\right\} \exp\left\{\frac{2\pi}{\lambda} \boldsymbol{\delta}^T \boldsymbol{d}_r\right\} \\ \iff G^3_{mlk} &= \sum_{r=1}^{R} A_r(\boldsymbol{b}_l)\, B_r(t_m)\, D_r(\boldsymbol{\delta}_k) \end{aligned} \tag{7}$$

where $k = 1, \ldots, K$ is the index of the translation with respect to the reference array, and $\boldsymbol{\delta}$ is the translational invariant repeating the array configuration over space. ESPRIT may be seen as a particular case of this diversity, when translation vector $\boldsymbol{\delta}$ relates $K = 2$ identical subarrays, whereas the tensor model itself can be applied to $K > 2$ identical configurations.

[6] For sake of simplicity, we'll refer to both the proper direction of arrival $\boldsymbol{d}_r$ and the angle of arrival θ_r with the acronym DoA.

4. *Polarization diversity*. Expressions (5), (6) and (7) refer to one-component (1C) sensors, but can be generalized to three-component (3C) sensors as well. Moreover, in the latter case, one can take advantage of an additional diversity related to polarization, as explained in [14, 9]:

$$\boldsymbol{g}^4(t) = \sum_{r=1}^{R} w_r(t)\, \boldsymbol{p}(\boldsymbol{\Phi}_r) \iff G^4_{mk} = \sum_{r=1}^{R} B_r(t_m)\, p_k(\boldsymbol{\Phi}_r) \tag{8}$$

where $\boldsymbol{p}(\cdot)$ is the polarization response vector of the considered sensor and $\boldsymbol{\Phi}_r = [\psi_r, \theta_r, \alpha_r, \beta_r]^T$ contains the polarization information of the r-th impinging wavefield with respect to the sensor position: the azimuth ψ, the elevation θ, the orientation angle α and the ellipticity angle β. In particular, the elevation θ refers to the angle of arrival when assumption **A**1 is made. Discrete index $j \in \{1, 2, 3\}$ refers to one out of three components of the 3C sensor vector $\boldsymbol{p}$.

5. *Repetition diversity* is a possible extension of dimensionality whenever we deal with multiple events describing the same physical phenomenon. The additional assumption of a linear relation between events from the same cluster has to be made, so that multiple events describing the same physical phenomenon are related to each other by a complex coefficient:

$$g^5(j, t) = \sum_{r=1}^{R} \gamma(j)\, w_r(t) \iff G^5_{jm} = \sum_{r=1}^{R} \gamma(j)\, B_r(t_m) \tag{9}$$

where discrete index $j = 1, \ldots, J$ is the cardinality of the event of the cluster (1 being the first recorded occurrence and I the last one), and γ a (generally unknown) complex coefficient. Repetition diversity has been already used in other contexts, in [15] for instance.

6. *Propagation speed diversity* may be seen as equivalent to a frequency diversity between P and S waves, under assumption **A**2 and properties **P**1, **P**2 and **P**3. The complete model on which this paper is based will be subsequently developed on the basis of the linear decomposition below:

$$g^6(f, t) = \sum_{r=1}^{R} \alpha(f)\, w_r(t) \iff G^6_{mk} = \sum_{r=1}^{R} B_r(t_m)\, C_r(f_k) \tag{10}$$

where again the second equation represents the discretization of the first one, f_k denote working frequencies for P and S waves ($k = 1$ for the P wave and $k = 2$ for the S wave), as defined in **P**2, chosen by the user as a function of propagation velocities v_1 and v_2, and $\alpha(\cdot)$ a complex unknown coefficient.

The focus of this paper consists in integrating at the same time the spatial, temporal and speed diversity, respectively embodied by matrices $\boldsymbol{A} \in \mathbb{C}^{L \times R}$, $\boldsymbol{B} \in \mathbb{C}^{M \times R}$ and $\boldsymbol{C} \in \mathbb{C}^{2 \times R}$.

3.1. Traditional matrix-based solutions

Traditional 2-D solutions only employ the concepts of *time diversity* and *space diversity*: base-band signals are stored in a

data matrix $\boldsymbol{Y}$ of size $L \times M$, where L is the number of sensors and M the number of time samples:

$$\boldsymbol{Y}(t_m) = \begin{bmatrix} \boldsymbol{y}(t_1) \\ \vdots \\ \boldsymbol{y}(t_M) \end{bmatrix}, \qquad m = 1, ..., M$$

where m refers to the time sample. In the particular instance of assumptions **A**1 and **A**2, each scalar element is given by

$$y_l(t_m) = \sum_{r=1}^{R} A_{lr} w_r(t_m) + n_l(t_m), \qquad l = 1, ..., L$$

where

$$A_{lr} = \exp\left\{ i \frac{\omega}{c} (\boldsymbol{b}_l^T \boldsymbol{d}_r) \right\}$$

where $\boldsymbol{b} \in \mathbb{R}^3$ is the sensor position within the acquisition system and $\boldsymbol{d}_r$ is the unit vector pointing to the r-th source.

Vector $\boldsymbol{y}(t_m)$ depends on two main factors, as in (6): the $L \times R$ steering matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots \boldsymbol{a}_R]$, and the $R \times 1$ source waveform vector $\boldsymbol{s}(t_m) = [\ w_1(t_m) \quad \cdots \quad w_R(t_m)\ ]^T$. This can be written in the following compact form:

$$\boldsymbol{y}(t_m) = \boldsymbol{A}\boldsymbol{s}(t_m) + \boldsymbol{n}(t_m), \qquad m = 1, ..., M$$

or:

$$\boldsymbol{Y} = \left[\ \boldsymbol{y}(t_1), \quad \cdots, \quad \boldsymbol{y}(t_M)\ \right] = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{N}$$

where appears the $R \times M$ source waveform matrix $\boldsymbol{S} = [\boldsymbol{s}(t_1), \cdots, \boldsymbol{s}(t_M)]$.

This formulation is the basis of MUSIC and ESPRIT algorithms, which are detailed in Appendix B.

3.2. Speed diversity in tensor format

P and S waves are generated by the same physical source but propagate at two different, v_1 and v_2 respectively. Thus, P and S signals received at position $\boldsymbol{b}$ and at time t are respectively given by

$$\begin{cases} q^P(\boldsymbol{b}, t) = \sum_{r=1}^{R} w_r^P(t - \tau_r^P(\boldsymbol{b})) + n(t, \boldsymbol{b}) \\ q^S(\boldsymbol{b}, t) = \sum_{r=1}^{R} w_r^S(t - \tau_r^S(\boldsymbol{b})) + n(t, \boldsymbol{b}) \end{cases}$$

where $\{\tau_r^P, \tau_r^S\}$ express the respective delays for P and S signals recorded at position $\boldsymbol{b}$.

The signal received at a point $\boldsymbol{b}$ in space contains R sources of interest plus an additive noise: under the assumption **A**1 and **A**2, the time delays of P and S waves correspond to the following phase shifts respectively: $\psi_r^P(\boldsymbol{b}) = \frac{\omega}{v_1}(\boldsymbol{b}^T \boldsymbol{d}_r)$, $\psi_r^S(\boldsymbol{b}) = \frac{\omega}{v_2}(\boldsymbol{b}^T \boldsymbol{d}_r)$.

With property **P**2, records produced by the r-th source contain two delayed narrowband waves propagating around frequencies f_1, f_2 at velocities v_1 and v_2, yielding two resulting waves with the same wavelength λ. This model is equivalent to the general three-variable expression

$$q(t, \boldsymbol{b}, f) = \sum_{r=1}^{R} \alpha_r(f)\, w_r(t) \exp\left\{ i \frac{2\pi}{\lambda} (\boldsymbol{b}^T \boldsymbol{d}_r) \right\} + n(t, \boldsymbol{b})$$

Thus, in the absence of noise $n(t, \boldsymbol{b})$, function $q(t, \boldsymbol{b}, f)$ decomposes into a sum of R functions whose variables separate.

Discretization yields an array of finite dimensions $2 \times L \times M$ (L sensors located at $\boldsymbol{b}_\ell$, M time samples t_m and $K = 2$ propagation velocities v_k), so that the data array writes:

$$Z_{lmk} = \sum_{r=1}^{R} \alpha_r(f_k) w_r(t_m) \exp\left\{ i \frac{2\pi}{\lambda} \boldsymbol{b}_l^T \boldsymbol{d}_r \right\} + n_l(t_m)$$

where frequency f_k, $k = 1, 2$ is fixed according to property **P**2 $\left(\lambda = \frac{v_1}{f_1} = \frac{v_2}{f_2}\right)$. In the case of an ULA, we just have that $\boldsymbol{b}_l^T \boldsymbol{d}_r = (l-1)\Delta \sin\theta_r$, where θ_r refers to the angle of arrival of the r-th source $\boldsymbol{d}_r$, and Δ is the distance between two consecutive sensors composing the ULA.

