Tensor decomposition of polarized seismic waves
Francesca Raimondi, Pierre Comon

To cite this version:
Francesca Raimondi, Pierre Comon. Tensor decomposition of polarized seismic waves. GRETSI 2015 - XXVème Colloque francophone de traitement du signal et des images, Sep 2015, Lyon, France. pp.4. hal-01164363

HAL Id: hal-01164363
https://hal.science/hal-01164363
Submitted on 16 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
1 Introduction

Starting from the ideas on vector sensor array developed for seismic waves in [1] from the more general model for polarized waves in [2] and [3], we state the observation model in tensor form. Next we compute the Cramér-Rao Bounds (CRB) for the joint estimation of the four parameters of polarized seismic waves. The ultimate estimation performances are then compared to the CRB as a function of the Signal to Noise Ratio (SNR). A deterministic approach based on tensor decomposition has been introduced in [4]. 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 (as opposed to traditional high resolution algorithms such as MUSIC [5] and ESPRIT [6]). CRB for the low-rank decompositions of multidimensional array was derived in [7] and extended in [8]. Polarization of waves has been first introduced in [9] as a multidimensional diversity in the tensor approach. The same authors in [10] explore the concept of polarization separation and its influence on performances.

**Notation** We shall assume throughout the following notations: matrices, column vectors and scalars are denoted respectively in bold uppercase, e.g. $A$, bold lowercase, e.g. $v$, and plain lowercase; in particular array entries are written e.g. $v_i$ or $A_{ij}$. Transposition, complex conjugation and Hermitian transposition are denoted by $\mathbf{A}^\mathsf{T}$, $(\cdot)^*$ and $(\cdot)^\mathsf{H}$, respectively. Arrays with more than two indices are referred to as tensors, with some abuse of terminology [11], and are denoted in bold calligraphic, as $\mathcal{T}$. The outer (tensor) product between two vectors is denoted by $\mathbf{u} \otimes \mathbf{v}$. Finally, $\| \cdot \|_F$ refers to the Frobenius norm; $\otimes$ denote Kronecker product. For the sake of conciseness, $a_i$ will represent the $i$-th column of matrix $A$.

2 Physical model

2.1 A 4-parameter far-field model

The physical quantity measured is the particle displacement vector recorded by a three-component particle displacement sensor (or geophone), located at a given point in space, in the direction of the $x$-, $y$- and $z$-axes of its reference system. The $z$-axis is required to be perpendicular to the earth’s surface. The following parameterization is based on the definition of four angular parameters. First, the unit vector pointing to the source is given by

$$\mathbf{u} = \begin{bmatrix} \cos \theta \cos \psi \\ \sin \theta \cos \psi \\ \sin \psi \end{bmatrix}$$

where $\theta \in (-\pi, \pi]$ refers to the azimuth and $\psi \in [-\pi/2, \pi/2]$ to the elevation of the source. Second, the polarization ellipse is described by the orientation angle $\alpha \in (-\pi/2, \pi/2]$ and the ellipticity angle $\beta \in [-\pi/4, \pi/4]$. Two models can be drawn respectively for transverse (TR)
waves and tilted generalized Rayleigh (TGR) waves [1].
For a transverse wave, the polarization ellipse lies in the
space orthogonal to the direction of propagation \( \mathbf{u} \), and is
spanned by the columns \( \{ \mathbf{v}_1, \mathbf{v}_2 \} \) of the matrix:
\[
\mathbf{V}^{TR} = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix}
-\sin \theta & -\cos \theta \sin \psi \\
\cos \theta & -\sin \theta \sin \psi \\
0 & \cos \psi
\end{bmatrix}
\] (1)
In the TGR model, the polarization ellipse is confined in the
plane spanned by the columns of \( \mathbf{V}^{TGR} \):
\[
\mathbf{V}^{TGR} = [\mathbf{u}, \mathbf{v}_2] = \begin{bmatrix}
\cos \theta \cos \psi & -\cos \theta \sin \psi \\
\sin \theta \cos \psi & -\sin \theta \sin \psi \\
\sin \psi & \cos \psi
\end{bmatrix}
\]
If the complex envelope of the source signal is denoted by
\( s(t) \), the general data model can be written as
\[
y(t) = \mathbf{V}(\theta, \psi) \mathbf{Q}(\alpha) \mathbf{w}(\beta) s(t) \in \mathbb{C}^{3 \times 1}
\] (2)
where \( \mathbf{V}(\theta, \psi) \) is one of the above matrices,
\[
\mathbf{Q} = \begin{bmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix}
\cos \beta \\
\sin \beta
\end{bmatrix}
\]
in the absence of noise, and \( \iota = \sqrt{-1} \).

