Seismic Velocity and Polarization Estimation for Wavefield Separation

Daniela Donno, *Student Member, IEEE*, Arye Nehorai, *Fellow, IEEE*, and Umberto Spagnolini, *Senior Member, IEEE*

***Abstract*—We address the problem of estimating the shape parameters of seismic wavefields using linear arrays of three-component (3C) vector sensors with uncertain acquisition geometry. The goal is to separate the different seismic waves, which is of practical need for oil exploration and geophysics. We present a parametric model for multiple wideband polarized signals received by an array of three-component sensors with positional and rotational calibration errors, and derive the Cramér–Rao lower bounds on the performance of the model parameters for both the exact physical model and the model with uncertain acquisition geometry. We propose a method for jointly estimating the velocity and polarization parameters based on the shift-invariance properties of multiple wavefields impinging on the linear array. We then remove the interfering surface waves by using a beamforming filter designed to exploit the velocity and polarization diversity of the different seismic waves, after clustering of the shape-parameter estimates. Examples using simulated and experimental data illustrate the applicability of the proposed methodology.**

***Index Terms*—Array signal processing, Cramér–Rao bound, velocity/polarization estimation, vector-sensor calibration errors, vector-sensor broadband beamforming.**

I. Introduction

In VARIOUS research areas, such as seismic prospecting, wireless communication, and remote sensing, the use of multicomponent sensors to characterize and exploit the polarization state of wavefields is increasing. This work is motivated by the practical need for velocity and polarization estimation for wavefield separation in oil exploration and geophysics [1].

Geophysical exploration maps the characteristics of subsurface layers from seismic wavefields registered by surface sensors. Among the many types of seismic waves registered as superimposed at the surface, only the volume waves reflected by the subsurface layers contain useful information for the subsurface imaging. However, many seismic data (especially those acquired on land) are seriously contaminated by surface waves that interfere with processing and analysis of volume waves. The purpose of this paper is to estimate and separate out spurious surface waves (technically referred as *ground-roll*) from seismic recordings to make reflected volume waves significantly easier to recognize and use for depth imaging. In this paper, we extend our earlier work in [2] by dealing also with the problem of sensor-calibration errors, which often affects real seismic surveys.

Compared to volume waves, surface waves are characterized by lower velocity and higher polarization (they are elliptically polarized with ellipticity $\varepsilon > 0.5$). Therefore, surface wave filtering methods that exploit both the velocity and polarization characteristics are expected to be effective. Conventional methods for surface-wave suppression take into account velocity (e.g., by f-k filters [1] or Radon transform [3]) or polarization [4]–[5] only. In the f-k filtering or Radon transform techniques, the separation process is accomplished by selecting an appropriate mask in the transformed domain. In [4], polarization is computed in the wavelet domain, whereas in [5] singular value decomposition (SVD) is used to estimate and separate interfering polarized wavefields. Unlike the conventional methods in [3]–[5], in this paper we propose to employ the *joint* information on velocity and polarization to extend the dimensionality of the separation space, thus improving the ability to separate surface waves from volume waves. A first attempt to separate different seismic wavefields jointly using velocity and polarization is to be found in [6].

Electromagnetic wavefield analysis is known to be somewhat simpler than elastic wavefield analysis [7]; this justifies intense investigations in this area. In fact, for electromagnetic wavefields several narrowband direction-finding and polarization estimation algorithms have been investigated. A comprehensive model for vector-sensor array processing has been proposed in [8], and direction/polarization estimation is analyzed in [9]. A maximum-likelihood estimation (MLE) of polarization parameters has been discussed in [10] and [11], whereas in [12] the authors presented a method for incorporating signal polarization in Capon's minimum variance estimator. Subspace-based methods, which are known to be computational efficient but still consistent, have been also proposed: a MUSIC-based approach is discussed in [13], and ESPRIT-based direction/polarization estimations are found in [14] for polarized signals with two-component sensors, in [15] and [16] using a sparse array of electromagnetic vector-sensors, and in [17] using a single electromagnetic vector sensor.

Since seismic wavefields are wideband in nature, we propose herein a subspace-based method for the joint estimation of velocity and polarization which holds for wideband elastic data. It is well known [18] that methods for narrowband electromag-

Manuscript received October 22, 2007; revised April 15, 2008. First published June 20, 2008; current version published September 17, 2008. This work was supported in part by the MURI under Grant FA9550-05-1-0443 and the National Science Foundation under Grants CCR-0330342 and CCF-0630734. Part of this work was presented at the IEEE International Conference on Acoustic, Speech, and Signal Processing, May 15–19, 2006, Toulouse, France. The associated editor coordinating the review of this manuscript and approving it for publication was Dr. Shahram Shahbazpanahi.

D. Donno and U. Spagnolini are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milano, Italy (e-mail: daniela.donno@gmail.com; spagnoli@elet.polimi.it).

A. Nehorai is with the Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: nehorai@ese.wustl.edu).

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

Digital Object Identifier 10.1109/TSP.2008.927459

netic wavefields can also be used in the wideband case by processing each frequency bin independently or by employing focussing matrices [19] that aim at referring the narrowband components of a signal to a central frequency. However, since the spectrum of seismic waveforms is too wide to allow for focalization to a unique central frequency, herein we propose a shift-invariance-based method that processes each frequency sample independently and then opportunely combines the estimates at different frequencies.

In real seismic surveys, array shape and sensor orientation are affected by uncertainty. Sensor-calibration errors would become a source of noise, interfering with the data-acquisition accuracy as well as the parameter estimation [20]–[21]. In this paper, the effect of uncertainty in sensor location and orientation on the accuracy of the velocity/polarization estimation is analyzed through a comparison with the Cramér–Rao lower bounds, showing that the estimates attain the bound in the presence of sensor-calibration errors.

This paper is organized as follows. In Section II, we present a parametric model for wideband polarized signals received by an array of three-component (3C) sensors. This model is based on the seismic polarization model in [22], but it is further extended to allow for modeling of sensor-calibration uncertainty as is often the case in practical experiments. In Section III, we evaluate the Cramér–Rao bound on the estimation accuracy that accounts also for model uncertainties. We propose in Section IV an algorithm that exploits the shift-invariance properties of the data for the joint estimation of velocity and polarization. This method is based on the principle proposed in [23] but it uses a subspace-based approach. In Section V, we define a constrained optimization problem for filtering surface waves from seismic recordings by exploiting the velocity/polarization estimates. Moreover, smoothing constraints are introduced to allow for reliable filtering when estimates are affected by uncertainty. In Section VI, we compare the performance of the proposed estimation algorithm with Cramér–Rao bounds. Simulated and experimental data examples prove that the velocity/polarization estimation method is effective for surface-wave separation when the joint estimates of the wavefield parameters are exploited in spatial filtering design.

II. Parametric Model for Wideband Polarized Seismic Waves

The measurement model proposed herein is based on parameters describing the polarization properties of the waveforms. As proposed in [8] and [22], the polarized signal can be modeled by a set of four angles to describe the polarization state of the waveforms (experimental motivations for this model can be found, e.g., in [24]). Unlike the model in [22], the model presented here holds for *wideband* seismic wavefields and it is extended to allow for modeling of sensor-calibration uncertainties.

This section is organized as follows: we propose (Section II-A) the geometric model for one polarized wavefield impinging on a single three-component sensor; then (Section II-B) we present the multiple-wavefield multiple-sensor model; finally (Section II-C), we extend the model to allow for positional and rotational uncertainties of the acquisition geometry.

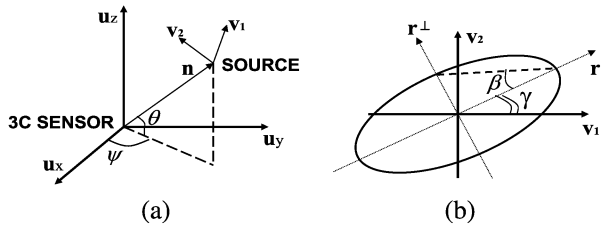

Fig. 1. Geometrical model of polarized signals. (a) The azimuth ψ and elevation θ relate the reference coordinate systems of the sensor $\{\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z\}$ and the source $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{n}\}$. (b) The polarization ellipse in the plane of polarization with the ellipse orientation angle γ and the ellipticity angle β.

A. Single-Wavefield Single-Sensor Model

For the single-wavefield single-sensor model in Fig. 1(a), the $\{\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z\}$ right orthonormal triad is referred to the three-components sensor (3C geophone), while the $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{n}\}$ right orthonormal triad is referred to the wavefield source. Throughout this section, the following assumptions hold.

- A1) The wavefield is propagating in a *homogeneous* and *isotropic* medium: this is equivalent to suppose the seismic radiation propagates along straight-lines, which is of practical use in the context of oil exploration and geophysics [25].
- A2) The wavefield is propagating in a *nondispersive* medium: the velocity of the wavefield is independent of frequency.
- A3) *Omnidirectional sensors*: the radiation pattern is constant in the horizontal plane.
- A4) *Plane-wave at the sensor*: this is equivalent to a *far-field* assumption (or a maximum wavelength that is much smaller than the source-to-sensor distance), a point-source assumption (i.e., the source size is much smaller than the source-to-sensor distance), and a point-like sensor assumption (i.e., the sensor's dimensions are small compared with the minimum wavelength). Notice that this is a common assumption in seismic wave analysis and processing to approximate seismic wavefronts by locally planar wavefronts (e.g., refer to [25] and [26]).

Let $\mathbf{y}(t) = [y_x(t), y_y(t), y_z(t)]^T \in \mathbb{C}^{3\times 1}$ be the noise-free complex analytical signal at the output of the three-component sensor in one observation period of duration T. Let the plane of polarization (i.e., the plane in which the particle moves) shown in Fig. 1 be defined by its normal vector $\mathbf{n}(\psi, \theta)$, whose *azimuth* angle is $\psi \in [-\pi, \pi]$ and *elevation* angle is $\theta \in [-\pi/2, \pi/2]$. Let $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2]$ be the base of the polarization plane of the waveform with respect to the $\{\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z\}$ triad, as follows:

$$\mathbf{V} = \begin{bmatrix} -\sin\psi & -\cos\psi\sin\theta \\ \cos\psi & -\sin\psi\sin\theta \\ 0 & \cos\theta \end{bmatrix}. \tag{1}$$

The polarized signal can be written as

$$\mathbf{y}(t) = \mathbf{V}\boldsymbol{\xi}(t) \tag{2}$$

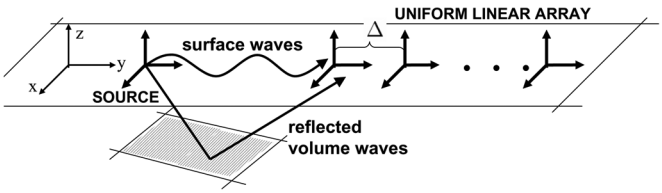

Fig. 2. Geometry of the scenario for the multiple-wavefield multiple-sensor model. This is a typical geometry of seismic experiments which consist of recording the seismic waves generated by a point source at the surface (a man-made explosion or a vibrator) using a uniform linear array of geophones.

and the two-component signal $\boldsymbol{\xi}(t) \in \mathbb{C}^{2\times 1}$ that fully determines the components of $\mathbf{y}(t)$ in the plane of polarization is

$$\boldsymbol{\xi}(t) = \mathbf{Q}\mathbf{b}x(t) \tag{3}$$

where $x(t)$ is the source signal at the transmitter. The vector $\mathbf{b} = [\cos\beta, j\sin\beta]^T$ contains the information about the signal ellipticity, with β being the *ellipticity* angle (the ellipticity is $\varepsilon = \tan\beta \in [-1, 1]$). $\mathbf{Q}$ is the 2×2 rotation matrix of angle $\gamma \in [-\pi/2, \pi/2]$ between the polarization ellipse axes and the base $[\mathbf{v}_1, \mathbf{v}_2]$ [see Fig. 1(b)]. Therefore, the wideband polarized signal is

$$\mathbf{y}(t) = \mathbf{V}\mathbf{Q}\mathbf{b}x(t) = \mathbf{p}(\boldsymbol{\Phi})x(t) \tag{4}$$

where $\mathbf{p}(\boldsymbol{\Phi}) \in \mathbb{C}^{3\times 1}$ is the response vector of the sensor (or *polarization vector*) that contains all the information about the polarization state of the wavefields. The polarization vector depends on the set of angle parameters $\boldsymbol{\Phi} = [\psi, \theta, \beta, \gamma]^T$.

By computing the Fourier transform of the signal $\mathbf{y}(t)$ within the observation period T, the lth frequency of the output signal in additive Gaussian noise is

$$\mathbf{y}(f) = X(f)\mathbf{p}(\boldsymbol{\Phi}) + \mathbf{e}(f) \qquad f = 1, \ldots, L \tag{5}$$

where $X(f)$ is the source signal in the frequency domain. The noise is Gaussian and uncorrelated for each sensor component: $\mathbf{e}(f) \sim CN\left(0, \sigma^2\mathbf{I}_3\right)$. Notice that the frequency domain model in (5) consists of L equations corresponding to L different frequencies. The *wideband* signal has been described as a combination of narrowband components, thus reducing the wideband problem to L narrowband equivalent problems for each of the frequencies f (for $f = 1, \ldots, L$) of the input signal.