3.3. Exact CP decomposition

A CP decomposition is said to be *essentially* unique if there exist a unique set $\{\varsigma_r, \mathcal{D}(r),\ 1 \le r \le R\}$ such that equality holds in (2). However, even if the CP decomposition is unique, there exist several ways of writing (3). In fact, writing a decomposable tensor as the outer product of vectors is subject to scaling indeterminacies, which actually stem from multilinearity properties of tensors [10], since

$$\alpha\boldsymbol{a} \otimes \beta\boldsymbol{b} \otimes \gamma\boldsymbol{c} = \boldsymbol{a} \otimes \boldsymbol{b} \otimes \boldsymbol{c}$$

if $\alpha\beta\gamma = 1$. Even if the above vectors are of unit L^p norm, there remain two scaling indeterminacies of unit modulus. As a consequence, attention should be paid to the difference between CP or *essential* uniqueness, and uniqueness of matrix factors appearing in (3).

There exist sufficient conditions ensuring CP uniqueness, *e.g.* the Kruskal condition [8]:

$$\kappa_A + \kappa_B + \kappa_C \ge 2R + 2 \tag{11}$$

where the notation κ_A refers to the *Kruskal*-rank[7] of matrix $\boldsymbol{A}$. However, less stringent conditions guaranteeing almost surely a unique solution can be found, for instance [9, 10, 16]:

$$R(K + L + M - 2) < lmk$$

In the present paper, we deal with simpler examples when $R = 1$ as for glacier data (cf Section 6) and $R = 2$ relative to simulated data (cf Section 5) so that essential uniqueness of a decomposition is always guaranteed.

In practice, it is more convenient to fix *trivial* indeterminacies of unit modulus in Equation (3):

$$M_{lmk} = \sum_{r=1}^{R} \varsigma_r A_{lr} B_{mr} C_{kr}$$

Therefore, as suggested in [17], we fix $2R$ parameters without restricting the generality. More precisely, based on properties described in Section 2, we assume the following:

[7] The Kruskal rank of a matrix $\boldsymbol{A}$ is the largest number κ_A such that any subset of κ_A columns are lineraly independent.

1. From property **N**1

$$\exp\{i\psi_{1r}\} = 1\ \forall r \Longrightarrow A_{1r} = 1\ \forall r \tag{12}$$

2. From property **P**3:

$$\alpha_r(f_1) = 1,\ \forall r \Longrightarrow C_{1r} = 1\ \forall r \qquad \text{cf (3)} \tag{13}$$

Note that after fixing the scaling indeterminacies, the number of remaining free parameters is now $(K+L+M-2)R$ [17], and not $(K+L+M-3)R$ as assumed in [8].

3.4. Presence of noise: DoA Cramér-Rao bounds

Cramér-Rao Bound (CRB) represents the lower bound on the variance of any unbiased estimator of a deterministic parameter. We hereby detail the derivation of the CRB in order to asses the performance of the proposed method and then compare different methods with one another. If noise $\boldsymbol{n}$ is considered as an i.i.d. complex circularly symmetric Gaussian variable of known variance σ_n^2, Signal-to-Noise ratio (SNR) is defined as (cf. [18])

$$SNR = 10\log_{10}\frac{\|\mathcal{M}\|_F^2}{LMK\sigma_n^2}$$

where operator $\|\cdot\|_F^2$ indicates Frobenius norm: $\|\mathcal{M}\|_F = \sqrt{\text{trace}\{\boldsymbol{\mu}^H\boldsymbol{\mu}\}}$ and $\boldsymbol{\mu} = \text{vec}\{\mathcal{M}\}$. The problem of scale ambiguity of the CP decomposition is solved by the assumption of expressions (12) and (13) that the first row of matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ is normalized to $[1\cdots1]_{1\times R}$. Therefore, the unknown complex parameter vector $\boldsymbol{\vartheta} \in \mathbb{C}^{(K+L+M-2)R\times 1}$ has the following form:

$$\boldsymbol{\vartheta} = [\bar{\boldsymbol{a}}_1^T, \cdots, \bar{\boldsymbol{a}}_R^T, \boldsymbol{b}_1^T, \cdots, \boldsymbol{b}_R^T, \bar{\boldsymbol{c}}_1^T, \cdots, \bar{\boldsymbol{c}}_R^T]$$

where $\bar{\boldsymbol{a}}_r^T \overset{\text{def}}{=} [A_{2r}, \cdots, A_{Lr}]$, $\boldsymbol{b}_r^T \overset{\text{def}}{=} [B_{1r}, \cdots, B_{Lr}]$, $\bar{\boldsymbol{c}}_r^T \overset{\text{def}}{=} [C_{2r}, \cdots, C_{Lr}]$.

The likelihood function for a zero mean, circularly symmetric complex Gaussian noise with covariance $\sigma_n^2\boldsymbol{I}$ takes the form

$$L(\boldsymbol{z}|\boldsymbol{\vartheta}) = \frac{1}{(\pi\sigma_n^2)^{lmk}}\exp\left\{-\frac{1}{\sigma_n^2}(\boldsymbol{z}-\boldsymbol{\mu}(\boldsymbol{\vartheta}))^H(\boldsymbol{z}-\boldsymbol{\mu}(\boldsymbol{\vartheta}))\right\}$$

The corresponding log-likelihood function $f(\boldsymbol{\vartheta}) = \log(L(\boldsymbol{z}|\boldsymbol{\theta}))$ can be written as

$$f(\boldsymbol{\vartheta}) = -lmk\log(\pi\sigma_n^2) - \frac{1}{\sigma_n^2}(\boldsymbol{z}-\boldsymbol{\mu}(\boldsymbol{\vartheta}))^H(\boldsymbol{z}-\boldsymbol{\mu}(\boldsymbol{\vartheta})).$$

Then, the complex Fisher information matrix (FIM) is given by (cf. [18, 17])

$$\boldsymbol{\Phi}(\boldsymbol{\vartheta}) = \mathbb{E}\left\{\left(\frac{\partial f(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}\right)^H\left(\frac{\partial f(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}\right)\right\}.$$

A straightforward derivation leads to

$$\frac{\partial f(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}} = \frac{1}{\sigma_n^2}\left[\boldsymbol{n}^T\frac{\partial\boldsymbol{\mu}^*}{\partial\boldsymbol{\vartheta}} + \boldsymbol{n}^H\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\vartheta}}\right]$$

where $\boldsymbol{n} = \boldsymbol{z} - \boldsymbol{\mu}$.

Since noise is circularly-symmetric ($\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^H\} = \sigma_n^2\boldsymbol{I}_{lmk}$ and $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^T\} = \boldsymbol{0}$), the FIM reduces to [17]:

$$\boldsymbol{\Phi}(\boldsymbol{\vartheta}) = \frac{1}{\sigma_n^2}\left[\left(\frac{\partial\boldsymbol{\mu}^*}{\partial\boldsymbol{\vartheta}}\right)^H\left(\frac{\partial\boldsymbol{\mu}^*}{\partial\boldsymbol{\vartheta}}\right) + \left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\vartheta}}\right)^H\left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\vartheta}}\right)\right]$$

The CRB of any unbiased estimator of complex vector parameter $\boldsymbol{\vartheta}$ is is given by the inverse of the FIM (see Appendix C for more details about CRB of the DoA estimation).

4. Algorithms

4.1. Computation of the CP decomposition

We hereby recall the expressions of the R-term trilinear model for the three-way array $\mathcal{M}$:

$$\begin{aligned} M_{lmk} &= \sum_{r=1}^{R}\varsigma_r\, A_{lr}B_{mr}C_{kr} \\ \mathcal{M} &= \sum_{r=1}^{R}\varsigma_r\,\boldsymbol{a}_r\otimes\boldsymbol{b}_r\otimes\boldsymbol{c}_r \end{aligned} \tag{14}$$

An exact direct decomposition algorithm has been proposed in [19] for $K \geq 2$: it takes as inputs the K data slabs $\boldsymbol{Z}[:,:,1], \boldsymbol{Z}[:,:,2], \cdots, \boldsymbol{Z}[:,:,K]$ and the number of factors R, and returns the estimates of matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$. Since the third dimension K investigated in the present paper refers to the propagation speed diversity, it corresponds to $K = 2$ (P and S waves). The uniqueness of such a decomposition is guaranteed provided that the R column vectors corresponding to two of the ways are linearly independent, and the R column vectors associated with the third way have the property that no two are collinear (see below). The trilinear model (14) can be rewritten as K matrices $\boldsymbol{\zeta}_k$ of size $L\times M$, according to the *trilinear matrix equations* [20, 19]:

$$\boldsymbol{\zeta}_k = \sum_{r=1}^{R}(\boldsymbol{a}_r\otimes\boldsymbol{b}_r)\,c_r[k] = \sum_{f=1}^{F}\boldsymbol{a}_r\,c_r[k]\,\boldsymbol{b}_r^T = \boldsymbol{A}\boldsymbol{D}_k\boldsymbol{B}^T$$

where

$$\begin{aligned} \boldsymbol{A} &= [\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots\boldsymbol{a}_r], \\ \boldsymbol{B} &= [\boldsymbol{b}_1,\ \boldsymbol{b}_2,\ \cdots\boldsymbol{b}_r], \\ \boldsymbol{D}_k &= \text{Diag}\{c_1[k],\ c_2[k],\cdots,\ c_r[k]\},, \qquad k = 1,\ldots,K \end{aligned}$$

and coefficients ς_r were pulled in factor matrix $\boldsymbol{B}$.