2.2 Seismic waves and polarization

There exist several types of elastic waves associated with
seismic activities [1]. Primary waves (or P-Waves) are compresional elastic waves whose particle displacement
vector is parallel to the direction of propagation. For these waves, \( \alpha = \beta = 0 \), which leads to a linearly polarized
wavefront with particle motion along the direction of
propagation \( \mathbf{u} \). Rayleigh Waves are elliptically polarized
surface waves. Therefore, \( \psi = 0 \), provided that the \( xy \)-
plane corresponds to the earth surface. For these waves, it is obvious that \( \alpha = 0 \) and then \( \mathbf{Q} = \mathbf{I} \). Secondary or
Shear waves (or S-waves) are transversely elliptically polarized
in general. P-Waves and Rayleigh Waves are particular cases of the TGR model described in Section 2.1 for elliptically polarized waveforms: the direction of propagation is
located in the plane spanned by the ellipse major
and minor axes. For reasons of space, we shall concentrate
on TR waves in the remainder.

2.3 Tensor model

Now suppose that data are recorded on \( K \) polarized sen-
sors located at points in space defined by vectors \( g(k) :=
[g_k^x; g_k^y; g_k^z] \in \mathbb{R}^3, 1 \leq k \leq K \). Also suppose that \( R \) far-field narrow-band sources impinge on this vector sensor
array from direction \( \mathbf{u}(r), 1 \leq r \leq R \), and denote \( \omega \) their
common angular pulsation. We make the assumption that
impinging waves have elliptical polarization (neither linear
nor circular). Then from (2) we can assume the following
observation model in baseband about pulsation \( \omega \):
\[
\mathbf{T} = \mathbf{Z} + \mathbf{E}, \quad \mathbf{Z} = \sum_{r=1}^{R} \mathbf{a}(r) \otimes \mathbf{b}(r) \otimes \mathbf{s}(r)
\] (3)
where \( a_k(r) = \frac{1}{\lambda} \exp \{ \iota \omega g(k)^{T} \mathbf{u}(r) \} \) is the \( k \)-th entry of the steering vector, \( \lambda \) the wave celerity, \( \mathbf{b}(r) = \mathbf{V}(\theta_r, \psi_r) \mathbf{Q}(\alpha_r) \mathbf{w}(\beta_r) \in \mathbb{C}^{L \times 1} \) (\( L = 3 \)) characterizes the
propagation type, \( (\theta_r, \psi_r) \) refers to the Direction of Arrival (DoA) of the \( r \)-th source and \( (\alpha_r, \beta_r) \) its polarization,
and \( s_m(r) \) is the signal propagating from the \( r \)-th source
and received at time \( t_m \), \( 1 \leq m \leq M \). The additive noise \( \mathbf{E} \) is assumed to be i.i.d. circular Gaussian and independent of
the sources. In terms of arrays of coordinates, model (3)
rewrites:
\[
Z_{klm}(\theta, \psi, \alpha, \beta, \mathbf{s}) = \sum_{r=1}^{R} a_k(r) b_l(r) s_m(r)
\] (4)
or in column vector format:
\[
z := \mathbf{vec} \mathbf{Z} = \sum_{r=1}^{R} \mathbf{a}(r) \otimes \mathbf{b}(r) \otimes \mathbf{s}(r)
\] (5)