B. Multiple-Wavefield Multiple-Sensor Model

Assume that there are d wavefields impinging on a uniform linear array (ULA) of M three-component sensors, spaced by Δ, as represented in Fig. 2. In order to generalize the model for one wavefield to d wavefields, we need assumptions A1 to A4 to hold for each of the d wavefields in this model. Moreover, we assume there are no distortions of the wavefields (i.e., polarization and waveform remain unchanged along the array); thus, the signal observed along the array of sensors is a delayed copy of the wavefields observed at the first sensor.

The lth frequency of the signal from M sensors $\mathbf{Y}(f) \in \mathbb{C}^{M\times 3}$ in additive Gaussian noise is the sum of d polarized wavefields recorded along the 3C-sensors ULA array

$$\mathbf{Y}(f) = \sum_{k=1}^{d} \left[X_k(f)\,\mathbf{a}(f, q_k)\,\mathbf{p}(\boldsymbol{\Phi}_k)^T + \mathbf{e}(f)\right]. \tag{6}$$

For the kth wavefield (with $k = 1, \ldots, d$) the $M \times 1$ vector $\mathbf{a}(f, q_k)$, also called the *steering vector*, contains the array response for the delays over the array $\left[\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(M)}\right]$

$$\mathbf{a}(f, q_k) = [e^{-j2\pi\tau^{(1)}f}, e^{-j2\pi\tau^{(2)}f}, \ldots, e^{-j2\pi\tau^{(M)}f}]^T. \tag{7}$$

Under assumption A4, the differential time delay of the plane wavefronts along the array turns out to be *linear*. In this case, the delay at the ith sensor (with $i = 1, \ldots, M$) is linearly varying

$$\tau^{(i)} = \Delta\,(i-1)\,q_k \tag{8}$$

with q_k the slowness of the kth wavefield (inverse of the velocity $q_k = 1/v_k$), and the kth steering vector (7) simplifies to

$$\mathbf{a}(f, q_k) = \left[1, e^{-j\eta_k f}, \ldots, e^{-j\eta_k(M-1)f}\right]^T \tag{9}$$

with $\eta_k = 2\pi\Delta q_k$.

In a compact matrix form, (6) can be written as

$$\mathbf{Y}(f) = \mathbf{A}(f)\mathbf{X}(f)\mathbf{P}(\boldsymbol{\Theta}) + \mathbf{E}(f) \tag{10}$$

where $\mathbf{A}(f) = [\mathbf{a}(f, q_1), \ldots, \mathbf{a}(f, q_d)]$ is the $M \times d$ array response matrix for the overall d wavefields. Matrix $\mathbf{X}(f) = \mathrm{diag}\,[X_1(f), \ldots, X_d(f)]$ accounts for the amplitudes, $\mathbf{P}(\boldsymbol{\Theta}) = [\mathbf{p}(\boldsymbol{\Phi}_1), \ldots, \mathbf{p}(\boldsymbol{\Phi}_d)]^T$ is the $d \times 3$ polarization matrix, and $\mathbf{E}(f)$ is the noise term. Notice that the model in (10) can be easily extended to dispersive media (thus neglecting assumption A2) by allowing the velocity of each wavefield to vary with frequency in (9).

Stacking the columns of the $M \times 3$ data matrix $\mathbf{Y}(f)$ into the $3M \times 1$ vector $\mathbf{z}(f)$, we obtain

$$\mathbf{z}(f) = \mathrm{vec}(\mathbf{Y}(f)) = \mathbf{H}(f)\,\mathbf{x}(f) + \mathbf{n}(f) \tag{11}$$

where $\mathbf{x}(f) = \mathrm{vec}(\mathbf{X}(f))$, $\mathbf{n}(f) = \mathrm{vec}(\mathbf{E}(f))$, and $\mathbf{H}(f) = \left(\mathbf{P}(\boldsymbol{\Theta})^T \otimes \mathbf{A}(f)\right)$ describes the array manifold that depends on the polarization and the velocity of each of the d wavefields (symbol "$\otimes$" denotes the Kronecker matrix product; see [27] for the main properties). Relationship (10) decouples the terms that depend on velocity in $\mathbf{A}(f)$ from those that are related to the polarization in $\mathbf{P}(\boldsymbol{\Theta})$ and from the amplitude parameters in $\mathbf{X}(f)$. Consequently, once the velocity and polarization parameters have been estimated, the amplitude parameters can be estimated as a simple linear regression problem. The unknown shape parameters of the kth wavefield are the slowness q_k of the wavefield, the complex amplitude $X_k(f) = X_{R,k} + jX_{I,k}$, and the set of angle parameters $\boldsymbol{\Phi}_k = [\psi_k, \theta_k, \beta_k, \gamma_k]^T$ that characterizes the polarization vector $\mathbf{p}(\boldsymbol{\Phi}_k)$ of each wavefield. These parameters to be estimated can be summed up as

$$\boldsymbol{\Psi}_k = [q_k, X_{R,k}, X_{I,k}, \psi_k, \theta_k, \beta_k, \gamma_k]^T. \tag{12}$$

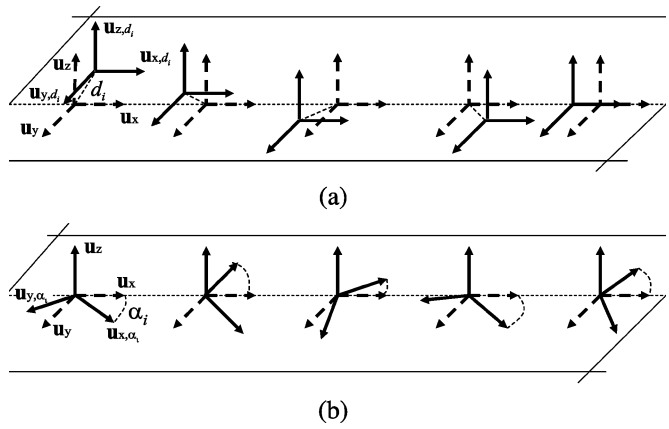

Fig. 3. Geometry of the array scenario for the multiple-wavefield multiple-sensor model with calibration errors. (a) *Positional calibration errors*: the actual vector-sensor triads $\{\mathbf{u}_{x,d_i}, \mathbf{u}_{y,d_i}, \mathbf{u}_{z,d_i}\}$ with positional calibration errors (in solid lines) with respect to the nominal vector-sensor triad $\{\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z\}$ (in dotted lines). (b) *Rotational calibration errors*: the horizontal components of the three-component sensors $\{\mathbf{u}_{x,\alpha_i}, \mathbf{u}_{y,\alpha_i}, \mathbf{u}_z\}$ (in solid lines) are misoriented with respect to the nominal vector-sensor triad $\{\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z\}$ (in dotted lines).

C. Parametric Model With Vector-Sensor Calibration Errors

In this section, we extend the multiple-wavefield multiple-sensor model to allow for both positional and rotational sensor calibration uncertainties.

1) Vector-Sensors With Positional Calibration Errors: Let us assume the ith three-component sensor $(i = 1, \ldots, M)$ of the nominally ULA array has a random, but time-invariant, displacement d_i from its nominal location $\Delta(i-1)$ [refers to Fig. 3(a)]. The positional errors d_i at each sensor are independent Gaussian random variables, all with zero mean and standard deviation σ_d $\left(d_i \sim N\left(0, \sigma_d^2\right)\right)$, with σ_d being small compared to the inter-sensor spacing Δ and to the distance between the source and the first sensor of the array.

When we introduce location uncertainties into the model in (10), only the steering matrix $\mathbf{A}(f)$ is affected by the location jitters d_i. The kth column $\mathbf{a}(f, q_k)$ of $\mathbf{A}(f)$ becomes (subscript "c" indicates calibration errors)

$$\mathbf{a}_c(f, q_k, \mathbf{d}) = \begin{bmatrix} 1 \cdot e^{-j2\pi d_1 q_k f} \\ e^{-j2\pi \Delta q_k f} \cdot e^{-j2\pi d_2 q_k f} \\ \vdots \\ e^{-j2\pi \Delta q_k (M-1) f} \cdot e^{-j2\pi d_M q_k f} \end{bmatrix} \tag{13}$$

or, in matrix form

$$\mathbf{a}_c(f, q_k, \mathbf{d}) = \mathbf{a}(f, q_k) \odot \mathbf{a}_{\text{pos}}(f, q_k, \mathbf{d}) \tag{14}$$

where $\mathbf{d} = [d_1, d_2, \ldots, d_M]^T$ is the vector of the sensor location uncertainties, and $\mathbf{a}_{\text{pos}}(f, q_k, \mathbf{d}) = \left[e^{-j2\pi d_1 q_k f}, e^{-j2\pi d_2 q_k f}, \ldots, e^{-j2\pi d_M q_k f}\right]^T$ is the kth steering vector which accounts for positional calibration errors (symbol "$\odot$" denotes the element-wise matrix product).

For the overall d wavefields, the steering matrix $\mathbf{A}(f)$ in the presence of sensor location uncertainties becomes

$$\mathbf{A}_c(f, \mathbf{d}) = \mathbf{A}(f) \odot \mathbf{A}_{\text{pos}}(f, \mathbf{d}) \tag{15}$$

where $\mathbf{A}_{\text{pos}}(f, \mathbf{d}) = [\mathbf{a}_{\text{pos}}(f, q_1, \mathbf{d}), \ldots, \mathbf{a}_{\text{pos}}(f, q_d, \mathbf{d})]^T$.

2) Vector-Sensors With Rotational Calibration Errors: We suppose the vertical component of the ith sensor to be correctly planted in the ground, while the two horizontal components (even if mutually orthogonal) are affected by a rotational error α_i with respect to their nominal angular position, as represented in Fig. 3(b). Let the M three-component sensors be mutually independent and the unknown angle of rotation α_i at the ith sensor be distributed as a zero-mean Gaussian random variable with standard deviation σ_α : $\alpha_i \sim N\left(0, \sigma_\alpha^2\right)$. The M-dimensional vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]^T$ accounts for the array rotational calibration errors.

Sensor-orientation uncertainties introduced into the model in (10) (and thus in (11)) are equivalent to a rotation of angle α_i (at the ith sensor) of the horizontal components of the polarization vector $\mathbf{p}(\boldsymbol{\Phi}_k)$ of the kth wavefield. The kth polarization vector with orientation uncertainties is $\mathbf{p}_c(\boldsymbol{\Phi}_k, \boldsymbol{\alpha})$

$$\mathbf{p}_c^T(\boldsymbol{\Phi}_k, \boldsymbol{\alpha}) = \mathbf{p}^T(\boldsymbol{\Phi}_k)\mathbf{R}(\boldsymbol{\alpha}) \tag{16}$$

where the rotation matrix $\mathbf{R}(\boldsymbol{\alpha})$ accounts for the overall orientation errors at the horizontal components of the M three-component sensors

$$\mathbf{R}(\boldsymbol{\alpha}) = \frac{1}{M}\begin{bmatrix} \sum_{i=1}^{M} \cos\alpha_i & \sum_{i=1}^{M} \sin\alpha_i & 0 \\ -\sum_{i=1}^{M} \sin\alpha_i & \sum_{i=1}^{M} \cos\alpha_i & 0 \\ 0 & 0 & M \end{bmatrix}. \tag{17}$$

In presence of sensor-orientation uncertainties, the polarization matrix $\mathbf{P}(\boldsymbol{\Theta})$ of the model (10) for the overall d wavefields becomes

$$\mathbf{P}_c(\boldsymbol{\Theta}, \boldsymbol{\alpha}) = \mathbf{P}(\boldsymbol{\Theta})\mathbf{R}(\boldsymbol{\alpha}). \tag{18}$$

3) Parametric Model With Positional and Rotational Calibration Errors: Summarizing, in the case of multiple-wavefield and multiple-sensor with calibration uncertainties of both position and orientation, the output signal $\mathbf{z}(f)$ in (11) is modified as

$$\mathbf{z}_c(f) = \mathbf{H}_c(f)\mathbf{x}(f) + \mathbf{n}(f) \tag{19}$$

where $\mathbf{H}_c(f) = \left(\mathbf{P}_c(\boldsymbol{\Theta}, \boldsymbol{\alpha})^T \otimes \mathbf{A}_c(f, \mathbf{d})\right)$ describes the array manifold for d wavefields in the presence of sensor-calibration errors, with $\mathbf{A}_c(f, \mathbf{d})$ and $\mathbf{P}_c(\boldsymbol{\Theta}, \boldsymbol{\alpha})$ defined as in (15) and (18), respectively. The unknown shape and sensor-calibration error for the kth wavefield can be summed up in the set of $(7 + 2M)$ parameters

$$\boldsymbol{\Upsilon}_k = [\boldsymbol{\Psi}_k, \mathbf{d}, \boldsymbol{\alpha}] \tag{20}$$

where $\boldsymbol{\Psi}_k$ accounts for the deterministic shape parameters of the kth wavefield defined in (12), and $\mathbf{d} \sim N\left(0, \sigma_d^2 \mathbf{I}_M\right)$, and $\boldsymbol{\alpha} \sim N\left(0, \sigma_\alpha^2 \mathbf{I}_M\right)$ are random calibration errors.

III. Cramér–Rao Bounds

The Cramér–Rao bound (CRB) provides a lower bound on the variance of any unbiased estimator [28], thus allowing for a benchmark against which we can compare the performance of an estimator. In this section, we present the Cramér–Rao bound

for the parameters of the physical model proposed in Section II. Within this section, dependence on frequency is omitted only for sake of simplifying the notation.