The following more constraining identifiability conditions are employed, which concurrently imply the Kruskal condition in (11):

IC1 The columns of $\boldsymbol{A}$ are linearly independent *i.e.* $\kappa_A \geq R$.

IC2 The columns of $\boldsymbol{B}$ are linearly independent *i.e.* $\kappa_B \geq R$.

IC3 Every pair of columns of $\boldsymbol{C}$ is linearly independent *i.e.* $\kappa_C \geq 2$.

Each condition refers to one way of the array: the first two conditions state that there must be at least R factors present in two ways. The third requires that no two factors are linked by a proportional relationship along the other way.

The decomposition proposed in [19] exploits the comparison between the following linear combinations, for every k-vector of weights $\boldsymbol{w}$:

$$\begin{cases} \bar{\boldsymbol{\zeta}}(\boldsymbol{w}) = \sum_{k=1}^{K} \boldsymbol{w}[k]\mu[k] \\ \bar{\boldsymbol{D}}(\boldsymbol{w}) = \sum_{k=1}^{K} \boldsymbol{w}[k]\boldsymbol{D}_k \end{cases},$$

so that $\boldsymbol{\zeta}_+ = \bar{\boldsymbol{\zeta}}(\mathbf{1}_K)$ and $\boldsymbol{D}_+ = \bar{\boldsymbol{D}}(\mathbf{1}_K)$. The following result constitutes the core of the decomposition algorithm (for more details, see [19]):

Theorem 4.1. *Let $\boldsymbol{USV}^T$ be the singular value decomposition of $\boldsymbol{\zeta}_+$, and define*

$$\boldsymbol{\eta}_k = \boldsymbol{U}\boldsymbol{\zeta}_k\boldsymbol{V}\boldsymbol{S}^{-1}, \qquad k = 1, ..., K$$

where $\boldsymbol{\eta}_k$ has size $R \times R$, $\boldsymbol{U}$ has size $R \times L$, $\boldsymbol{\zeta}_k$ has size $R \times R$, $\boldsymbol{V}$ has size $R \times M$. Then if

(i) the trilinear equation (3), and

(ii) the parametrization conventions (12) and (13)

hold, we have the following results:

(iii) if identifiability conditions IC1-IC2 hold, the columns of $\boldsymbol{Z} = \boldsymbol{U}^T\boldsymbol{A}$ are right eigenvectors of $\boldsymbol{\eta}_k$ with eigenvalues equivalent to the diagonal elements of $\boldsymbol{\Lambda}_k = \boldsymbol{D}_k\boldsymbol{D}_+^{-1}$, $k = 1, ..., K$;

(iv) if identifiability conditions IC1-IC3 hold, the columns of $\boldsymbol{Z} = \boldsymbol{U}^T\boldsymbol{A}$ are the only common eigenvectors of $\boldsymbol{\eta}_k$.

Notice that in the above description, $K = 2$ ($k = 1$ for the P wave and $k = 2$ for the S wave).

4.2. Extraction of DoA from matrix factors

Once the decomposition has been obtained, the parameters of the trilinear problem can be extracted as follows.

1. Direction of arrival of the r-th source θ_r is calculated from

$$A_{lr} = \exp\left\{i\frac{2\pi}{\lambda}(\boldsymbol{b}_l^T\boldsymbol{d}_r)\right\} = \exp\left\{i\frac{2\pi}{\lambda}\Delta(l-1)\sin(\theta_r)\right\}$$

Therefore,

$$\hat{\theta}_{lr} = \arcsin\left[\frac{2\pi}{\lambda}\log(A_{lr})\frac{1}{\Delta(l-1)}\right].$$

In order to estimate direction of arrival $\boldsymbol{\theta}$, matrix $\boldsymbol{A}$ can be considered as a *Vandermonde* matrix in the absence of noise if the antenna is ULA, with kernel $u_r = \exp\left\{i\frac{2\pi}{\lambda}\Delta\sin(\theta_r)\right\}$. Direction of arrival is then estimated through a Least Squares (LS) solution or the more robust Total Least Squares (TLS) solution.

2. Signal $w_r(t_m)$ of the r-th source is extracted up to a scaling factor directly from matrix $\boldsymbol{B}$:

$$\hat{w}_r(t_m) \propto B_{mr}, \qquad m = 1, ..., M$$

3. Complex multiplicative coefficient $\alpha_r(f_k)$ is extracted directly from the factor $\boldsymbol{C}$:

$$\hat{\alpha}_r(f_k) = C_{kr}$$

5. Numerical simulations

Signals were simulated according to sampling conditions, which are typically those selected for Argentière experiments (cf. [21, 22]) and emulate the normal seismic activity of a glacier. Estimation efficiency was evaluated in comparison with two other narrowband algorithms, ESPRIT [5] and MUSIC [3, 23], and with deterministic CRB as a benchmark [17]. The performance criterion is the *total mean square error* (total MSE) of the DoA: $\frac{1}{N}\sum_{n=1}^{N}\sum_{r=1}^{R}(\hat{\theta}_{rn} - \theta_r)^2$, where $\hat{\theta}_{rn}$ is the estimated DoA of source r at the n-th Monte-Carlo trial, N is the number of trials. The number of simultaneous sources was chosen to be $R = 2$. The SNR definition for P wave was chosen as:

$$SNR = 10log_{10}\frac{\| \boldsymbol{p} \|_F^2}{LM\sigma_n^2}$$

where $\| \boldsymbol{p} \|_F^2$ refers to the P wave energy. This is consistent with previous works, where the S wave data is traditionally thrown away even if received. We have:

1. $v_1 = 3600ms^{-1}$, $v_2 = 1610ms^{-1}$.
2. $\frac{v_1}{v_2} = \frac{f_1}{f_2} = 2.236$.
3. $f_1 = 193Hz$, $f_2 = 86Hz$.
4. Time duration of simulated records: $201ms$.
5. Configuration of the array: ULA with inter-sensor distance $\Delta = 10m$.

A simple source was simulated as approximately narrowband: a signal carried by a sinusoid modulated by a Kaiser window complex envelope $a_r(t_m)\exp\{i\phi_r(t_m)\}$

$$w_{rk}(t_m) = a_r(t_m)\exp\{i\phi_r(t_m)\}\exp\{i\omega_k t_m\}$$

Since narrowband ω_k is known for P and S waves, simulated signal $w_{rk}(t_m)$ is brought to the baseband through demodulation:

$$w_r(t_m) = a_r(t_m)\exp\{i\phi_r(t_m)\}$$

Monte Carlo simulations show a superior performance of CP deterministic decomposition with respect to MUSIC and ESPRIT algorithm, especially for a low number of sensors composing the array. When the array is composed by a large number of sensors, performances of MUSIC and CP decomposition become comparable. In particular, Figures 1 and 2 illustrates computer simulations for a variable number of sensors ($L \in \{3, 10, 50\}$) composing the array and for different configurations of sources impinging from a broadside (perpendicular to the array) or endfire (laterally to the array) perspective. Figure 3 represents the evolution of RMSE with respect to angular separation of the two simultaneous sources, for a given SNR ($SNR \in \{10, 20dB\}$): the tendency is an evident predominance in the performance of the CP decomposition algorithm, especially for closely spaced sources.

The performance gap between MUSIC RMSE and its CRB reference is due to the short length of useful signal effectively present in data records: this causes the statistical estimate of

signal covariance matrix to deteriorate progressively. A comparable difference can be noticed between the performance of the proposed tensorial method and its CRB. This loss can be explained by the 2-step nature of the estimation process, which consists on a CP decomposition followed by a TLS performed on the estimated steering matrix (cf. Paragraph 4.2).

6. Results on real data

6.1. Windowing of P and S recorded signals

Since recorded P and S waves are consecutive and distinct in time, they need to be selected and cropped, so as to form two $M \times L$ slices within the data sensor, where M is the number of data samples, and L the number of sensors. Whenever P and S waves are not aligned, the broadband general formulation provides for an additive model of a deterministic component and a stochastic component:

$$\begin{cases} y^{(P)}(t - t_l^{(P)}) = w_P(t - t_l^{(P)}) + n_l(t - t_l^{(P)}) \\ y^{(S)}(t - t_l^{(S)} + \xi_{PS}) = w_S(t - t_l^{(S)} + \xi_{PS}) + n_l(t - t_l^{(S)} + \xi_{PS}) \end{cases}$$

where $l = 1, ..., L$, and ξ_{PS} is the alignment error between P and S slices.