3 Parameter identification

3.1 Model identification

It is always possible to decompose the data tensor into
a sum of decomposable tensors \([1, 4]\) of the form \( \mathbf{D}(r) =
\mathbf{a}(r) \otimes \mathbf{b}(r) \otimes \mathbf{s}(r) \), that is, in terms of array of coordinates:
\[
D_{klm}(r) = a_k(r) b_l(r) s_m(r)
\]
Hence tensor \( \mathbf{Z} \) takes the form:
\[
\mathbf{Z} = \sum_{r=1}^{R} \xi_r \mathbf{D}(r)
\] (6)
where coefficients \( \xi_r \) can always be chosen to be real positive, and decomposable tensors \( \mathbf{D}(r) \) to have unit norm,
\textit{i.e.} for \( L^p \) norms, \( \| \mathbf{D} \| = \| \mathbf{a} \| \| \mathbf{b} \| \| \mathbf{c} \| = 1 \). The minimal value of \( R \) such that this decomposition holds is called
rank of \( \mathbf{Z} \). If \( R \) is not too large, the corresponding
decomposition is unique \([4, 12, 11, 13]\) and deserves to be referred to as Canonical Polyadic (CP); other terminologies
include rank decomposition or Candecomp/Parafac.

Note that decomposable tensors have a rank equal to 1. Because of the uniqueness of the CP decomposition, decomposable tensors of (4) and (6) coincide in the absence
of noise. This means that vectors \( \{ \mathbf{a}(r), \mathbf{b}(r), \mathbf{s}(r) \} \) coinci-
de up to some scaling factors \([4, 8, 11]\).

3.2 Model identifiability

There exist sufficient conditions ensuring uniqueness of
the exact CP, \textit{e.g.} the Kruskal condition \([4]\):
\[
\kappa_A + \kappa_B + \kappa_C \geq 2R + 2
\] (7)
where the notation \( \kappa_A \) refers to the Kruskal-rank* of matrix \( \mathbf{A} \). However, less stringent conditions guaranteeing
almost surely a unique solution can be found \([12, 11, 13]\):
\[
R(K + L - M - 2) < KL M
\]
\footnote{The Kruskal rank of a matrix \( \mathbf{A} \) is the largest number \( \kappa_A \) such that any subset of \( \kappa_A \) columns are linearly independent.}
This hold true when data are not corrupted by noise. However, if noise is present, we have to solve a best rank-R approximation problem:

$$\min_{a_r, b_r, s_r} \left\| T - \sum_{r=1}^R a_r \otimes b_r \otimes s_r \right\|_F^2$$

(8)

For $d \geq 3$, the best approximation of a $d$-partite function of a sum of $R$ product of $d$ separable functions does not exist in general [12], as a sequence of rank-$r$ functions can converge to a limit which is not rank-$r$. A sufficient condition ensuring existence of a solution to (8) via the definition of a physical constraint, the coherence, is derived in [12]. Unlike the Kusak rank, coherences are easy to compute and present the advantage of having a physical meaning, i.e. the best rank-$R$ approximation exists and is unique if either impinging signals are not too correlated, or their directions of arrivals and polarization states are not too close.