A. CRB for the Model Parameters With Known Sensor Location and Orientation

Let $\hat{\mathbf{\Psi}}$ be the estimator of the unknown set of shape parameters $\mathbf{\Psi} = [q, X_R, X_I, \psi, \theta, \gamma, \beta]^T$, as defined in Section II-B. We derive the CRB for the covariance matrix of the estimate error $(\hat{\mathbf{\Psi}} - \mathbf{\Psi})$ in the case of a single wavefield $(d = 1)$ and multiple sensors. The physical model (11) reduces to

$$\begin{aligned}\mathbf{z}(\mathbf{\Psi}) &= X\mathbf{p}(\mathbf{\Phi}) \otimes \mathbf{a}(q) + \mathbf{n} \\ &= \boldsymbol{\mu}(\mathbf{\Psi}) + \mathbf{n}. \qquad (3M \times 1).\end{aligned} \tag{21}$$

The Cramér–Rao bound follows from the Fisher information matrix (FIM) $\mathbf{J}(\mathbf{\Psi})$ as $\text{CRB}(\mathbf{\Psi}) = \mathbf{J}^{-1}(\mathbf{\Psi})$. Having supposed the noise to be i.i.d. zero-mean circular complex Gaussian and independent from the signal $\mathbf{n} \sim CN\left(0, \sigma^2\mathbf{I}\right)$, in this case [28] the FIM components are

$$[\mathbf{J}(\mathbf{\Psi})]_{ij} = \frac{2}{\sigma^2}\text{Re}\left[\left(\frac{\partial \boldsymbol{\mu}(\mathbf{\Psi})}{\partial \mathbf{\Psi}}\right)^H \left(\frac{\partial \boldsymbol{\mu}(\mathbf{\Psi})}{\partial \mathbf{\Psi}}\right)\right]. \tag{22}$$

After inverting the Fisher information matrix $\mathbf{J}(\mathbf{\Psi})$, the CRBs for the set $\mathbf{\Psi}$ of shape parameters are given by

$$\text{CRB}\,(\hat{q}) = \frac{6}{M\,(M^2-1)\,\phi^2\rho_f} \tag{23}$$

$$\text{CRB}(\hat{X}_R) = CRB(\hat{X}_I) = \frac{\kappa|X\,(f)\,|^2}{2M\,(M+1)\,\rho_f} \tag{24}$$

$$\text{CRB}(\hat{\psi}) = \frac{1 - \cos 2\gamma \cos 2\beta}{M\rho_f \cos^2\theta \sin^2 2\beta} \tag{25}$$

$$\text{CRB}(\hat{\theta}) = \frac{1 + \cos 2\gamma \cos 2\beta}{M\rho_f \sin^2 2\beta} \tag{26}$$

$$\text{CRB}(\hat{\gamma}) = \frac{\xi}{M\rho_f} \tag{27}$$

$$\text{CRB}(\hat{\beta}) = \frac{1}{2M\rho_f} \tag{28}$$

where the parameters ϕ, ρ_f, κ, and ξ are defined in (68)–(71) of Appendix A, respectively.

The matrix of the Cramér–Rao bounds is in fact block-diagonal (the overall entries of the CRB matrix are given in Appendix A) with the two nonzero block matrices consisting of the velocity-related parameters $[q, X_R, X_I]$ for the first block matrix and the parameters $[\psi, \theta, \gamma, \beta]$ that characterize the polarization for the second block matrix. Therefore, the velocity-related parameters are decoupled from the polarization parameters. This result was expected [29] since it is quite clear that, under the hypotheses of isotropic medium and planar wavefronts, the velocity parameters (which depend from the sensor positions and spacing, in matrix $\mathbf{A}$ of our model) are physically decoupled from the polarization parameters (which depend from the angles of arrival on the three-component sensors, in matrix $\mathbf{P}$).

Moreover, from (23) and (28), we notice that the lower bounds of the slowness q and the ellipticity angle β are decoupled from the polarization parameters ψ, θ, γ, and β. The CRB does not depend on the other parameters of the physical model but only on the geometry of the experiment. Therefore the CRB is an analytic tool to optimize the geometry of acquisition and to improve the estimation performance of the two parameters q and β.

B. CRB for the Model Parameters With Calibration Errors

When the sensor locations and orientations are uncertain, from (19) the output signal at the lth frequency for a single source wavefield is

$$\begin{aligned}\mathbf{z}_c(\mathbf{\Upsilon}) &= X\mathbf{p}_c(\mathbf{\Phi}, \mathbf{d}) \otimes \mathbf{a}_c(q, \boldsymbol{\alpha}) + \mathbf{n} \\ &= \boldsymbol{\mu}(\mathbf{\Upsilon}) + \mathbf{n}. \qquad (3M \times 1).\end{aligned} \tag{29}$$

We derive herein the CRB for the unknown set of model parameters $\mathbf{\Upsilon} = [\mathbf{\Psi}, \mathbf{d}, \boldsymbol{\alpha}]$, where the parameters of interest $\mathbf{\Psi}$ are mixed with the random nuisance parameters $\mathbf{d}$ and $\boldsymbol{\alpha}$.

When the unknown parameters are random variables with known prior distribution, the equivalent formula for the CRB (also known as posterior CRB [30]) is (see [20])

$$\text{CRB}(\mathbf{\Upsilon}) = \mathbf{J}(\mathbf{\Upsilon})^{-1} = (\mathbf{J}_1(\mathbf{\Upsilon}) + \mathbf{J}_2(\mathbf{\Upsilon}))^{-1} \tag{30}$$

where $\mathbf{J}_1$ and $\mathbf{J}_2$ are components of the Fisher information matrix $\mathbf{J}(\mathbf{\Upsilon})$ describing, respectively, the contribution of the observations and the prior statistics to the bound on the estimation error $(\hat{\mathbf{\Upsilon}} - \mathbf{\Upsilon})$. Their components are

$$[\mathbf{J}_1(\mathbf{\Upsilon})]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{z}_c|\mathbf{\Upsilon})}{\partial\Upsilon_i \partial\Upsilon_j}\right] \tag{31}$$

and

$$[\mathbf{J}_2(\mathbf{\Upsilon})]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{\Upsilon})}{\partial\Upsilon_i \partial\Upsilon_j}\right] \tag{32}$$

where $p(\mathbf{z}_c|\mathbf{\Upsilon})$ is the conditional density function of $\mathbf{z}_c$ for the given parameter vector $\mathbf{\Upsilon}$, and $p(\mathbf{\Upsilon})$ is the prior probability density of $\mathbf{\Upsilon}$.

It can be proved that (30) still holds for hybrid systems [18] with mixed deterministic $\mathbf{\Psi}$ and stochastic parameters ($\mathbf{d}$ and $\boldsymbol{\alpha}$). Of course, the terms of matrix $\mathbf{J}_2(\mathbf{\Upsilon})$ in (32) that involve the derivative with respect to $\mathbf{\Psi}$ are simply zero

$$\mathbf{J}_2(\mathbf{\Upsilon}) = 2\left[\begin{array}{c|c|c} & \mathbf{\Psi} & \mathbf{d} \qquad\qquad \boldsymbol{\alpha} \\ \hline \mathbf{\Psi} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times M} \qquad \mathbf{0}_{7\times M} \\ \mathbf{d} & \mathbf{0}_{M\times 7} & \frac{1}{\sigma_d^2}\mathbf{I}_M \qquad \mathbf{0}_{M\times M} \\ \boldsymbol{\alpha} & \mathbf{0}_{M\times 7} & \mathbf{0}_{M\times M} \qquad \frac{1}{\sigma_\alpha^2}\mathbf{I}_M \end{array}\right]. \tag{33}$$

For a circular complex Gaussian signal $(\mathbf{z}_c(\mathbf{\Upsilon}) \sim CN\left(\boldsymbol{\mu}\left(\mathbf{\Upsilon}\right), \mathbf{C}_\mathbf{z}\left(\mathbf{\Upsilon}\right)\right))$, the components of $\mathbf{J}_1(\mathbf{\Upsilon})$ can be written as [28]

$$\begin{aligned}[\mathbf{J}_1(\mathbf{\Upsilon})]_{ij} = \text{tr}&\left[\mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})^{-1}\frac{\partial \mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})}{\partial\Upsilon_i}\mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})^{-1}\frac{\partial \mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})}{\partial\Upsilon_j}\right] \\ &+2\text{Re}\left[\frac{\partial \boldsymbol{\mu}(\mathbf{\Upsilon})^H}{\partial\Upsilon_i}\mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})^{-1}\frac{\partial \boldsymbol{\mu}(\mathbf{\Upsilon})}{\partial\Upsilon_j}\right].\end{aligned} \tag{34}$$

where the derivatives of the covariance matrix $\mathbf{C}_\mathbf{z}(\mathbf{\Upsilon})$ required for the evaluation of the CRB are given in Appendix B. CRBs

for the parameters of interest $\mathbf{\Psi}$ have no simple compact form and are given in Section VI-B for the cases of interest.

IV. Joint Estimation of Velocity and Polarization

In this section, we propose a shift-invariance-based method to separate volume waves from surface waves by *jointly* estimating the velocity and polarization parameters. The interest of the subspace-based method is that this algorithm has attractive computational cost, compared with the prohibitive costs for the implementation of the MLE, because the velocity/polarization estimation with MLE is a nonlinear optimization problem over a large number of unknowns (only solvable using numerical techniques, such as the EM algorithm or the Newton–Raphson method [28]).

Unlike the ESPRIT [31] method which is limited to processing narrowband data, here we exploit the shift invariance over all the L frequency samples having the same shape parameters. Within this section, dependence on frequency is omitted for the sake of notational simplicity.

Let the singular value decomposition of the data matrix $\mathbf{Y}$ in (10) be

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{35}$$

where $\mathbf{U}$ and $\mathbf{V}$ are the unitary matrices of size $M \times M$ and 3×3, formed, respectively, with the left and right singular vectors of $\mathbf{Y}$, while $\mathbf{\Sigma}$ is an $M \times 3$ diagonal matrix that contains the singular values of $\mathbf{Y}$ sorted in decreasing order.

The singular vectors associated with the d largest singular values of $\mathbf{Y}$ are known to span the signal subspace (see, e.g., [32] for details). Therefore, matrix $\mathbf{Y}$ can be partitioned into signal and noise subspace matrices, resulting in

$$\mathbf{Y} = \mathbf{U}_S\mathbf{\Sigma}_S\mathbf{V}_S^H + \mathbf{U}_N\mathbf{\Sigma}_N\mathbf{V}_N^H. \tag{36}$$

The signal subspace $\mathbf{U}_S\mathbf{\Sigma}_S\mathbf{V}_S^H$ can be (approximately) identified with the noise-free signal of $\mathbf{Y}$, hence $\mathbf{U}_S\mathbf{\Sigma}_S\mathbf{V}_S^H \simeq \mathbf{AXP}$. It follows that the d left singular vectors of $\mathbf{Y}$ can be written as

$$\mathbf{U}_S = \mathbf{AXC} \tag{37}$$

where $\mathbf{C} = \mathbf{P}\mathbf{V}_S\mathbf{\Sigma}_S^{-1}$ is a $d \times d$ nonsingular matrix.[1] The interest of this decomposition is that the velocity and polarization information are preserved in $\mathbf{U}_S$. Moreover, $\mathbf{U}_S$ has a structure similar to that of $\mathbf{Y}$. Consequently, a shift-invariance technique can be applied to $\mathbf{U}_S$ to estimate the shape parameters of the wavefields.

Let $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ be two subsets of the data $\mathbf{U}_S$ that differ only by a shift. For an array of M elements, $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ are chosen so that one is the shifted copy of the other, by taking the first and the last $M-1$ sensors, respectively

$$\mathbf{U}_S^{(1)} = \mathbf{J}^{(1)}\mathbf{U}_S \text{ and } \mathbf{U}_S^{(2)} = \mathbf{J}^{(2)}\mathbf{U}_S \tag{38}$$

[1]Since the columns of $\mathbf{U}_S = \mathbf{AXC}$ span the same subspace of $\mathbf{A}$, the matrix $\mathbf{XC}$ is non singular. Then, since $\mathbf{X}$ is a $d \times d$ diagonal matrix, it follows that $\mathbf{C}$ is nonsingular.

where $\mathbf{J}^{(1)} = \left[\mathbf{I}_{M-1}, \mathbf{0}_{(M-1)\times 1}\right]$ and $\mathbf{J}^{(2)} = \left[\mathbf{0}_{(M-1)\times 1}, \mathbf{I}_{M-1}\right]$ are two selection matrices of dimension $(M-1) \times M$. Similarly, let $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ be the two $(M-1) \times d$ subsets of the array response $\mathbf{A}$

$$\mathbf{A}^{(1)} = \mathbf{J}^{(1)}\mathbf{A} \text{ and } \mathbf{A}^{(2)} = \mathbf{J}^{(2)}\mathbf{A}. \tag{39}$$

In accordance with the shift-invariance property between two shifted subsets, the relationship between $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ is linear and depends on the wavefield velocity and polarization. In fact, as with the ESPRIT method, the two subsets $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ differ only by a phase shift that can be collected (for all the d wavefields) into the diagonal matrix $\mathbf{D} = \operatorname{diag}\left[e^{-j\eta_1 f}, \ldots, e^{-j\eta_d f}\right]$, where $\eta_k = 2\pi\Delta/v_k$ (with $k = 1, \ldots, d$).