The narrowband approximation (**A**2) at a given ostensible wavelength $\lambda = \frac{v_1}{f_1} = \frac{v_2}{f_2}$ and property **P**3 gives

$$\begin{cases} y^{(P)}(t - t_l^{(P)}) \simeq \alpha_P\, w(t) \exp\{-i\psi_l\} + n_l(t - t_l^{(P)}) \\ y^{(S)}(t - t_l^{(S)} + \xi_{PS}) \simeq \alpha_S\, w(t + \xi_{PS}) \exp\{-i\psi_l\} + n_l(t - t_l^{(S)} + \xi_{PS}) \end{cases}$$

where $\psi_l = \frac{2\pi}{\lambda}\Delta(l-1)sin\theta$.

Alignment is pursued through cross-correlation between P and S narrowband complex envelopes after the detection process, which is theoretically justified provided signals are jointly stationary and decorrelated to noise $n_l(t)$:

$$\begin{aligned} R_l^{PS}(\tau, \xi_{PS}) &= R_l^{PS}(\tau + \xi_{PS}) = \\ &= \mathbb{E}\{y^{(P)}(t - t_l^{(P)}) y^{(S)}(t - t_l^{(S)} + \tau + \xi_{PS})\} = \\ &= \alpha_P\, \alpha_S\ \exp\{-i2\psi_l\} R_l^{ww}(\tau + \xi_{PS}) \end{aligned}$$

If the alignment is fulfilled on the reference sensor ($l = 1$),

$$R_1^{PS}(\tau + \xi_{PS}) = \alpha_P\, \alpha_S\, R_1^{ww}(\tau + \xi_{PS})$$

which attains its maximum for $\tau_{max} = -\xi_{PS}$. Once τ_{max} was determined, the best alignment can be performed to the P and S waves.

6.2. A dataset of 26 events from the same cluster at Argentière glacier

Argentière glacier is a $10km$ long glacier located in the French side of Mont Blanc massif, covering a $19km^2$ surface. It is characterized by high seismic activity, as stated in [24, 25]. Temperate Alpine glaciers are characterized by ice remaining at melting point, deforming by three main mechanisms: plastic deformation of ice, plastic deformation of underlying bedrock, and basal sliding friction of the former upon the latter. The underneath flow of water plays an important role in the third phenomenon, because it modifies pressure and temperature parameters. Furthermore, brittle behavior of ice is the reason of the sudden openings of crevasses and falls of seracs on the surface of a glacier, with indirect effects on glacial hydrology by means of water transfer from the surface to greater depth zones. An array composed of 9 velocimeters pointing in the direction perpendicular to glacier motion at $1kHz$ sampling frequency was placed on the glacier surface at $2400m$ above sea level. The sensors within the array were spaced out at $10m$ intervals. The general case involves R far field sources impinging on the array, at a large unknown distance [21], in the presence of noise, thus allowing us to use the plane wave approximation **A**1 of the wavefields impinging on the array. Signals resulting of seismic events within the glacier are assumed zero-mean, non stationary and broadband stochastic processes over the observation time, and are assumed to be uncorrelated to environmental noise. The DoA technique evaluated via simulated signals was applied to the 26 deep events recorded by the array during November 2012. These events were associated with the same cluster of deep events [21], on the basis of their waveform similarities through cross-correlation. We use values for P and S wave velocities through ice, as in [26, 21]: $v_1 = 3600\ m\ s^{-1}$ and $v_2 = 1610\ m\ s^{-1}$, according to Property **P**1. An example of one deep event recorded by the 9 sensors with P and S waves is provided in Figure 6.

The frequency optimization was achieved throughout the dataset from the same cluster of events (see Figure 4 for results), after preprocessing of recorded signals (see Appendix D for details). Median dominant frequency $\bar{f}_1$ of P wave is then calculated. Thus, all the P waves from the cluster are filtered around $\bar{f}_1 = 203Hz$ and all the S waves are filtered around $\bar{f}_2 = \frac{v_2}{v_1}\bar{f}_1 = 100Hz$ before tensor decomposition. This result is coherent with spectral analysis of P and S waves. DoA estimation is performed over the dataset with different methods: MUSIC and ESPRIT over P waves, and joint tensor decomposition with speed diversity. Results are shown in Figure 5: the dispersion of the DOA estimates through the joint CP decomposition of P and S waves is smaller than that of MUSIC and ESPRIT for P waves separately.

7. Conclusion and perspectives

Throughout this paper, we developed a tensor decomposition model for seismic data exploiting propagation speed diversity of P and S waves. A physical model was traced, followed by simulations and statistical comparisons with ESPRIT and MUSIC, and theoretical CRBs. Our approach was also tested on real data of seismic activity of an alpine glacier, and different techniques were compared in terms of localization efficacy. The strength of our method lies in the integration of the double information conveyed by P and S waves of distant events impinging in succession on the array: to the traditional dimensions of array processing (recording samples for time and sensor locations for space), we added the dual content transmitted by the P and S waves, temporally distinct by virtue of the diversity of

propagation speed. The effect of adding a way to the data array is evident in terms of estimation performances (lower CRBs) and localization precision of a cluster of real events originating from the same source, especially for short data durations and for arrays composed by a small number of sensors.

An alternative to the exact tensor decomposition is represented by tensor approximation minimizing R-rank approximation error Υ, as in [17]:

$$\Upsilon(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}; \boldsymbol{\Lambda}) = \left\| z - \sum_{r=1}^{R} \varsigma_r \, \boldsymbol{a}_r \otimes \boldsymbol{b}_r \otimes \boldsymbol{c}_r \right\|^2 = \Upsilon(\boldsymbol{x}; \boldsymbol{\Lambda}) \quad (15)$$

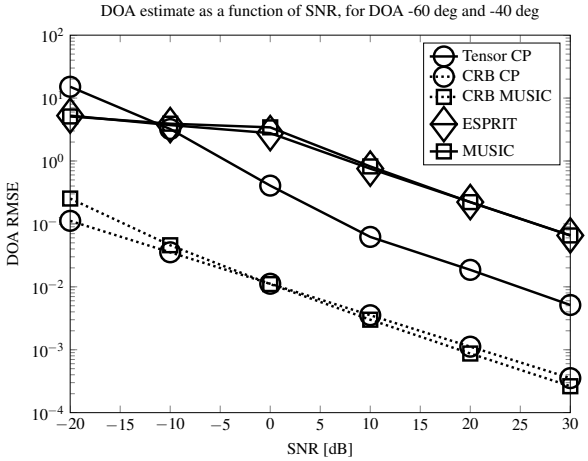
with $z = \text{vec}\{\mathcal{Z}\}$ and $\boldsymbol{x} = \text{vec}\{[\boldsymbol{A}^T, \boldsymbol{B}^T, \boldsymbol{C}^T]\}$. An iterative minimization problem (*e.g.* conjugate gradient descent method) in order to find $(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}})$ could be initialized with the exact decomposition in [19], and is stated as:

$$(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}) = \arg \min_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}} (\Upsilon(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}; \boldsymbol{\Lambda})) = \arg \min_{\boldsymbol{x}} (\Upsilon(\boldsymbol{x}; \boldsymbol{\Lambda}))$$

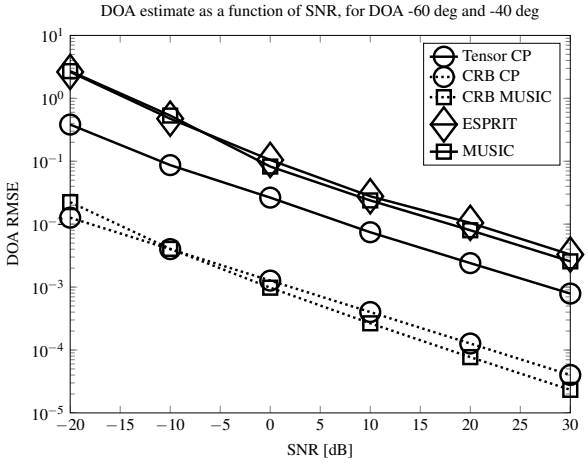
An alternative to the two-stage derivation of DoA estimates (tensor CP decomposition followed by an estimation of $\boldsymbol{\theta}$ from matrix $\boldsymbol{A}(\boldsymbol{\theta})$) lies in a direct parametrization of the objective function in (15) with DoAs with $\boldsymbol{x} = \text{vec}\{[\boldsymbol{A}^T(\boldsymbol{\theta}), \boldsymbol{B}^T, \boldsymbol{C}^T]\}$:

$$(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}) = \arg \min_{\boldsymbol{\theta},\boldsymbol{B},\boldsymbol{C}} (\Upsilon(\boldsymbol{A}(\boldsymbol{\theta}), \boldsymbol{B}, \boldsymbol{C}; \boldsymbol{\Lambda})) = \arg \min_{\boldsymbol{x}} (\Upsilon(\boldsymbol{x}(\boldsymbol{\theta}); \boldsymbol{\Lambda}))$$

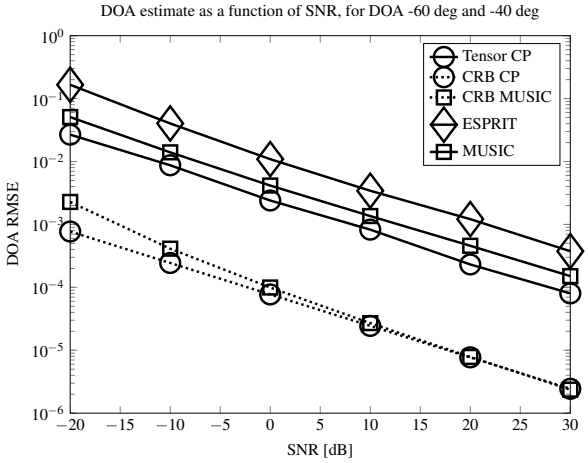
This would allow performances to get closer to the CR bounds. Further research is expected upon deterministic models of repetition and polarization diversities, in the context of array processing. Calibration error concerning the configuration of sensors in the array may possibly be included into the model, with consequences upon the statistical performances of the proposed method.