4 Performances

4.1 Mixed real-complex gradients

Since the parameter of the array processing model are complex, a definition of the derivative of a real function $h(z) \in \mathbb{R}^p$ with respect to the complex variable $z \in \mathbb{C}^n$, $z = x + iy$, $x, y \in \mathbb{R}$, needs to be introduced. [7] presents a derivation of Cramér-Rao bounds related to the CP decomposition of multidimensional arrays, using the same definition of complex derivative as in [8, 14]:

$$\frac{\partial h}{\partial z} = \frac{1}{2} \frac{\partial h}{\partial x} - i \frac{\partial h}{\partial y}$$

(9)

For clarity, our notation of the derivative of a scalar function $Y(z)$ with respect to a column vector $z \in \mathbb{C}^n$ is:

$$\frac{\partial Y}{\partial z}$$

is a column vector

$$\frac{\partial^T Y}{\partial z^T}$$

is a line vector

Given an holomorphic column vector function $f(z) \in \mathbb{C}^m$, we define $\left(\frac{\partial f}{\partial z}\right)_i = \frac{\partial f_i}{\partial z}$ so that

$$\left\{ \begin{array}{l}
\frac{\partial f}{\partial z} \\
\frac{\partial f}{\partial z^T}
\end{array} \right\}$$

is a $m \times n$ matrix

is a $n \times m$ matrix

In the sequel, we shall need a complex derivative chain rule. Given a scalar function $Y(z) \in \mathbb{R}$, a complex function $z(\theta) = x + iy \in \mathbb{C}$, and a real variable $\theta \in \mathbb{R}$, we have from the real derivative chain rule:

$$\frac{\partial Y}{\partial \theta^1} = \frac{\partial Y}{\partial x} \frac{\partial x}{\partial \theta^1} + \frac{\partial Y}{\partial y} \frac{\partial y}{\partial \theta^1}$$

which, using (9), yields the chain rule:

$$\frac{\partial Y}{\partial \theta^1} = 2R \left\{ \frac{\partial Y}{\partial z} \left\{ \frac{\partial z}{\partial \theta^1} \right\} \right\}$$

(10)

4.2 Gradient calculation

In order to compute Cramér-Rao bounds, we shall need the gradients of the log-likelihood, which turn out to be the same as those of the cost function $f$ defined below, deduced from (5), if noise is i.i.d. circular Gaussian:

$$f(\theta) = \frac{1}{\sigma_n^2} \left\| t - z(\theta) \right\|^2_2, \quad \theta := [\theta; \psi; \alpha; \beta; \text{vec} \mathbf{S}; \text{vec} \mathbf{S}^*]$$

(11)

where $\sigma_n^2$ denotes its variance. The gradient expressions will also be subsequently useful to implement descent algorithms. According to the chain rule (10) and definition (9), the partial derivatives of the cost function with respect to DoA parameters are given by

$$\frac{\partial f}{\partial \theta^r} = 2R \left\{ \frac{\partial f}{\partial a_r} \left\{ \frac{\partial a_r}{\partial \theta^r} \right\} + \frac{\partial f}{\partial b_r} \left\{ \frac{\partial b_r}{\partial \theta^r} \right\} \right\}, \quad \frac{\partial f}{\partial \theta^r} = 2R \left\{ \frac{\partial f}{\partial \theta^r} \left\{ \frac{\partial \theta^r}{\partial \theta^r} \right\} \right\}$$

with

$$\frac{\partial a_r}{\partial \theta^r} = \left( z - \sum_{r=1}^R a_r \otimes b_r \otimes s_r \right)^H (-I_K \otimes b_r \otimes s_r)$$

$$\frac{\partial b_r}{\partial \theta^r} = \left( z - \sum_{r=1}^R a_r \otimes b_r \otimes s_r \right)^H (-a_r \otimes I_K \otimes s_r)$$

$$\frac{\partial a_r}{\partial \theta^r} = \left( \frac{\omega}{v} (-g_k^\theta \sin \psi_r + g_k^\theta \cos \psi_r) A^k \right)_{k=1}^K$$

$$\frac{\partial V^R}{\partial \theta^r} = \begin{bmatrix} -\cos \theta_r & \sin \theta_r & \sin \psi_r \\
-\sin \theta_r & -\cos \theta_r & \sin \psi_r \\
0 & 0 & 0 \end{bmatrix}$$