Recalling the structure of the array response matrix $\mathbf{A}$, we can state that $\mathbf{A}^{(2)} = \mathbf{A}^{(1)}\mathbf{D}$. Moreover, according to (37), subsets $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ can be modeled as

$$\mathbf{U}_S^{(1)} = \mathbf{A}^{(1)}\mathbf{XC} \tag{40}$$

$$\mathbf{U}_S^{(2)} = \mathbf{A}^{(2)}\mathbf{XC} = \mathbf{A}^{(1)}\mathbf{DXC} = \mathbf{A}^{(1)}\mathbf{XDC} \tag{41}$$

where, in (41), $\mathbf{X}$ and $\mathbf{D}$ can be switched since both are diagonal matrices. The relationship between $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ can be reduced to $\mathbf{U}_S^{(1)}\mathbf{T} = \mathbf{U}_S^{(2)}$, where the $d \times d$ shifting matrix $\mathbf{T}$ has to be estimated. Rearranging the subsets (40) and (41) and considering full-rank matrices, we see that

$$\mathbf{T} = \mathbf{C}^{-1}\mathbf{DC} \tag{42}$$

thus matrix $\mathbf{T}$ contains the information of polarization and velocity for each of the d wavefields.

In particular, eigenvalues $\{\lambda_k\}_{k=1}^d$ of $\mathbf{T}$ in $\mathbf{D}$ are unit-amplitude complex values whose phases are related to the velocity estimates $\hat{v}_k$ as $\measuredangle\hat{\lambda}_k = \hat{\phi}_k = 2\pi\Delta/\hat{v}_k$ (symbol "$\measuredangle$" denotes the phase angle), while the eigenvectors of $\mathbf{T}$ in $\mathbf{C}^{-1}$ are related to the polarization of the wavefields by $\hat{\mathbf{P}} = \mathbf{C}\mathbf{\Sigma}_S\mathbf{V}_S^H$. As a result, for each frequency sample, velocity and polarization are estimated and automatically paired. In Section VI, when discussing experimental results, we will show that once the shape parameters have been estimated, the surface waves can be separated from volume waves in the velocity/ellipticity plane.

Notice that matrix $\mathbf{T}$ in (42) can be estimated by the *least squares* (LS) criterion. However, it is common practice in the ESPRIT method to adopt the *total least squares* (TLS) criterion [33] as it allows both $\mathbf{U}_S^{(1)}$ and $\mathbf{U}_S^{(2)}$ to be noisy. For large signal-to-noise ratio (SNR), the difference between the LS and the TLS estimates is small; while at low SNRs, unlike for the LS estimates, the TLS estimates have been shown [34] to be strongly consistent (converging with probability one to the true value).

1) Estimation of the Polarization Vector: The estimate of the kth polarization vector $\hat{\mathbf{p}}_k$ provided by the algorithm proposed in Section IV does not include an unknown complex scaling term ν, such that $\nu\hat{\mathbf{p}}_k = \mathbf{p}_k$, where $\mathbf{p}_k$ represents the exact polarization vector for the kth wavefield.

The real and imaginary components of $\hat{\mathbf{p}}_k$ belong to the plane of the polarization ellipse, but they are not assured to be orthogonal. The value of the constant $\nu = |\nu| e^{j\varphi_k}$ guarantees obtaining the exact principal axes of the polarization ellipse $\mathbf{r}$ and $\mathbf{r}^{\perp}$. Since we are interested only in the direction of the principal axes, the rotation $e^{j\varphi_k}$ in the complex plane follows by solving for the condition of orthogonality

$$\langle \mathrm{Re}(\hat{\mathbf{p}}_k e^{j\varphi_k}), \mathrm{Im}(\hat{\mathbf{p}}_k e^{j\varphi_k}) \rangle = 0 \tag{43}$$

with respect to φ_k (symbol "$\langle \cdot, \cdot \rangle$" denotes the inner product).

The solution of (43) gives the angle φ_k (see proof in Appendix C)

$$\tan(2\varphi_k) = -\frac{2\langle \mathrm{Re}(\hat{\mathbf{p}}_k), \mathrm{Im}(\hat{\mathbf{p}}_k) \rangle}{\| \mathrm{Re}(\hat{\mathbf{p}}_k) \|^2 - \| \mathrm{Im}(\hat{\mathbf{p}}_k) \|^2}. \tag{44}$$

Once φ_k is given from (44), the principal axes $\mathbf{r}$ and $\mathbf{r}^{\perp}$ are given by $\mathrm{Re}(\hat{\mathbf{p}}_k e^{j\varphi_k})$ and $\mathrm{Im}(\hat{\mathbf{p}}_k e^{j\varphi_k})$, and the ellipticity ε_k can be calculated as the ratio between the norm of the minor axis and the norm of the major axis

$$\varepsilon_k = \frac{\min\left(\| \mathrm{Re}(\hat{\mathbf{p}}_k e^{j\varphi_k}) \|, \| \mathrm{Im}(\hat{\mathbf{p}}_k e^{j\varphi_k}) \| \right)}{\max\left(\| \mathrm{Re}(\hat{\mathbf{p}}_k e^{j\varphi_k}) \|, \| \mathrm{Im}(\hat{\mathbf{p}}_k e^{j\varphi_k}) \| \right)}. \tag{45}$$

2) Data Extension: To increase the number d of wavefields that can be estimated and to obtain a more accurate estimate of the signal subspace, it is common practice to use smoothing techniques (see, for instance, [35]). We propose adapting the spatial smoothing techniques currently used for DOA estimation to increase the number of columns of the data matrix by reducing the number of sensors. The idea of spatial smoothing is to use the redundancy of the M sensors to increase the number of observations. This technique is based on the assumption that the shifted observation of one sensor is uncorrelated with the original observation.

Let M_1 be the spatial smoothing coefficient, defined as the number of shifted copies of the data. As a consequence of spatial smoothing, the number of effective sensors reduces to $(M - M_1 + 1)$. The $(M - M_1 + 1) \times 3M_1$ new data matrix $\mathbf{Y}_S$ is formed by horizontally stacking M_1 shifted copies $\left\{\mathbf{Y}^{(k)}\right\}_{k=1}^{M_1}$ of the data $\mathbf{Y}$

$$\mathbf{Y}_S = \left[\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \ldots, \mathbf{Y}^{(M_1)} \right]. \tag{46}$$

The kth subset $\mathbf{Y}^{(k)}$ of $\mathbf{Y}_S$ is defined as

$$\mathbf{Y}^{(k)} = \mathbf{J}_{M_1}^{(k)} \mathbf{Y} = \mathbf{J}_{M_1}^{(k)} \mathbf{A} \mathbf{X} \mathbf{P} \tag{47}$$

where $\mathbf{J}_{M_1}^{(k)}$ is the $(M - M_1 + 1) \times M$ selection matrix that keeps the $(M - M_1 + 1)$ lines of the kth column of $\mathbf{Y}$.

Using the property of Vandermonde matrix $\mathbf{A}$, it follows that

$$\mathbf{Y}^{(k)} = \mathbf{A}_{M_1}^{(1)} \mathbf{D}^{k-1} \mathbf{X} \mathbf{P} = \mathbf{A}_{M_1}^{(1)} \mathbf{X} \mathbf{D}^{k-1} \mathbf{P} \tag{48}$$

where $\mathbf{D}^{k-1}$ and $\mathbf{X}$ can be exchanged because they are diagonal matrices (matrix $\mathbf{D}$ was defined in Section IV), and where the $(M - M_1 + 1) \times d$ matrix $\mathbf{A}_{M_1}^{(1)}$ is a subset of $\mathbf{A}$, with $(M - M_1 + 1)$ rows instead of M. The matrix $\mathbf{Y}_S = \left[\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \ldots, \mathbf{Y}^{(M_1)} \right]$ can be finally written as

$$\mathbf{Y}_S = \left[\mathbf{A}_{M_1}^{(1)} \mathbf{X} \mathbf{P}, \mathbf{A}_{M_1}^{(1)} \mathbf{X} \mathbf{D} \mathbf{P}, \ldots, \mathbf{A}_{M_1}^{(1)} \mathbf{X} \mathbf{D}^{M_1 - 1} \mathbf{P} \right] \tag{49}$$

$$= \mathbf{A}_{M_1}^{(1)} \mathbf{X} [\mathbf{P}, \mathbf{D}\mathbf{P}, \ldots, \mathbf{D}^{M_1 - 1} \mathbf{P}] \tag{50}$$

$$= \mathbf{A}_{M_1}^{(1)} \mathbf{X} \mathbf{P}_S \tag{51}$$

where the $d \times 3M_1$ matrix $\mathbf{P}_S = \left[\mathbf{P}, \mathbf{D}\mathbf{P}, \ldots, \mathbf{D}^{M_1 - 1} \mathbf{P} \right]$ is an extended version of the polarization matrix $\mathbf{P}$. Note that the first block of $\mathbf{P}_S$ remains equal to the polarization matrix $\mathbf{P}$.

The estimation method described in the previous section can now be applied to $\mathbf{Y}_S$, since the structure needed to apply the shift-invariance algorithm is preserved. From the structure in (51), the spatial smoothing assures that the column rank of the data matrix $\mathbf{Y}_S$ is $\min(d, 3M_1)$. Note that seismic sources are known to be highly correlated and even coherent, therefore spatial smoothing becomes necessary in the estimation algorithm workflow, since it essentially "decorrelates" the signals, thus eliminating the difficulties encountered with coherent signals as in [35].

3) Frequency Averaging of the Estimates: Since the recorded signal $\mathbf{Y}(f)$ (with varying frequency $f = 1, \ldots, L$) is wideband, we employ the shift-invariance-based technique to each frequency sample independently. Now, the estimates at different frequencies must be properly combined. As a result of the estimation, for the kth wavefield, L sets of estimated parameters $\left\{ [\hat{v}_k, \hat{\mathbf{p}}_k](f) \right\}_{f=1,\ldots,L}$ are obtained, one for each frequency. For the specific seismic application, we apply the algorithm only at the frequency samples belonging to a certain range around the central frequency f_c of the Ricker waveform[2] [36], and then we average these estimates. In this paper, frequencies in the range of $0.7 f_c$ up to $1.3 f_c$ have been used, as the signal energy is largest. However, when we deal with real seismic data the heuristic analysis of the data spectrum allows to select the range of frequencies for averaging the estimates.

Note that when $d > 1$ wavefields must be estimated, the estimates need to be clustered before the averaging, in order to group the estimates belonging to the same wavefield. It is known that the seismic surface waves are characterized by an elliptical polarization (high ellipticity values) and a low velocity, while the volume waves are linearly polarized (low ellipticity values) and have a higher velocity with respect to surface waves. Moreover, the surface waves show a lower frequency content than the volume waves. Therefore, this *a priori* information on the nature of the wavefields, together with a clustering algorithm (such as k-means [37] or others [38]) would allow for clustering of the velocity/polarization estimates (refer to Fig. 10 for an example).

It should be also noted that, as regards to the assumption A2 of nondispersive medium, the shape-parameter estimation method proposed herein would give reliable results also when employed with dispersive real data. In fact, since our estimation method processes each frequency sample independently, it would allow the velocity of each wavefield to vary with frequency. In such a case, the only difference with respect to the proposed method is about the average of the estimates. In fact, to allow for dispersive data, it would be suitable to employ a clustering algorithm and

[2]The Ricker waveform is commonly used in geophysics [1]–[25] to model seismic data. The time and frequency response of the Ricker waveform are

$$s(t) = \left(1 - 2\pi^2 f_c^2 t^2\right) e^{-\pi^2 f_c^2 t^2} \longleftrightarrow S(f) = \frac{2}{\sqrt{\pi}} \frac{f^2}{f_c^2} e^{-f^2/f_c^2} \tag{52}$$

where f_c is referred as central frequency.

averaging technique which use the typical dispersion curve of surface waves (a plot of wave velocity as a function of frequency [25]) as an *a priori*, to appropriately cluster and average together the estimates that belong to the same wavefield.

V. Velocity/Polarization-Constrained Wideband Beamforming

In this section, we propose the linearly constrained minimum variance (LCMV) beamforming problem to separate the interfering surface wavefields from seismic recordings. The basic idea of beamforming is to linearly constrain the response of the spatial filter so that signals of interest are not affected by attenuation, while suppressing the interfering wavefields.

The spatial filtering accounts for finding the matrix of weights $\mathbf{w}$ such that the desired signal $\mathbf{x}$ is estimated from the linear combination of the observed data vector $\mathbf{z} = \left[\mathbf{z}(f_1)^T, \ldots, \mathbf{z}(f_l)^T, \ldots, \mathbf{z}(f_L)^T\right]^T$ (with $\mathbf{z}(f_l)$ being defined as in (11))

$$\hat{\mathbf{x}} = \boldsymbol{w}^H \mathbf{z} \tag{53}$$

where $\mathbf{w} = \left[\mathbf{w}_1{}^T, \ldots, \mathbf{w}_l{}^T, \ldots, \mathbf{w}_L{}^T\right]^T$ is the matrix of weights for the overall L frequency samples. Filtering of the interfering wavefields is achieved by linearly constrained minimum variance (LCMV) beamforming, which estimates the optimal weights $\hat{\mathbf{w}}$ to minimize the output power of the beamformer

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \left\{\mathbf{w}^H \mathbf{R}_z \mathbf{w}\right\} \quad \text{subject to} \quad \mathbf{K}^H \mathbf{w} = \mathbf{F} \tag{54}$$

where $\mathbf{R}_z = E\left\{\mathbf{z}\mathbf{z}^H\right\}$ is the correlation matrix of the data matrix $\mathbf{z}$, and matrix $\mathbf{K}$ accounts for the constraints and $\mathbf{F}$ for the desired response. The method of Lagrange multipliers can be used to solve (54), resulting in

$$\hat{\mathbf{w}}^H = \mathbf{F}^H (\mathbf{K}^H \mathbf{R}_z^{\dagger} \mathbf{K})^{\dagger} \mathbf{K}^H \mathbf{R}_z^{\dagger} \tag{55}$$

where the symbol "†" indicates the pseudoinverse of the corresponding matrix. Note that, unlike with the classical LCMV beamformer [39], here we employ pseudoinverse matrices to allow for a low-rank covariance matrix $\mathbf{R}_z$ and linearly dependent constraints in matrix $\mathbf{K}$ [40].