(a) endfire perspective - array of $L = 3$ sensors

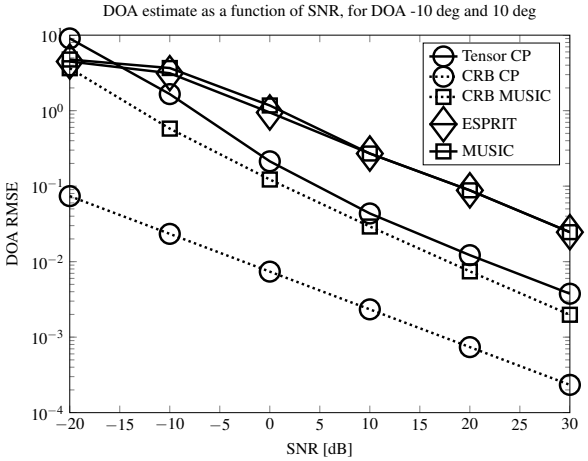


(b) endfire perspective - array of $L = 10$ sensors

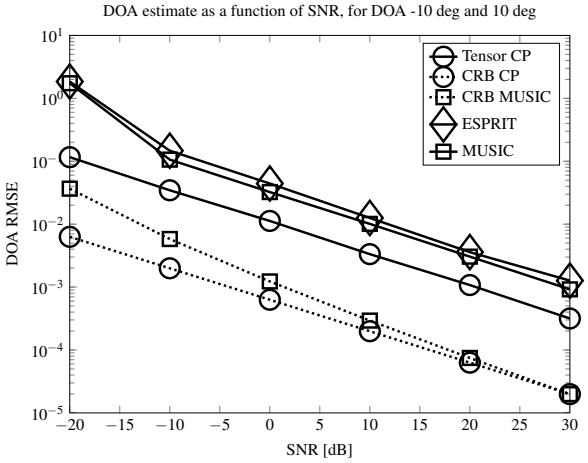


(c) endfire perspective - array of $L = 50$ sensors

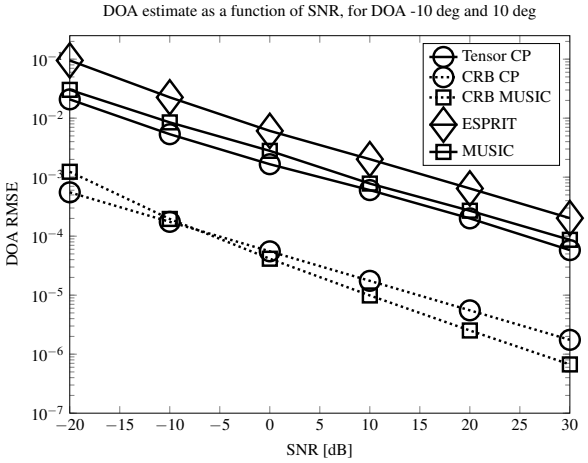
Figure 1: MSE vs SNR



(a) broadside perspective - array of $L = 3$ sensors

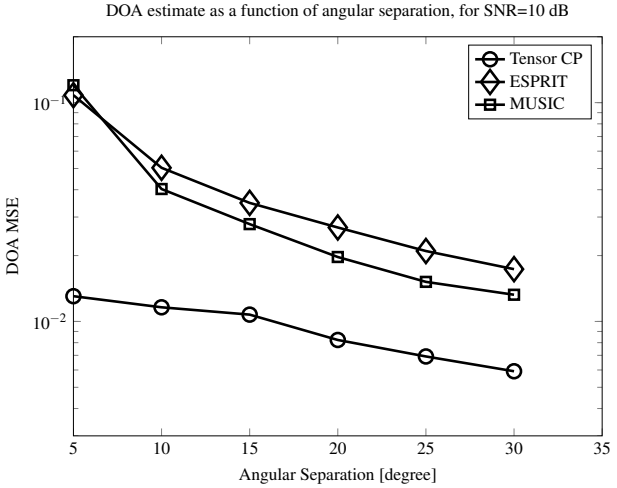


(b) broadside perspective - array of $L = 10$ sensors

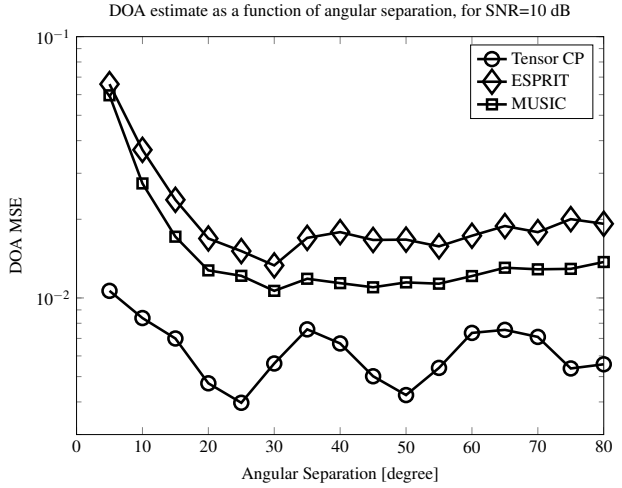


(c) broadside perspective - array of $L = 50$ sensors

Figure 2: MSE vs SNR



(a) MSE vs angular separation - endfire perspective - 10dB around 33°



(b) MSE vs angular separation - broadside perspective - 10dB around 0°

Figure 3: MSE vs angular separation

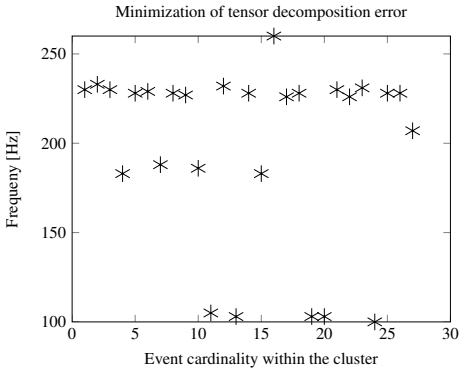


Figure 4: Optimized working frequency for P wave

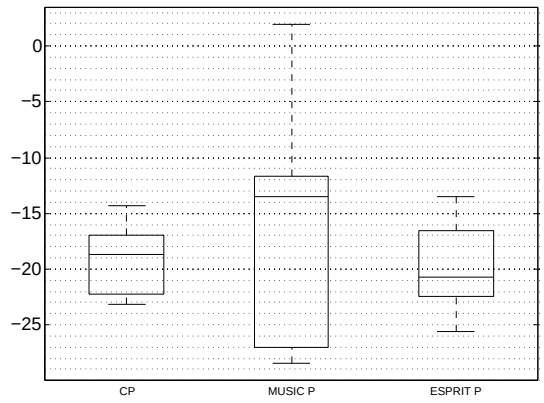
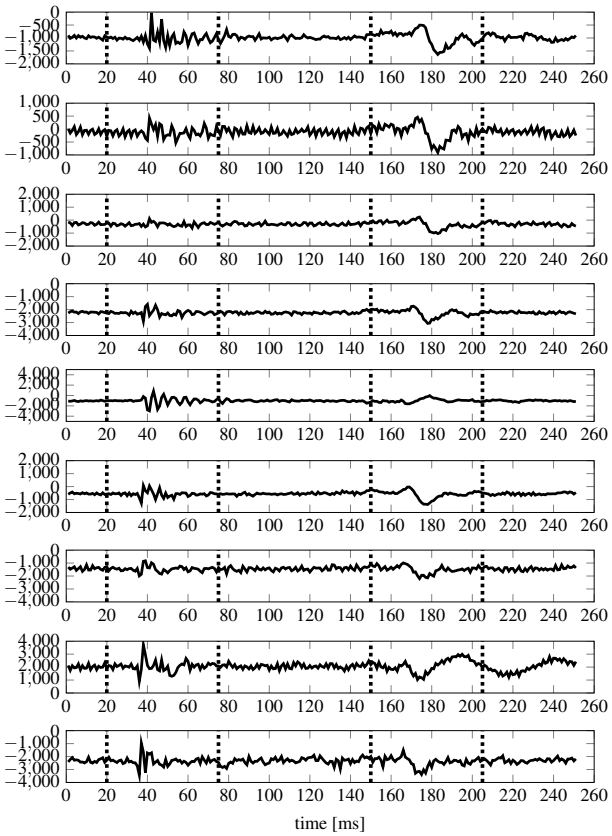