$$\frac{\partial A_k}{\partial \theta^r} = \frac{1}{K} \exp \left\{ i \frac{\omega}{v} \left[ g_k^\theta \cos \theta_r \cos \psi_r + g_k^\theta \sin \theta_r \cos \psi_r + g_k^\theta \sin \psi_r \right] \right\}$$

$$\frac{\partial V^R}{\partial \theta^r} = \begin{bmatrix} 0 & -\cos \theta_r & \cos \psi_r \\
0 & -\sin \theta_r & \cos \psi_r \\
0 & 0 & -\sin \psi_r \end{bmatrix}$$

$$\frac{dQ_r}{d\alpha_r} = \begin{bmatrix} -\sin \alpha_r & \cos \alpha_r \\
-\cos \alpha_r & -\sin \alpha_r \end{bmatrix}, \quad \frac{d\alpha_r}{d\beta_r} = \begin{bmatrix} -\sin \beta_r \\
\cos \beta_r \end{bmatrix}$$

4.3 Cramér-Rao Bounds

Cramér-Rao Bounds (CRB) represent the lower bound on the variance of any unbiased estimator of a deterministic parameter. Define the Signal-to-Noise ratio (SNR) as [7]:

$$\text{SNR} = 10 \log_{10} \frac{\| Z \|^2_2}{KLMS_n}$$
where operator $\| \cdot \|_F^2$ indicates Frobenius norm. For a zero-mean, circularly complex Gaussian noise with covariance $\sigma_n^2 I$ the log-likelihood takes the form (11) up to an additive constant. Then, the mixed real-complex Fisher Information Matrix (FIM) can be shown to be given by [7, 8]:

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

The CRB of any unbiased estimator of a vector parameter $\theta$ is given by the inverse of the FIM. It is useful to separate parameters to be estimated in three vectors: one real, $[\theta; \psi; \alpha; \beta]$, and two complex, $\text{vec} S$ and $\text{vec} S^\ast$. With this organization, the FIM has 9 blocks [8]:

$$
\Phi = \frac{1}{\sigma_n^2} \begin{bmatrix}
2\Re \{G_{11}\} & G_{12}^\ast & G_{12}

G_{12} & G_{22} & 0

G_{12}^\ast & 0 & G_{22}^\ast
\end{bmatrix}
$$

where $G_{ij} = \left( \frac{\partial \theta_i}{\partial \theta_j} \right)^H \left( \frac{\partial \theta_i}{\partial \theta_j} \right)$.

5 Computer experiments

Signals were simulated according to realistic sampling conditions (1kHz sampling frequency, $M = 42$ time samples, $K = 9$ sensors, $L = 3$ polarization components). Ultimate performances have been evaluated by running a gradient descent initialized with the true values slightly corrupted by noise. A comparison with deterministic CRB is shown in Figure 1. The performance criterion is the total mean square error (total MSE) of each DoA and polarization parameter $\theta$: $
\frac{1}{N} \sum_{n=1}^{N} \sum_{r=1}^{R} (\hat{\theta}_{rn} - \theta_r)^2,
$ where $\hat{\theta}_{rn}$ is the estimated parameter of the $r$-th source at the $n$-th Monte-Carlo trial, $N = 99$. The number of simultaneous sources was chosen to be $R = 2$, with the following parameters:

$$
\begin{align*}
\theta_1 &= -\pi/3, \theta_2 = \pi/6, \quad \psi_1 = -\pi/4, \quad \psi_2 = \pi/7 \\
\alpha_1 &= \pi/5, \quad \alpha_2 = \pi/7, \quad \beta_1 = -\pi/7, \quad \beta_2 = \pi/5
\end{align*}
$$

**FIG. 1:** MSE vs SNR

CRBs are obtained by summing the diagonal entries in the inverse of the first block, $2\Re \{G_{11}\}$, in the FIM. This means source signals were considered as known and not as nuisances (which implies the obtention of slightly smaller bounds).

References