The solution (55) is valid for arbitrary wideband signals. However, if we choose the constraint matrix $\mathbf{K}$ in (54) to be block diagonal $\mathbf{K} = \operatorname{diag}[\mathbf{K}_1, \ldots, \mathbf{K}_l, \ldots, \mathbf{K}_L]$, then the weights $\{\mathbf{w}_l\}_1^L$ are constrained independently and the wideband constraint $\mathbf{K}^H \mathbf{w} = \mathbf{F}$ can be reduced to a set of L independent constraints $\mathbf{K}_l^H \mathbf{w}_l = \mathbf{F}_l$, one for each of the L frequencies. Therefore, the broadband beamformer can be viewed as L narrowband beamformers that produce the frequency-domain filtered data $\hat{\mathbf{x}}(f) = \mathbf{w}_l^H \mathbf{z}(f)$ for $f = 1, \ldots, L$. The solution for the lth narrowband problem is similar to (55) $\hat{\mathbf{w}}_l^H = \mathbf{F}_l{}^H (\mathbf{K}_l^H \mathbf{R}_{zl}^{\dagger} \mathbf{K}_l)^{\dagger} \mathbf{K}_l^H \mathbf{R}_{zl}^{\dagger}$. The choices of the matrices $\mathbf{F}_l$ and $\mathbf{K}_l$ for the lth narrowband beamformer to satisfy the constraint in (54) are specified in Sections V-A-I and II.

1) Velocity/Polarization Constraint: A beamformer, in its classical formulation, behaves as a spatial filter. However, in the case of seismic wavefields, the assumption of spatial separation between signal and interference does not always hold. Therefore it is not possible to discriminate the wavefields using only their direction of arrival. In this paper, we propose exploiting the velocity/polarization diversity of volume waves and surface waves to constrain the beamformer.

Let the estimated polarization matrix $\hat{\mathbf{P}}$ be partitioned as $\hat{\mathbf{P}} = \begin{bmatrix} \hat{\mathbf{P}}_{\text{sur}} \\ \hat{\mathbf{P}}_{\text{vol}} \end{bmatrix}$, where $\hat{\mathbf{P}}_{\text{sur}}$ and $\hat{\mathbf{P}}_{\text{vol}}$ are the polarization matrices for surface and volume waves, respectively. The same partition for the lth frequency sample of the estimated steering matrix $\hat{\mathbf{A}}_l$ results in $\hat{\mathbf{A}}_l = [\,\hat{\mathbf{A}}_{\text{sur}} \quad \hat{\mathbf{A}}_{\text{vol}}\,]$. Note that the dimensions of the partition matrices and the classification of the shape-parameters estimates (as relative to surface or volume waves) are chosen after employing clustering techniques [38] together with the *a priori* information on the nature of the different seismic wavefields, as it will be shown in Section VI.

The beamformer response is constrained to have unitary gain for the velocity/polarization pairs relative to volume waves and zero gain for the estimates of surface waves. If we have only one wavefield, the constraints would be

$$\left(\hat{\mathbf{p}}_{\text{vol}}^T \otimes \hat{\mathbf{a}}_{\text{vol}}\right)^H \mathbf{w}_l = 1 \quad \text{and} \quad \left(\hat{\mathbf{p}}_{\text{sur}}^T \otimes \hat{\mathbf{a}}_{\text{sur}}\right)^H \mathbf{w}_l \approx 0 \tag{56}$$

where $\hat{\mathbf{p}}_{\text{vol}}$ and $\hat{\mathbf{p}}_{\text{sur}}$ are the 3×1 estimated polarization vectors defined as in (5), while $\hat{\mathbf{a}}_{\text{vol}}$ and $\hat{\mathbf{a}}_{\text{sur}}$ are the $M \times 1$ estimated steering vectors.

With more than one wavefield, multiple constraints as in (56) can be written as

$$\left(\begin{bmatrix} \hat{\mathbf{P}}_{\text{sur}} \\ \hat{\mathbf{P}}_{\text{vol}} \end{bmatrix}^T \otimes [\,\hat{\mathbf{A}}_{\text{sur}} \quad \hat{\mathbf{A}}_{\text{vol}}\,]\right)^H \mathbf{w}_l = \mathbf{F}_l \tag{57}$$

where $\mathbf{F}_l$ is the matrix of desired response, which has unitary entries in correspondence to volume waves and zeros in correspondence to surface waves.

2) Smooth Constraints: The velocity/polarization constraints (56) alone are inadequate to ensure good performance in the presence of errors in the estimation of the shape parameters as well as in the calibration of the sensors [e.g., if the shape-parameters relative to one volume wave are not well estimated, the volume wave itself might be eliminated by means of the constraint in (56)]. For this reason, we investigate smooth constraint methods to improve the robustness of beamforming in presence of uncertainties. We add the *derivative* constraints [41] by forcing the derivative of the array manifold $\mathbf{H}(f) = (\mathbf{P}^T \otimes \mathbf{A}(f))$ with respect to velocity v and polarization angle β to be zero in order to achieve a maximally flat response of the filter over a region around the peaks (or nulls) of the beamformer response. Therefore, the matrix of constraints $\mathbf{K}_{l,\text{der}}$ has the form

$$\mathbf{K}_{l,\text{der}} = \left[\mathbf{P}^T \otimes \mathbf{A}, \frac{d}{dv}\left(\mathbf{P}^T \otimes \mathbf{A}\right), \frac{d}{d\beta}\left(\mathbf{P}^T \otimes \mathbf{A}\right)\right] \tag{58}$$

$$= \left[\mathbf{P}^T \otimes \mathbf{A}, \mathbf{P}^T \otimes \frac{d\mathbf{A}}{dv}, \frac{d\mathbf{P}^T}{d\beta} \otimes \mathbf{A}\right] \tag{59}$$

and the desired matrix response $\mathbf{F}_{l,der}$ imposes zeros in correspondence to the derivatives of the array manifold $\mathbf{H}(f)$

$$\mathbf{F}_{l,der} = \begin{bmatrix} \mathbf{F}_l \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{60}$$

Other methods could be used for robust beamforming such as the optimization in the worst case scenario [42], [43] which would allow the use of velocity/polarization constraints in the form of inequality, thus simplifying the choice of the wavefield shape-parameters for the separation of seismic waves.

VI. Performance Analysis and Experimental Results

In this section, we analyze the performance of the estimation algorithm proposed in Section IV in terms of accuracy (Section VI-A) as well as robustness of estimation when sensor-calibration errors are introduced in the parametric model (Section VI-B). Moreover, in Section VI-C, we show examples on simulated and field data to support the applicability of the approach proposed in this paper to estimate the shape parameters of seismic waves and to separate polarized wavefields.

A. Performance Analysis of the Estimation Algorithm

In this section, we use simulated data to evaluate the performance accuracy of the proposed method for the joint velocity/polarization estimation (Section IV), and we compare the results with the CRB derived in Section III-A.

The CRB for the slowness q and the ellipticity angle β have been shown to vary only with the geometry-dependent parameters; here, we consider three different acquisition geometry configurations. The first geometry configuration consists of $M = 10$ vector sensors spaced $\Delta = 20$ m apart, the second one has $M = 30$ sensors with $\Delta = 25$ m; and the third one has $M = 50$ sensors with $\Delta = 30$ m. The simulated two-wave-traveltime data have been generated with different shape parameters, namely the surface wave (*parameter set 1*) and the reflected volume wave (*parameter set 2*). The parameter set 1 is characterized by slowness $q = 1.2\times10^{-3}$ sec/m ($v \simeq 833$ m/sec) and elliptically polarized wavefield ($\gamma = \pi/18$ rad, $\beta = \pi/5$ rad, $\varepsilon \simeq 0.72$) that lies at the free surface of the medium ($\theta = 0$ rad) and is cross-line with the array of sensors ($\psi = \pi/2$ rad); parameter set 2 has $q = 9 \times 10^{-4}$ sec/m ($v \simeq 1110$ m/sec), $\gamma = \pi/36$ rad, $\beta \simeq 0$ rad ($\varepsilon \simeq 0$), $\theta = \pi/18$ rad, and $\psi = \pi/2$ rad. For both parameter sets, the sampling period is $T = 4\times10^{-3}$ s, the central frequency of the Ricker waveform is $f_c = 15$ Hz, and a zero-mean uncorrelated Gaussian noise with noise power σ^2 is added. The spatial smoothing coefficient is set to $M_1 = M/2$. Note that the geometry configuration parameters (M, Δ) and the sampling period T are deterministic and known, while the shape-parameters of the seismic waves (q, γ, β, θ, ψ) are deterministic and unknown.

Fig. 4 shows the root mean square error (RMSE) of the estimates versus SNR (SNR $= E/\sigma^2$, where E is the energy of the seismic waveform) for the three geometry configurations. The performance of the proposed method is compared with the square-root of the analytical CRB from Section III-A. For the

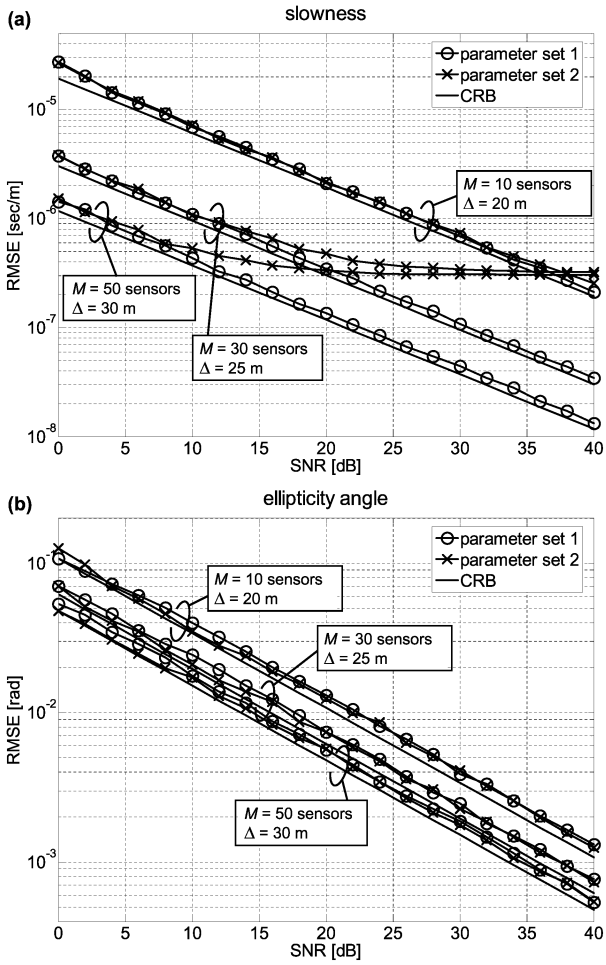

Fig. 4. RMSE versus SNR for (a) the slowness and (b) ellipticity angle estimates (lines with markers), compared with the square-root of the CRB (solid lines) for three different geometry configurations and two parameter sets of seismic polarized wavefields.

surface wave (*parameter set 1*), the RMSE of the slowness and ellipticity estimates (lines with circle markers) reaches the CRB (solid line) very closely across the entire SNR range and all geometry configurations. In particular, for the slowness estimate, Fig. 4(a), the lower bound is reached with a negligible loss (smaller than 1 dB in SNR). The ellipticity angle estimate, Fig. 4(b), experiences a loss of 2 dB in SNR with respect to the CRB, since in the shift-invariance method, the estimation of the ellipticity angle $\hat{\beta}$ (or ellipticity $\hat{\varepsilon}$) follows from the estimate of the polarization vector $\hat{\mathbf{p}}$, as shown in Section IV-A-1. Moreover, we note that, as expected, increasing the length of the array (in terms of number of sensors M and sensor inter-distance Δ) the performances improve in the case of the surface wave. In particular, for the slowness estimation Fig. 4(a), the passage from 10 to 30 sensors induces a gain of almost 16 dB in SNR, while from 30 to 50 sensors, the gain is 9 dB in SNR. On the other hand, for the ellipticity estimation Fig. 4(b), the gain improvements are smaller: 4 dB in SNR from 10 to 30 sensors and almost 2 dB in SNR from 30 to 50.