Figure 5: DoA estimation throughout the cluster of deep events



(a) Record from the array with distinct P and S waves

Figure 6: Example of an event from Argentière glacier - 2012

Appendix A. Justification of P3

The resulting signal propagating from a given source r and impinging on a given sensor of the array is composed of two contributions before addition to background and instrumentation noise: the P phenomenon, generally followed by the S phenomenon. An additional assumption **A**5 identifies the impulse responses of particle motion as the only source of variability between received P and S waves:

$$\begin{aligned} w^1(t) &\approx h^P(t) * w(t)m(t) \\ w^2(t) &\approx h^S(t) * w(t)m(t) \end{aligned}$$

where $h^P(t)$ is the impulse response of the given medium to the P wave and $h^S(t)$ to the S wave. In particular, notice that the source signal can be decomposed into a product of a low-frequency part (the complex envelope $w(t)$) and a high-frequency contribution $m(t)$:

$$w^i(t) \approx h^i(t) * w(t)m(t)$$

This corresponds to the base-band complex signal

$$w_b^i(t) \approx h_b^i(t) * w(t)M_+ \exp\{-j\,2\pi\Delta f_i\, t\}$$

where $\exp\{-j\,2\pi\Delta f_i\, t\}$ is a frequency shift due to different demodulation for P and S waves from working frequency f_1 and f_2 respectively, and M_+ is the complex amplitude of m(t). In the frequency domain, we have then

$$W_b^i(f) = \mathcal{F}\{w_b^i(t)\}(f) = H_b^i(f)M_+ \left(W(f) * \delta_{\Delta f_i}(f)\right)$$

For a generally low-pass transfer function, $H^i(f)$ can be considered as a constant H_b^i,

$$W_b^i(f) = H_b^i M_+ W(f - \Delta_{f_i})$$

Since band-pass filtering around frequency f_i and a joint base-band translation are equivalent to a base-band translation from frequency f_i followed by a low-pass filtering with window $\Pi(f)$, we have

$$W_{bf}^i(f) \approx \Pi(f)\, H_b^i\, M_+ W(f - \Delta_{f_i})$$

Then we have in time domain, for ideal low-pass filtering $\Pi(f) \approx \Pi$ in the support of interest

$$w_{bf}^i(t) \approx \left(\Pi\, H_b^i\, M_+\right)\, w(t)$$

provided that $W(f - \Delta_{f_1}) \approx W(f - \Delta_{f_2}) \approx W(f)$.

Finally, resulting P wave and S wave complex envelopes, after base-band translation and filtering, are related by a proportionality relationship

$$w_{bf}^2(t) \approx \frac{H_b^2}{H_b^1}\, w_{bf}^1(t)$$

Appendix B. Details of MUSIC and ESPRIT algorithms

Appendix B.1. MUSIC

The analysis of the $L \times L$ signal covariance matrix $\boldsymbol{R}_{YY}$ is the foundation of the MUSIC algorithm [3]:

$$\boldsymbol{R}_{YY} = \mathbb{E}\{\boldsymbol{Y}\boldsymbol{Y}^H\} = \boldsymbol{A}\mathbb{E}\{\boldsymbol{S}\boldsymbol{S}^H\}\boldsymbol{A}^H + \mathbb{E}\{\boldsymbol{N}\boldsymbol{N}^H\}$$

Under the assumptions **H**2, **H**3 and **H**4,

$$\boldsymbol{R}_{YY} = \boldsymbol{A}\boldsymbol{R}_{SS}\boldsymbol{A}^H + \sigma_n^2\boldsymbol{I}$$

Since $\mathrm{rank}\{\boldsymbol{A}\boldsymbol{R}_{SS}\boldsymbol{A}^H\} = R$, the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_R > \lambda_L$ of $\boldsymbol{R}_{YY}$ are

$$\begin{cases} \lambda_i > \sigma_n^2 & \text{for } i = 1, \ldots, R \\ \lambda_i = \sigma_n^2 & \text{for } i = R+1, \ldots, L \end{cases}$$

If the noise spatial covariance is not proportional to identity, but equal to $\sigma_n^2\boldsymbol{G}$ where $\boldsymbol{G}$ is known up to the scaling factor σ_n^2, the same reasoning applies where λ_i are generalized eigenvalues of the matrix pencil $(\boldsymbol{R}_{YY}, \boldsymbol{G})$. Let $\boldsymbol{E}_S$ denote the $L \times R$ matrix whose columns are the R first (generalized) eigenvectors of $\boldsymbol{R}_{YY}$, and $\boldsymbol{E}_N$ the $L \times (L-R)$ matrix containing the $(L-R)$ noise eigenvectors: $\boldsymbol{E}_S = [\boldsymbol{e}_1|\cdots|\boldsymbol{e}_R]$ and $\boldsymbol{E}_N = [\boldsymbol{e}_{R+1}|\cdots|\boldsymbol{e}_L]$. Thus, the squared Euclidean distance from a vector $\boldsymbol{x}$ to the signal subspace $\mathrm{Span}\{\boldsymbol{E}_S\} = \mathrm{Span}\{\boldsymbol{A}\}$ is $d^2 = \boldsymbol{x}^H\boldsymbol{E}_N\boldsymbol{E}_N^T\boldsymbol{x}$. The MUSIC algorithm aims at finding the values of angle of arrival θ for which the distance $\boldsymbol{a}(\theta)^H\boldsymbol{E}_N\boldsymbol{E}_N^T\boldsymbol{a}(\theta)$ between the array manifold $\mathcal{A}_\theta$ and the signal subspace is minimized. For this purpose, an exhaustive search is performed on a grid $\{\theta_i, 1 \leq i \leq \Theta\}$ of arbitrary precision. This is equivalent to maximizing the so-called MUSIC "spectrum":

$$P_{MU}(\theta) = \frac{\boldsymbol{a}(\theta)^H\boldsymbol{a}(\theta)}{\boldsymbol{a}(\theta)^H\boldsymbol{E}_N\boldsymbol{E}_n^T\boldsymbol{a}(\theta)}$$

whose numerator is a normalization factor. Measurements in the presence of signal $\boldsymbol{S}$ are used to estimate $\boldsymbol{R}_{YY}$:

$$\hat{\boldsymbol{R}}_{YY} = \frac{1}{M}\sum_{m=1}^{M}\boldsymbol{y}(t_m)\boldsymbol{y}^H(t_m) = \frac{1}{M}\boldsymbol{Y}\boldsymbol{Y}^H$$

If noise is absent, parameter estimates are given by the intersections of the array manifold $\mathcal{A}_\theta$ and the signal subspace $\mathrm{Span}\{\boldsymbol{E}_S\}$. In the presence of noise, there are no intersections with probability one [5]: parameters estimates are thus given by the R largest peaks of the MUSIC spectrum.

Under assumptions **H**3, **H**4 and **H**5, the generic CRB for angles of arrival $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_R]$ and noise covariance $\sigma_n^2\boldsymbol{I}$ is given by [23], if noise is circular Gaussian:

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{\sigma_n^2}{2}\left\{\sum_{m=1}^{M}\Re\left[\boldsymbol{S}^H(t_m)\boldsymbol{D}^H[\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H]\boldsymbol{D}\boldsymbol{S}(t_m)\right]\right\}^{-1}, \tag{B.1}$$

where

$$\boldsymbol{S}(t_m) = \begin{bmatrix} w_1(t_m) & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & w_R(t_m) \end{bmatrix} \in \mathbb{C}^{R\times R}$$

$$\boldsymbol{D} = [d(\theta_1), \cdots, d(\theta_R)] \in \mathbb{C}^{L\times R}$$

whose r-th element is expressed as the derivative of steering vectors with respect to directional elements

$$d(\theta_r) = \frac{d\boldsymbol{a}(\theta_r)}{d\theta_r} \in \mathbb{C}^{L\times 1}$$