For the volume reflected wavefield (*parameter set 2*), the RMSE of the ellipticity estimates [lines with cross markers in Fig. 4(b)] reaches the CRB (solid line) very closely and

mostly overlaps the parameter set 1's ellipticity estimate across the entire SNR range and all geometry configurations. For the slowness estimates [lines with cross markers in Fig. 4(a)] the lower bound is reached with a negligible loss (smaller than 1 dB in SNR), for the geometry configuration with $M = 10$ sensors and $\Delta = 20$ m. However, when we increase the length of the array, the slowness estimates move away from the CRB. This happens because for reflected waves, the assumption about the linearity of the delay [in (8)] is not satisfied, because the traveltime is hyperbolic. So, the error due to the linearization of the traveltime increases with the array length. To overcome this limitation, for reflected waves, linear traveltime can be assumed locally, by reducing the length of the array aperture where the estimation is performed.

B. Performance Analysis of the Velocity/Polarization Estimation Algorithm in Case of Sensor-Calibration Errors

In this section, we evaluate the effects of sensor-calibration errors on the performance of the velocity/polarization estimation method proposed in Section IV, in order to demonstrate its robustness.

The simulated data generated for this analysis is a surface wave with added noise power σ^2 having the following characteristics: the slowness is $q = 10^{-3}$ sec/m ($v = 1000$ m/sec) and the elliptically polarized wavefield ($\beta = \pi/5$ rad, $\gamma = \pi/18$ rad) lies at the free surface of the medium ($\theta = 0$ rad) and is cross-line with the sensor array ($\psi = \pi/2$ rad). The array comprises M sensors spaced $\Delta = 25$ m apart with the same settings as Section VI-A: $T = 4 \times 10^{-3}$ s, $f_c = 15$ Hz, $M_1 = M/2$. As in the previous section, (M, Δ, T) are deterministic and known, while (q, γ, β, θ, ψ) are deterministic and unknown.

Fig. 5 is the RMSE of the estimates versus SNR with $M = 30$ sensors and varying positional sensor-calibration errors σ_d^2. The reference setting has no positional error ($\sigma_d^2 = 0$) with the performance as in Fig. 4, while the other configurations have positional errors with variance $\sigma_d^2 = 0.5$ m^2, 1 m^2, and 1.5 m^2. The performance of the estimation method is compared with the square-root of the analytical CRBs as derived in Section III-B, in the presence of sensor-calibration errors. From these results, we note that the RMSE of the slowness and ellipticity estimates (lines with markers) reaches the CRBs (solid lines) very closely across the entire SNR range and all geometry configurations. Moreover, for the ellipticity estimate, Fig. 5(b), we note that the loss of performance in the case of positional error with variance $\sigma_d^2 = 1.5$ m^2 is small (about 1 dB in SNR) with respect to the case with no positional errors; thus, we can conclude that the ellipticity estimation is independent of positional errors. This result was expected, since positional errors affect only the steering matrix $\mathbf{A}(f)$ (refer to Section II-C-I) that is independent on the polarization parameters. Instead, for the slowness estimate, Fig. 5(a), the performance worsens with increases in the positional calibration errors, even though the slowness estimates still attain the CRBs across the entire SNR range.

The same conclusions can be inferred from Fig. 6(a) and (b), where the RMSE of the estimates are plotted versus positional errors (with variance σ_d^2 ranging from 0 to 3m^2), varying the

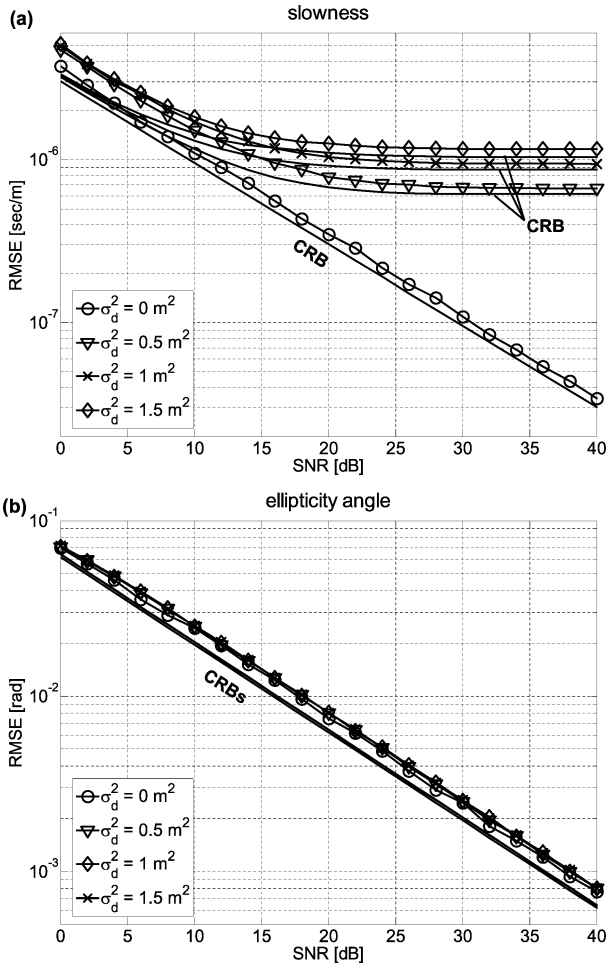

Fig. 5. RMSE versus SNR for the (a) slowness and (b) ellipticity angle estimates (lines with markers), compared with the square-root of the CRBs (solid lines) for different values of *positional* calibration errors.

length of the sensor array from $M = 10$ to 50 sensors. The performance of the estimation method (lines with markers) is compared with the square-root of the analytical CRBs (solid lines) from Section III-B ($SNR = 30$ dB for this figures). Note that both estimates improve when we increase the length of the array of sensors M.

Fig. 7 is the RMSE of the estimates versus SNR with $M = 30$ sensors and varying rotational sensor-calibration errors σ_α^2. The reference setting with no rotational error $(\sigma_\alpha^2 = 0)$ is as in Fig. 4; the others configurations have rotational errors with variance from $\sigma_\alpha^2 = 1.2 \times 10^{-3}$ rad^2 (4 deg^2) up to 3.6×10^{-3} rad^2 (12 deg^2). The performance of the estimation method is compared with the square-root of the analytical CRBs derived in Section III-B. Here, we note that the RMSE of the slowness and ellipticity estimates (lines with markers) approaches the CRBs (solid lines) very closely over the entire SNR range and across all the geometry configurations. Unlike in the previous analysis with positional calibration errors, the slowness estimates, Fig. 7(a), with rotational errors have small loss of performance (about 1 dB in SNR) with respect to the case with no rotational errors, while for the ellipticity estimates, Fig. 7(b), the performance worsens with increasing rotational calibration errors. This result, too, was expected,

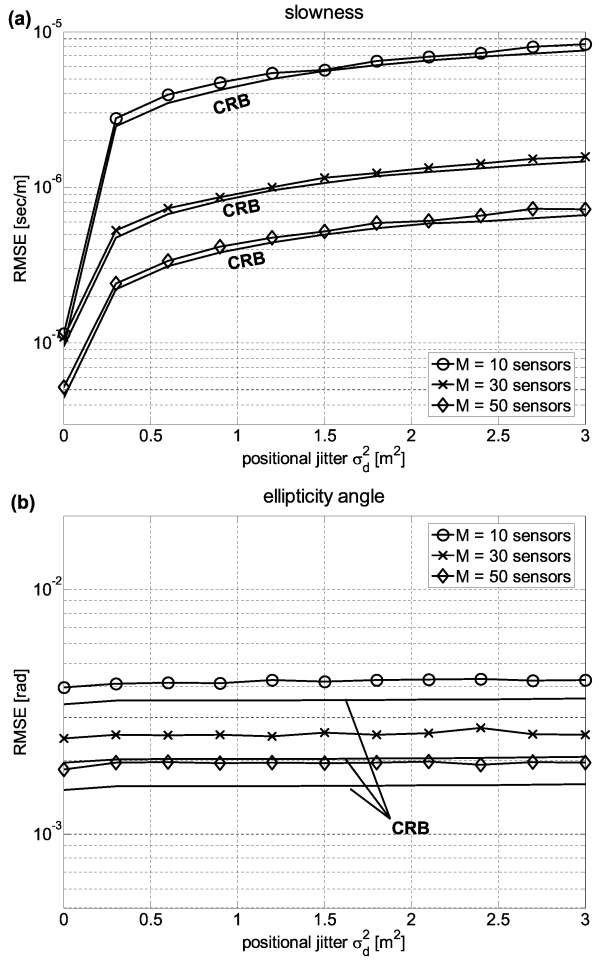

Fig. 6. RMSE versus *positional* errors for (a) the slowness and (b) ellipticity angle estimates (lines with markers), compared with the square-root of the CRBs (solid lines) for different sensor array lenghts M.

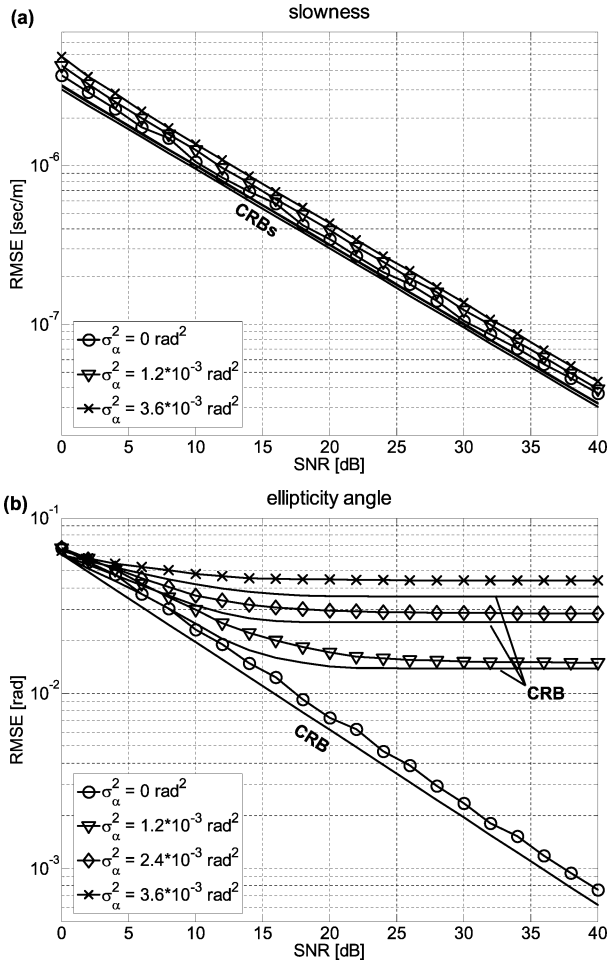

Fig. 7. RMSE versus SNR for (a) the slowness and (b) ellipticity angle estimates (lines with markers), compared with the square-root of the CRBs (solid lines) for different values of *rotational* calibration errors.

since rotational calibration errors dominate in the polarization matrix $\mathbf{P}$ (refer to Section II-C-II).

Referring to Fig. 8(a) and (b), we can conclude that the slowness estimation is independent of the rotational calibration errors. In these figures, the RMSE of the estimates are plotted versus rotational errors with variance σ_α^2 ranging from 0 to 4.8×10^{-3} rad^2 (0 to 16 deg^2), varying the length of the sensor array from $M = 10$ to 50 sensors, with $SNR = 30$ dB. The performance of the estimation method (lines with markers) is compared with the square-root of the analytical CRBs (solid lines) from Section III-B. Moreover, for the slowness estimates, Fig. 8(a), we note that performances improve when we increase the length of the array of sensors M. Instead, for the ellipticity estimates, Fig. 8(b), the estimation gains only 1 dB when we increase the length of the array from $M = 10$ to $M = 30$ sensors, and achieves the same gain of 1 dB from $M = 30$ to $M = 50$ sensors.

These numerical examples allow to understand the respective influence of the calibration errors on the parameters of interest. In summary, the slowness estimate is strongly influenced by the positional calibration errors and weakly by the rotational error; on the other hand, the ellipticity estimate is only affected by the rotational calibration errors.

C. Experimental Results

In this section, examples on simulated and experimental data show the applicability of the proposed methods to estimate the shape parameters of seismic waves and to separate polarized wavefields.

The three components $[y_x(t), y_y(t), y_z(t)]$ of the simulated data are shown in Fig. 9. For each component, on the vertical axis is the time and on the horizontal axis is the distance of each of the M sensors from the source. For each sensor position, the vertical trace represents the record of the ground motion (as a function of the time) at that sensor location, which is usually referred to as *seismogram*. The experiment consists of an in-line acquisition that produces two reflected volume waves and three surface waves. The Ricker waveform has been modeled using different frequencies [1] for surface waves ($f_c = 20$ Hz) and volume waves ($f_c = 30$ Hz). Sampling period is $T = 5 \times 10^{-3}$ s. The sensor array is made of $M = 40$ 3C geophones spaced $\Delta = 20$ m apart. Gaussian noise power σ^2 is added, with SNR $= 10$ dB. Moreover, we suppose array perturbations, with each sensor being affected by an angular rotation of the horizontal sensor components with variance $\sigma_\alpha^2 = 6 \times 10^{-3}$ rad^2.

Fig. 10 shows the velocity versus ellipticity estimates, obtained using the shift-invariance method, versus the frequency

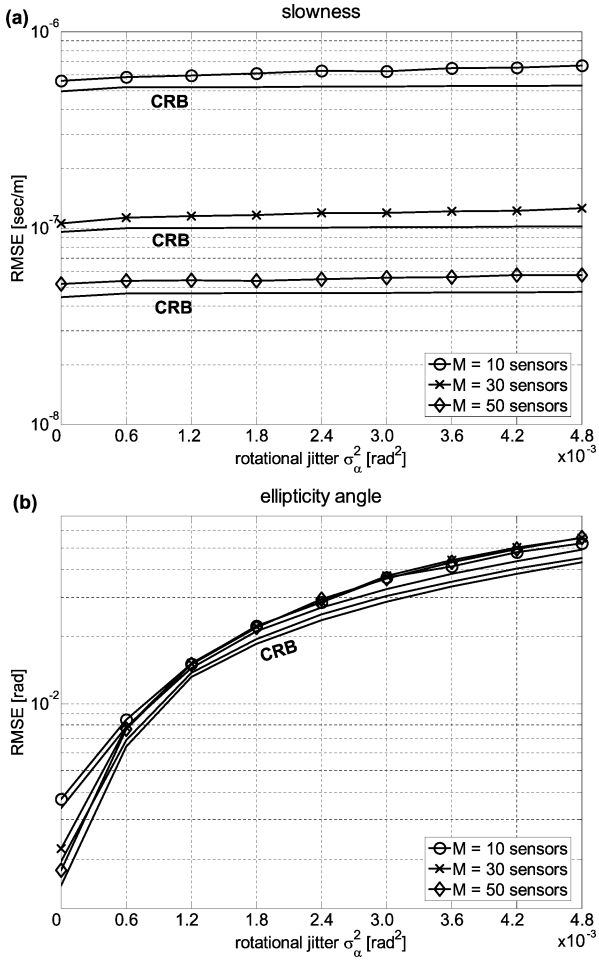

Fig. 8. RMSE versus *rotational* errors for (a) the slowness and (b) ellipticity angle estimates (lines with markers), compared with the square-root of the CRBs (solid lines) for different sensor array lengths M.

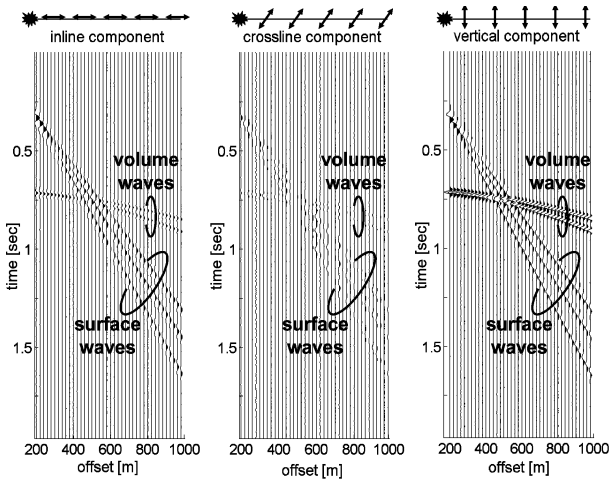

Fig. 9. Original simulated data with mispositioning: inline $\mathbf{y}_x(t)$ (left), crossline $\mathbf{y}_y(t)$ (center), and vertical $\mathbf{y}_z(t)$ (right) components.

by considering $d = 5$ wavefields. The projections of these estimates onto the velocity/ellipticity plane (the red points in Fig. 10) show that a simple method based on a velocity/polarization threshold is enough to separate surface waves from volume waves. Moreover, clustering techniques [38] can be used

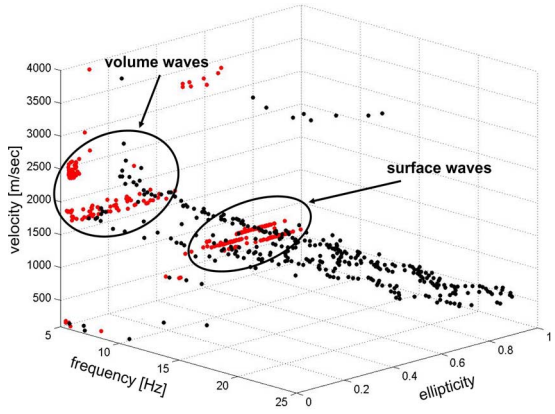

Fig. 10. Wavefield separation ($d = 5$) in the velocity versus ellipticity plane $\{\hat{v}_k, \hat{\varepsilon}_k\}$ for several frequency samples. Red points are the projection of the frequency-dependent estimates (in black) onto the velocity-ellipticity plane.

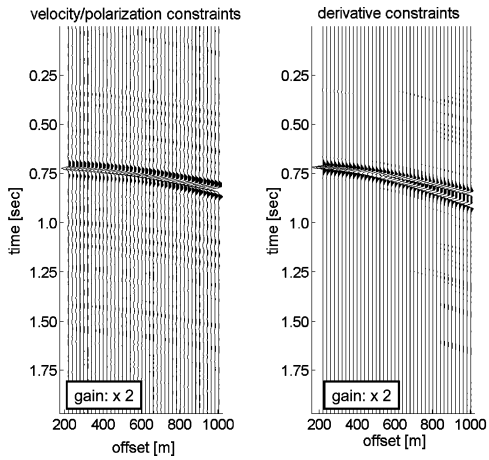

Fig. 11. Vertical component $\mathbf{y}_z(t)$ of the simulated data: after application of *velocity/polarization*-constrained beamformer (left) and after application of *derivative*-constrained beamformer (right).

to create a partition of the estimates and find the dimensions of the two subgroups. For the purpose of this paper we have used the k-means clustering technique [37]. Clustering approach can be further improved by constraining classification algorithms with the typical physical-patterns of seismic waves as an *a priori* (not considered in this paper).

The results after applying beamforming to the simulated data of Fig. 9 are in Fig. 11. The three surface wavefields have been totally removed. However, in the result after the application of the *velocity/polarization*-constrained beamformer (on the left) only one of the two volume waves is present. This happens because in an environment with mispositioning of the sensors, the estimation of the shape parameters suffers from higher uncertainty, thus affecting the construction of the filter whose main beam position does not coincide with the actual position of the wavefield to be preserved. The result after applying the *derivative*-constrained beamformer (on the right of Fig. 11) preserves the two volume waves as the beam pattern in correspondence of peaks and nulls is smoother, thus being less sensitive to errors in sensors' positioning.

The field-collected seismic data in Fig. 12 is a land measurement from an inline survey. The array of sensors is composed

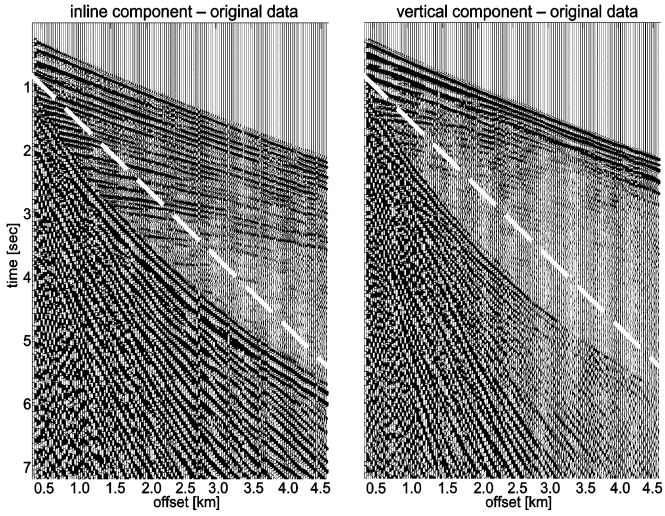

Fig. 12. Experimental seismic measurements: inline (left) and vertical (right) components of the original real data (surface waves are below the dashed line).

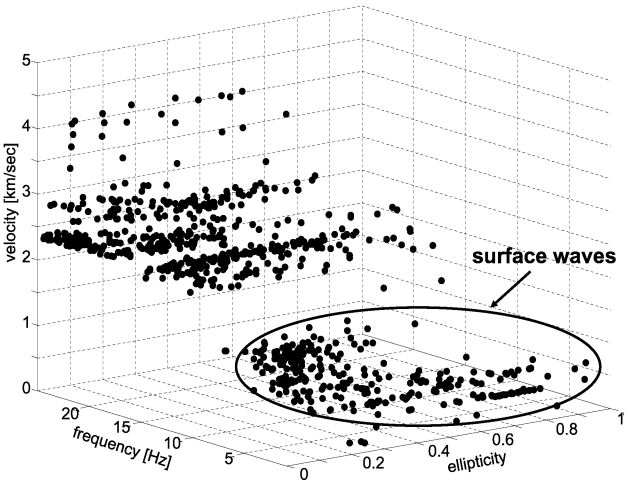

Fig. 13. Experimental seismic measurements: wavefield separation in the velocity versus ellipticity plane $\{\hat{v_k}, \hat{\varepsilon_k}\}$ for several frequency samples, ranging from 1 to 25 Hz.

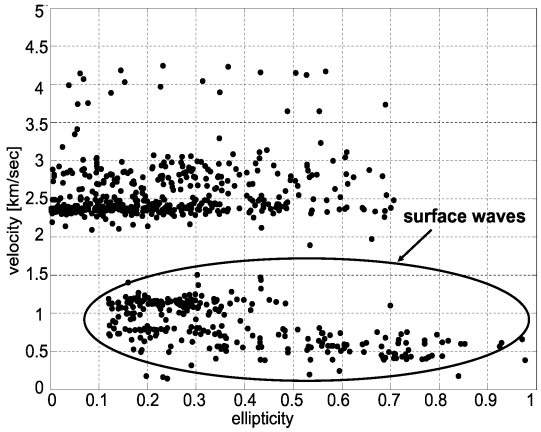

Fig. 14. Experimental seismic measurements: wavefield separation in the velocity versus ellipticity plane, where each point represents the projection of the frequency-dependent estimates (in Fig. 13) onto the velocity-ellipticity plane.

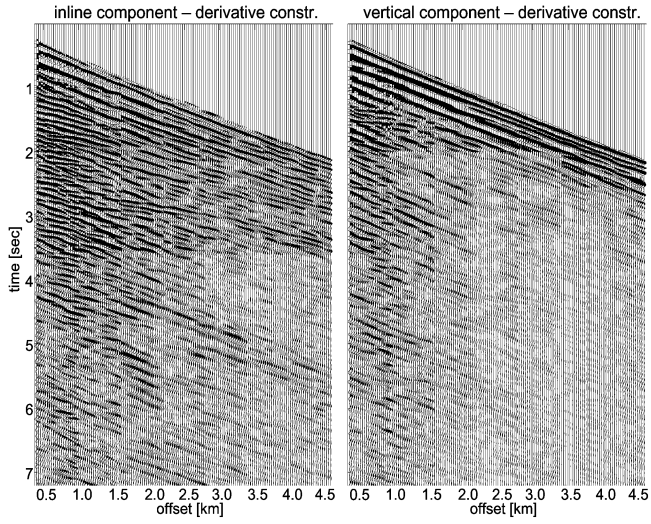

Fig. 15. Experimental seismic measurements: inline (left) and vertical (right) components of the real data after applying the *derivative*-constrained beamformer.

of $M = 140$ three-component sensors spaced $\Delta = 30$ m apart and sampling period $T = 4 \times 10^{-3}$ s. The data are corrupted by strong surface waves with velocities from 238 m/s to 834 m/s that basically saturate the data image. Other linear events with higher velocity (around 2100 m/s) are visible, and we can also see the reflected events we are interested in enhancing.

In Fig. 13, the velocity estimates are shown versus ellipticity for several frequency samples (from 1 to 25 Hz). Fig. 14 represents the projection of the frequency-dependent estimates of Fig. 13 onto the velocity-ellipticity plane to illustrate the effect of the clustering algorithm. Note that surface wave estimates can be reliably separated from volume wave estimates, even if these estimates are clearly more scattered when compared with the simulated data. Moreover, each estimate can be classified into a group employing clustering algorithms.

The result after application of the derivative-constrained beamformer is shown in Fig. 15. The surface waves have been totally removed from the central section of the recording for both the inline and vertical components. The volume waves result to be enhanced with respect to the original real data, as expected. However, some artifacts are visible due to a poor estimation of the parameters of interest and to the consequent beamforming filter that is inadequate to fully remove the interference. Of course, fine tuning of the filtering algorithm requires to set application specific constraint as part of the future activity.

VII. Conclusion

In this paper, we proposed a method to separate seismic surface waves by jointly exploiting the diversity of the polarization state and propagation velocity between wavefields (refer to the result in Fig. 14).

We presented a parametric model for the multicomponent wideband polarized signal in uncertain acquisition geometry and we analytically derived Cramér–Rao bounds for the seismic shape parameters. Our estimation method exploits the shift invariance of linear arrays to estimate and automatically pair the velocity and polarization parameters of the wavefields. We have

shown that this method is robust in the presence of both positional and rotational sensor-calibration errors, and it is effective to separate mixed seismic wavefields into their constituent wave modes. We proposed a velocity/polarization-constrained wideband beamformer to suppress the interference from surface waves and we introduced additional smoothing constraints to allow for reliable filtering when estimates are affected by uncertainty. It is clear from experimental testing that the constraint selection in the beamforming technique is the key step for obtaining reliable results with real data. Moreover, techniques for clustering of the velocity/polarization estimates is one of the open issues that needs to be further investigated.