Appendix B.2. ESPRIT

The ESPRIT algorithm reduces the computational complexity, by exploiting the structure of the sensor array when available [5]: the array is composed by two identical known subarrays (here denoted with letters x and y) displaced from each other by a (possibly unknown) displacement vector $\boldsymbol{\delta}$. Around frequency f_0, the observation model is

$$\begin{cases} \boldsymbol{x}(t_m) = \boldsymbol{A}\boldsymbol{s}(t_m) + \boldsymbol{n}_x(t_m), & m = 1, \ldots, M \\ \boldsymbol{y}(t_m) = \boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{s}(t_m) + \boldsymbol{n}_y(t_m), & m = 1, \ldots, M \end{cases}$$

where, in the ULA configuration, $\boldsymbol{\Phi} = \text{Diag}\{e^{j2\pi f_0 \delta sin\theta_1/c}, \cdots, e^{j2\pi f_0 \delta sin\theta_R/c}\} \in \mathbb{C}^{R\times R}$ is a unitary operator that relates the measurements from subarray $\boldsymbol{X} \in \mathbb{C}^{(L-1)\times M}$ to those from subarray $\boldsymbol{Y} \in \mathbb{C}^{(L-1)\times M}$. The joint output vector is then defined as

$$\begin{aligned} \boldsymbol{z}(t_m) &= \begin{bmatrix} \boldsymbol{x}(t_m) \\ \boldsymbol{y}(t_m) \end{bmatrix} = \bar{\boldsymbol{A}}\boldsymbol{s}(t_m) + \boldsymbol{n}_z(t_m) \\ \bar{\boldsymbol{A}} &= \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{\Phi}\boldsymbol{A} \end{bmatrix}, \qquad \boldsymbol{n}_z(t_m) = \begin{bmatrix} \boldsymbol{n}_x(t_m) \\ \boldsymbol{n}_y(t_m) \end{bmatrix} \end{aligned} \tag{B.2}$$

The signal subspace $\text{Span}\{\boldsymbol{E}_S\}$ is obtained by the first R eigenvectors of the covariance of the measurements $\boldsymbol{R}_{ZZ} = \bar{\boldsymbol{A}}\boldsymbol{R}_{SS}\bar{\boldsymbol{A}}^T + \sigma^2\boldsymbol{I}$. From equation (B.2) $\boldsymbol{E}_S$ can be further decomposed in the couple

$$\boldsymbol{E}_S = \begin{bmatrix} \boldsymbol{E}_X \\ \boldsymbol{E}_Y \end{bmatrix}$$

The solution to the problem $\boldsymbol{\Psi}\boldsymbol{E}_X = \boldsymbol{E}_Y$ is given by a Least Squares (LS) solution

$$\hat{\boldsymbol{\Psi}} = \boldsymbol{E}_X^{\dagger}\boldsymbol{E}_Y$$

or by Total Least Squares (TLS) solution:

$$[\boldsymbol{E}_X + \boldsymbol{Q}_X]\hat{\boldsymbol{\Psi}} = [\boldsymbol{E}_Y + \boldsymbol{Q}_Y]$$

where residual matrices $\boldsymbol{Q}_X$ and $\boldsymbol{Q}_Y$ have minimum Frobenius norm and the eigenvalues of $\boldsymbol{\Psi}$ are equal to the diagonal elements of $\boldsymbol{\Phi}$.

Appendix C. Cramér-Rao Bound of the DoA estimation

The parameter vector of the general model of R sources impinging on the array from angles of arrival $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_R]$ is expressed by

$$\begin{aligned} \boldsymbol{\vartheta} &= [\theta_1, \cdots, \theta_R, \bar{\boldsymbol{b}}_1^T, \cdots, \bar{\boldsymbol{b}}_R^T, \bar{\boldsymbol{c}}_1^T, \cdots, \bar{\boldsymbol{c}}_R^T, \bar{\boldsymbol{b}}_1^H, \cdots, \bar{\boldsymbol{c}}_R^H] \\ &= [\boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\xi}^*] \end{aligned}$$

with $\boldsymbol{\xi} = [\bar{\boldsymbol{b}}_1^T, \cdots, \bar{\boldsymbol{b}}_R^T, \bar{\boldsymbol{c}}_1^T, \cdots, \bar{\boldsymbol{c}}_R^T]$. In particular, if the noise is circularly complex, the parameter vector and the FIM simplify to [17]:

$$\boldsymbol{\vartheta} = [\boldsymbol{\theta}, \boldsymbol{\xi}]$$

and

$$\boldsymbol{\Phi} = \frac{1}{\sigma_n^2}\begin{pmatrix} 2\Re\{\boldsymbol{G}_{11}\} & \boldsymbol{G}_{12} & \boldsymbol{G}_{12}^* \\ \boldsymbol{G}_{12}^H & \boldsymbol{G}_{22} & \boldsymbol{0} \\ \boldsymbol{G}_{12}^T & \boldsymbol{0} & \boldsymbol{G}_{22}^* \end{pmatrix}$$

where $\boldsymbol{G}_{ij} = \left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\vartheta}_i}\right)^H \left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\vartheta}_i}\right)$, $(i, j) \in \{1, 2\} \times \{1, 2\}$, $\boldsymbol{\vartheta}_1 = \boldsymbol{\theta}$ and $\boldsymbol{\vartheta}_1 = \boldsymbol{\xi}$. Using the chain rule and complex derivative formula (cf. [17]),

$$\begin{aligned} \frac{\partial\boldsymbol{\mu}}{\partial\theta_r} &= \left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{a}_r^T}\right)\left(\frac{\partial\boldsymbol{a}_r^T}{\partial\theta_r}\right) = \\ &= i\frac{2\pi}{\lambda}cos\theta_r(\boldsymbol{I}_L \otimes \boldsymbol{b}_r \otimes \boldsymbol{c}_r)(\boldsymbol{a}_r \boxdot \boldsymbol{v}_L) \equiv \boldsymbol{\phi}_{\theta_r} \end{aligned}$$

where $\boldsymbol{v}_L = [0, 1, \cdots, L-1]$. Thus,

$$\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}} = [\boldsymbol{\phi}_{\theta_1}, \cdots, \boldsymbol{\phi}_{\theta_R}] \in \mathbb{C}^{lmk\times R}$$

As for nuisance cross-terms of the FIM,

$$\frac{\partial\boldsymbol{\mu}}{\partial\bar{\boldsymbol{b}}_r^T} = (\boldsymbol{a}_r \otimes \boldsymbol{I}_{MK})(\boldsymbol{I}_M \otimes \boldsymbol{c}_r)\boldsymbol{J}_M \equiv \boldsymbol{\phi}_{\bar{\boldsymbol{b}}_r} \in \mathbb{C}^{lmk\times(M-1)}$$

$$\frac{\partial\boldsymbol{\mu}}{\partial\bar{\boldsymbol{c}}_r^T} = \boldsymbol{a}_r \otimes \boldsymbol{b}_r \otimes \boldsymbol{I}_K \equiv \boldsymbol{\phi}_{\bar{\boldsymbol{c}}_r} \in \mathbb{C}^{lmk\times K}$$

with selection matrix $\boldsymbol{J}_M = [\boldsymbol{0}_{(M-1),1}\boldsymbol{I}_{(M-1)}]$. Thus,

$$\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\xi}} = [\boldsymbol{\phi}_{\bar{\boldsymbol{b}}_1}, \cdots, \boldsymbol{\phi}_{\bar{\boldsymbol{b}}_R}, \boldsymbol{\phi}_{\bar{\boldsymbol{c}}_1}, \cdots, \boldsymbol{\phi}_{\bar{\boldsymbol{c}}_R}] \in \mathbb{C}^{lmk\times R(K+M-1)}$$

Once the FIM is calculated, it can be inverted, so that the first leading $R \times R$ block in matrix $\boldsymbol{\Phi}^{-1}$ corresponds to the CRB of unbiased DoA estimators.

Appendix D. Preprocessing of seismic data

A time series $x(t)$ can be considered a *pass-band* signal if its spectral support is limited and it does not include the origin:

$$[-f_0 - W, -f_0 + W] \cup [f_0 - W, f_0 + W], \text{with } \infty > f_0 > W > 0$$

Every *pass-band* signal can be associated with an *analytic* signal $\hat{x}(t)$ whose support does not contain positive frequencies:

$$\hat{X}(f) = \sqrt{(2)}U_+(f)X(f) \longleftrightarrow \hat{x}(t) = \frac{1}{2}[x(t) + iH\{x(t)\}]$$

where $U_+(\cdot)$ indicates the Heaviside step function, and $H\{x(t)\}$ refers to the Hilbert transform of real signal $x(t)$.

Since a real record is characterized by even spectral symmetry, it can be represented by its analytical signal without loss of information.

The complex envelope of $X(f)$ around frequency f_0 is obtained from the *base-band* analytical signal by a mere translation in frequency:

$$\tilde{X}(f) = \hat{X}(f + f_0) \longleftrightarrow \tilde{x}(t) = \hat{x}(t)e^{-i2\pi f_0 t}$$

For a given carrier at frequency f_0, a complex envelope around f_0 is in bijection with a complex number representing the modulus and the phase of the carrier.

Thus, recorded data have to be filtered in order to satisfy the narrowband assumption **A**2.

Moreover, an approximation has to be made with respect to the spectral content of signals, in case the ratio of the dominant frequencies does not reflect the ratio of the velocities of P and S waves, according to property **P**2. In order to maintain a constant wavelength λ, which is a necessary condition for separability of the multilinear model, a compromise needs to be reached in extracting the dominant frequencies, as a consequence of the ideal condition $\lambda = \frac{v_1}{f_1} = \frac{v_2}{f_2}$ in property **P**2. Within the set of four parameters determining wavelength, $\{v_1, v_2, f_1, f_2\}$ one only needs 3 degrees of freedom to determine optimal working conditions. P and S wave propagation velocities v_1 and v_2 are given by geophysical active analysis of known reflecting waves from the surface to the glacier bed [21], [27]. The only parameter subject to optimization is then P or S wave frequency f_1 or f_2. If f_1 is chosen as the free parameter and optimized, then f_2 directly derives from property **P**2

$$f_2 = \frac{C_2}{C_1} f_1 = \frac{C_2}{\lambda} \tag{D.1}$$

Optimization is fulfilled by minimizing the Frobenius norm of tensor decomposition error:

$$\hat{f_1} = \arg\min_{f_1} \left\| \boldsymbol{z}(f_1) - \sum_{r=1}^{R} \lambda_r \boldsymbol{a}_r(f_1) \otimes \boldsymbol{b}_r \otimes \boldsymbol{c}_r(f_1) \right\|_F^2$$

where $\boldsymbol{z}(f_1)$ is the data array after narrowband filtering around f_1 and $f_1 \in (0Hz, f_S/2 = 500Hz)$. Then, from Equation (D.1), we have $\hat{f_2} = \frac{v_2}{v_1}\hat{f_1}$.

References

[1] W. Liu, S. Weiss, Wideband beamforming: concepts and techniques, Vol. 17, John Wiley & Sons, 2010.

[2] H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach, Signal Processing Magazine, IEEE 13 (4) (1996) 67–94.

[3] R. O. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antenna Propagation 34 (3) (1986) 276–280.

[4] A. Barabell, Improving the resolution performance of eigenstructure-based direction-finding algorithms, in: ICASSP'83., Vol. 8, IEEE, 1983, pp. 336–339.

[5] R. Roy, T. Kailath, Esprit-estimation of signal parameters via rotational invariance techniques, Acoustics, Speech and Signal Processing, IEEE Transactions on 37 (7) (1989) 984–995.

[6] B. D. Van Veen, K. M. Buckley, Beamforming: A versatile approach to spatial filtering, IEEE assp magazine 5 (2) (1988) 4–24.

[7] H. Wang, M. Kaveh, Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources, Acoustics, Speech and Signal Processing, IEEE Transactions on 33 (4) (1985) 823–831.

[8] N. D. Sidiropoulos, R. Bro, G. B. Giannakis, Parallel factor analysis in sensor array processing, IEEE Trans. Sig. Proc. 48 (8) (2000) 2377–2388.

[9] L.-H. Lim, P. Comon, Blind multilinear identification, IEEE Trans. Inf. Theory 60 (2) (2014) 1260–1280, open access. hal-00763275.

[10] P. Comon, Tensors: a brief introduction, in: IEEE Sig. Proc. Magazine, Vol. 31, 2014, pp. 44–53.

[11] E. Bedrosian, A product theorem for hilbert transforms, Research Memorandum, RM-3439-PR, Rand Corporation.

[12] P. J. Oonincx, Automatic phase detection in seismic data using the discrete wavelet transform, CWI Technical Report PNA-R 9811 (1998). URL `http://oai.cwi.nl/oai/asset/4624/04624D.pdf`

[13] R. E. Sheriff, L. P. Geldart, Exploration seismology. Volume 1: History, theory, and data acquisition, Cambridge University Press, New York, NY, 1983.

[14] X. Guo, S. Miron, D. Brie, S. Zhu, X. Liao, A Candecomp/Parafac perspective on uniqueness of doa estimation using a vector sensor array, IEEE Trans. Signal Processing 59 (7) (2011) 3475–3481.

[15] M. Niknazar, H. Becker, B. Rivet, C. Jutten, P. Comon, Blind source separation of underdetermined mixtures of event-related sources, Signal Processing 101 (2014) 52–64, hal-00952039.

[16] L. Chiantini, G. Ottaviani, N. Vannieuwenhoven, An algorithm for generic and low-rank specific identifiability of complex tensors, SIAM J. matrix Ana. Appl. 35 (4) (2014) 1265–1287.

[17] S. Sahnoun, P. Comon, Tensor polyadic decomposition for antenna array processing, in: 21st Int. Conf. Comput. Stat (CompStat), Geneva, 2014, pp. 233–240, hal-00986973.

[18] X. Liu, N. D. Sidiropoulos, Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays, Signal Processing, IEEE Transactions on 49 (9) (2001) 2074–2086.

[19] S. Leurgans, R. Ross, R. Abel, A decomposition for three-way arrays, SIAM Journal on Matrix Analysis and Applications 14 (4) (1993) 1064–1083.

[20] J. B. Kruskal, Three-way arrays: Rank and uniqueness of trilinear decompositions, Linear Algebra and Applications 18 (1977) 95–138.

[21] P. C. A. Helmstetter, B. Nicolas, M. Gay, Basal icequakes recorded beneath an Alpine glacier (Glacier d'Argentière, Mont Blanc, France) : evidence for stick-slip motion?, submitted to J. Geophys. Res.

[22] P. C. A. Helmstetter, B. Nicolas, M. Gay, Intermediate-depth icequakes and harmonic tremor in an alpine glacier (Glacier d'Argentière, Mont Blanc, France) : evidence for hydraulic fracturing?, submitted to J. Geophys. Res.

[23] P. Stoica, N. Arye, MUSIC, maximum likelihood, and Cramér-Rao bound, Acoustics, Speech and Signal Processing, IEEE Transactions on 37 (5) (1989) 720–741.

[24] P.-F. Roux, Méthodes sismologiques pour l'étude de la fracturation dans les glaciers alpins: glaciers d'Argentière et du Gorner, Ph.D. thesis, Université de Savoie (2008).

[25] P.-F. Roux, D. Marsan, J.-P. Mtaxian, G. O'Brien, L. Moreau, Microseismic activity within a serac zone in an alpine glacier (glacier d'argentière, mont blanc, france), Journal of glaciology 54 (184) (2008) 157–168.

[26] M. Vallon, Contribution à l'étude de la Mer de Glace - Alpes françaises, Theses, Faculté des Sciences de l'Université de Grenoble (May 1967). URL `https://tel.archives-ouvertes.fr/tel-00693943`

[27] F. Walter, N. Deichmann, M. Funk, Basal icequakes during changing subglacial water pressures beneath Gornergletscher, Switzerland, Journal of Glaciology 54 (186) (2008) 511–521.

[28] F. Walter, J. F. Clinton, N. Deichmann, D. S. Dreger, S. E. Minson, M. Funk, Moment tensor inversions of icequakes on gornergletscher, switzerland, Bulletin of the Seismological Society of America 99 (2A) (2009) 852–870.

[29] F. Walter, D. S. Dreger, J. F. Clinton, N. Deichmann, M. Funk, Evidence for near-horizontal tensile faulting at the base of Gornergletscher, a Swiss Alpine glacier, Bulletin of the Seismological Society of America 100 (2) (2010) 458–472.

[30] N. Deichmann, J. Ansorge, F. Scherbaum, A. Aschwanden, F. Bernardi, G. H. Gudmundsson, Evidence for deep icequakes in an Alpine glacier, Annals of Glaciology 31 (1) (2000) 85–90.

[31] A. Gusmeroli, R. A. Clark, T. Murray, A. D. Booth, B. Kulessa, B. E. Barrett, Seismic wave attenuation in the uppermost glacier ice of storglaciaren, sweden, Journal of Glaciology 56 (196) (2010) 249–256.

[32] C. Knapp, G. C. Carter, The generalized correlation method for estimation of time delay, Acoustics, Speech and Signal Processing, IEEE Transactions on 24 (4) (1976) 320–327.

[33] S. Sahnoun, P. Comon, Deterministic blind identification in antenna array processing, in: 8th IEEE SAM Workshop, A Coruna, Spain, 2014, hal-00957357.

[34] D. Donno, A. Nehorai, U. Spagnolini, Seismic velocity and polarization estimation for wavefield separation, Signal Processing, IEEE Transactions on 56 (10) (2008) 4794–4809.

[35] J. M. ten Berge, Kruskal's polynomial for 2× 2× 2 arrays and a generalization to 2× n× n arrays, Psychometrika 56 (4) (1991) 631–636.

[36] L. De Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization, SIAM Journal on Matrix Analysis and Applications 28 (3) (2006) 642–666.

[37] R. Fletcher, Generalized inverse methods for the best least squares solution of systems of non-linear equations, The Computer Journal 10 (4) (1968) 392–399.

[38] O. M. H. Clergeot, New simple implementation of the coherent signal subspace method for wide band direction of arrival estimation, in: Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference on, IEEE, 1989, pp. 2764–2767.