Appendix A
CRB Derivations for the Physical Model With Known Acquisition Geometry

From (22), we can easily derive the Fisher information matrix $\mathbf{J}(\mathbf{\Psi})$ by employing the following relations:

$$\mathbf{p}^H\mathbf{p} = 1 \tag{61}$$

$$\mathbf{a}^H\mathbf{a} = M \tag{62}$$

$$\left(\frac{\partial \mathbf{a}}{\partial q}\right)^H \left(\frac{\partial \mathbf{a}}{\partial q}\right) = \phi^2 \frac{M(M-1)(2M-1)}{6} \tag{63}$$

$$\left(\frac{\partial \mathbf{a}}{\partial q}\right)^H \mathbf{a} = j\phi \frac{M(M-1)}{2}. \tag{64}$$

The Cramér–Rao matrix for the shape parameters $\mathbf{\Psi} = [q, X_R, X_I, \psi, \theta, \gamma, \beta]$, $\mathrm{CRB}(\mathbf{\Psi}) = \mathbf{J}^{-1}(\mathbf{\Psi})$, is found to be block-diagonal

$$\mathrm{CRB}(\mathbf{\Psi}) = \begin{bmatrix} \mathrm{CRB}_1(q, X_R, X_I) & \mathbf{0}_{3\times 4} \\ \mathbf{0}_{4\times 3} & \mathrm{CRB}_2(\psi, \theta, \gamma, \beta) \end{bmatrix} \tag{65}$$

where $\mathrm{CRB}_1(q, X_R, X_I)$ and $\mathrm{CRB}_2(\psi, \theta, \gamma, \beta)$ are the nonzero block matrices, with $\mathrm{CRB}_1(q, X_R, X_I)$ being equal to

$$\begin{array}{ccc} q & X_R & X_I \\ \left[\begin{array}{c|c|c} \frac{6}{Q(M-1)\phi^2\rho_f} & -\frac{3X_I}{Q\phi\rho_f} & \frac{3X_R}{Q\phi\rho_f} \\ -\frac{3X_I}{Q\phi\rho_f} & \frac{\kappa|X(f)|^2}{2Q\rho_f} & -\frac{\zeta X_I X_R}{2Q\rho_f \cos^2 2\beta} \\ \frac{3X_R}{Q\phi\rho_f} & -\frac{\zeta X_I X_R}{2Q\rho_f \cos^2 2\beta} & \frac{\kappa|X(f)|^2}{2Q\rho_f} \end{array}\right] \end{array} \tag{66}$$

and $\mathrm{CRB}_2(\psi, \theta, \gamma, \beta)$ appearing in (67), as shown at the bottom of the page, with

$$\phi = 2\pi f \Delta \tag{68}$$

$$\rho_f = \frac{|X(f)|^2}{\sigma^2} \tag{69}$$

$$\kappa = M\left(\tan^2 2\beta + 5\right) + \tan^2 2\beta - 1 \tag{70}$$

$$\xi = \left(\frac{1}{2\cos^2 2\beta} + \frac{(1 - \cos 2\gamma \cos 2\beta)\sin^2\theta}{\cos^2\theta \sin^2 2\beta}\right) \tag{71}$$

$$\zeta = (M-1)(1 + 2\cos^2 2\beta) + 2\sin^2 2\beta \tag{72}$$

$$\mathrm{Q} = M(M+1). \tag{73}$$

Appendix B
CRB Derivations for the Physical Model With Uncertain Acquisition Geometry

Let the output signal from the vector-sensor array with calibration uncertainties be as in (29). The Cramér–Rao bounds for the set of model parameters $\mathbf{\Upsilon} = [\mathbf{\Psi}, \mathbf{d}, \boldsymbol{\alpha}]$ is $\mathrm{CRB}(\mathbf{\Upsilon}) = (\mathbf{J}_1(\mathbf{\Upsilon}) + \mathbf{J}_2(\mathbf{\Upsilon}))^{-1}$, where $\mathbf{J}_1(\mathbf{\Upsilon})$ and $\mathbf{J}_2(\mathbf{\Upsilon})$ are given by (34) and (33), respectively. The relevant steps for the derivation of the covariance matrix $\mathbf{C_z}(\mathbf{\Upsilon})$ and its derivatives are listed next.

The covariance matrix $\mathbf{C_z}(\mathbf{\Upsilon})$ for the model in (29) is

$$\mathbf{C_z}(\mathbf{\Upsilon}) = |X|^2 \mathbf{C}_{\boldsymbol{\mu}}(\mathbf{\Upsilon}) + \sigma^2 \mathbf{I}_{3M} \tag{74}$$

where $\mathbf{C}_{\boldsymbol{\mu}}(\mathbf{\Upsilon})$ is the covariance matrix of the array response manifold $\boldsymbol{\mu}(\mathbf{\Upsilon}) = \mathbf{p}_c(\mathbf{\Phi}, \mathbf{d}) \otimes \mathbf{a}_c(q, \boldsymbol{\alpha})$. After some algebra, we find

$$\mathbf{C}_{\boldsymbol{\mu}}(\mathbf{\Upsilon}) = \left(e^{j2\pi f q \sigma_d^2} - 1\right) \mathbf{\Gamma} \odot \left[\mathbf{p}\mathbf{p}^H \otimes \mathbf{a}\mathbf{a}^H\right] \tag{75}$$

where

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{-\sigma_\alpha^2}{4} & \frac{-\sigma_\alpha^2}{4} & \frac{(1-M)\left(2-\sigma_\alpha^2\right)}{2M} \\ \frac{-\sigma_\alpha^2}{4} & \frac{-\sigma_\alpha^2}{4} & \frac{(1-M)\left(2-\sigma_\alpha^2\right)}{2M} \\ \frac{(1-M)\left(2-\sigma_\alpha^2\right)}{2M} & \frac{(1-M)\left(2-\sigma_\alpha^2\right)}{2M} & 0 \end{bmatrix}. \tag{76}$$

The derivatives of $\mathbf{C_z}(\mathbf{\Upsilon})$ with respect to the calibration parameters $[\mathbf{d}, \boldsymbol{\alpha}]$ are all zeros

$$\frac{\partial \mathbf{C_z}}{\partial d_i} = \frac{\partial \mathbf{C_z}}{\partial \alpha_i} = 0 \tag{77}$$

$$\mathrm{CRB}_2(\psi, \theta, \gamma, \beta) = \begin{array}{cccc} \psi & \theta & \gamma & \beta \\ \left[\begin{array}{c|c|c|c} \frac{1-\cos 2\gamma \cos 2\beta}{M\cos^2\theta \sin^2 2\beta\rho_f} & \frac{\sin 2\gamma \cos 2\beta}{M\cos\theta \sin^2 2\beta\rho_f} & \frac{\sin\theta(1-\cos 2\gamma\cos 2\beta)}{M\cos^2\theta\sin^2 2\beta\rho_f} & 0 \\ \frac{\sin 2\gamma\cos 2\beta}{M\cos\theta\sin^2 2\beta\rho_f} & \frac{1+\cos 2\gamma\cos 2\beta}{M\sin^2 2\beta\rho_f} & \frac{\sin 2\gamma\cos 2\beta\tan\theta}{M\sin^2 2\beta\rho_f} & 0 \\ \frac{\sin\theta(1-\cos 2\gamma\cos 2\beta)}{M\cos^2\theta\sin^2 2\beta\rho_f} & \frac{\sin 2\gamma\cos 2\beta\tan\theta}{M\sin^2 2\beta\rho_f} & \frac{\xi}{M\rho_f} & 0 \\ 0 & 0 & 0 & \frac{1}{2M\rho_f} \end{array}\right] \end{array} \tag{67}$$

while the derivatives of $\mathbf{C_z}(\boldsymbol{\Upsilon})$ with respect to the parameters of interest $\boldsymbol{\Psi} = [q, X_R, X_I, \boldsymbol{\Phi}]$, with $\boldsymbol{\Phi} = [\psi, \theta, \gamma, \beta]$, are

$$\frac{\partial \mathbf{C_z}}{\partial q} = |X|^2 \left(j2\pi f \sigma_d^2 e^{j2\pi f q \sigma_d^2} \right) \boldsymbol{\Gamma} \odot \left[\mathbf{p}\mathbf{p}^H \otimes \mathbf{a}\mathbf{a}^H \right] + |X|^2 \left(e^{j2\pi f q \sigma_d^2} - 1 \right) \boldsymbol{\Gamma} \odot \left[\mathbf{p}\mathbf{p}^H \otimes \frac{\partial (\mathbf{a}\mathbf{a}^H)}{\partial q} \right] \tag{78}$$

$$\frac{\partial \mathbf{C_z}}{\partial X_R} = 2X_R \left(e^{j2\pi f q \sigma_d^2} - 1 \right) \boldsymbol{\Gamma} \odot \left[\mathbf{p}\mathbf{p}^H \otimes \mathbf{a}\mathbf{a}^H \right] \tag{79}$$

$$\frac{\partial \mathbf{C_z}}{\partial X_I} = 2X_I \left(e^{j2\pi f q \sigma_d^2} - 1 \right) \boldsymbol{\Gamma} \odot \left[\mathbf{p}\mathbf{p}^H \otimes \mathbf{a}\mathbf{a}^H \right] \tag{80}$$

$$\frac{\partial \mathbf{C_z}}{\partial \Phi_i} = |X|^2 \left(e^{j2\pi f q \sigma_d^2} - 1 \right) \boldsymbol{\Gamma} \odot \left[\frac{\partial (\mathbf{p}\mathbf{p}^H)}{\partial \Phi_i} \otimes \mathbf{a}\mathbf{a}^H \right]. \tag{81}$$

We note that the first matrix term in (34) is block-diagonal, with the only nonzero block being the 7×7 block-matrix related to the parameters of interest $\boldsymbol{\Psi}$.

Appendix C
Proof of (44)

Let the complex-valued estimated kth polarization vector be $\hat{\mathbf{p}}_k = \mathbf{p} = \mathbf{p}_r + j\mathbf{p}_r$, where $\mathbf{p}_r$ and $\mathbf{p}_i$ are real valued. We want to determine the complex rotation $e^{j\varphi}$ such that the real and imaginary components of $\mathbf{p}' = e^{j\varphi}\mathbf{p}$ are orthogonal and such that

$$\left\langle \mathrm{Re}(\mathbf{p}e^{j\varphi}), \mathrm{Im}(\mathbf{p}e^{j\varphi}) \right\rangle = \langle \mathbf{p'}_r, \mathbf{p'}_i \rangle = 0 \tag{82}$$

where $\mathbf{p'}_r$ and $\mathbf{p'}_i$ are the real and the imaginary parts of $\mathbf{p}'$, defined as

$$\begin{cases} \mathbf{p'}_r = \mathrm{Re}(\mathbf{p}e^{j\varphi}) = \mathbf{p}_r \cos\varphi - \mathbf{p}_i \sin\varphi \\ \mathbf{p'}_i = \mathrm{Im}(\mathbf{p}e^{j\varphi}) = \mathbf{p}_r \sin\varphi + \mathbf{p}_i \cos\varphi. \end{cases} \tag{83}$$

The scalar product $\langle \mathbf{p'}_r, \mathbf{p'}_i \rangle$ of the real and imaginary components of $\mathbf{p}'$ can be written as

$$\langle \mathbf{p}'_r, \mathbf{p}'_i \rangle = \frac{\sin 2\varphi}{2} (||\mathbf{p}_r||^2 - ||\mathbf{p}_i||^2) + \cos 2\varphi \langle \mathbf{p}_r, \mathbf{p}_i \rangle. \tag{84}$$

After nullifying the previous expression, the solution to the problem (82) is given by

$$\tan(2\varphi) = -\frac{2\langle \mathbf{p}_r, \mathbf{p}_i \rangle}{\| \mathbf{p}_r \|^2 - \| \mathbf{p}_i \|^2} \quad \text{if} \quad \| \mathbf{p}_r \| \neq \| \mathbf{p}_i \| \tag{85}$$

$$\varphi = \frac{\pi}{4} \quad \text{if} \quad \| \mathbf{p}_r \| = \| \mathbf{p}_i \| \tag{86}$$

References

[1] O. Yilmaz, *Seismic Data Analysis*. Tulsa, OK: Society of Exploration Geophysicists, 2001.

[2] D. Donno, A. Nehorai, and U. Spagnolini, "Velocity/polarization estimation of seismic waves and polarized wavefield separation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'06)*, May 2006, vol. 4, pp. 1181–1184.

[3] T. J. Ulrych, M. D. Sacchi, and J. M. Graul, "Signal and noise separation: Art and science," *Geophys.*, vol. 64, pp. 1648–1656, Sep. 1999.

[4] A. Roueff, J. Chanussot, and J. Mars, "Efficient oblique polarization filter with cross scalogram," *SEG Expanded Abstracts*, vol. 21, pp. 1069–1072, 2002.

[5] R. de Franco and G. Musacchio, "Polarization filter with singular value decomposition," *Geophys.*, vol. 66, pp. 932–938, May 2001.

[6] J. Picheral, "High Resolution Methods for Joint Angle and Delay Estimation," Ph.D. dissertation, Univ. de Paris XI—SUPELEC, Paris, France, 2003.

[7] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. New York: Wiley-IEEE Press, 2001.

[8] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.

[9] B. Hochwald and A. Nehorai, "Polarimetric modeling and parameter estimation with applications to remote sensing," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1923–1935, Aug. 1995.

[10] I. Ziskind and M. Wax, "Maximum-likelihood localization of diversely polarized sources by simulated annealing," *IEEE Trans. Antennas Propag.*, vol. 38, no. 7, pp. 1111–1114, Jul. 1990.

[11] D. Rahamim, J. Tabrikian, and R. Shavit, "Source localization using vector sensor array in a multipath environment," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3096–3103, Nov. 2004.

[12] E. Ferrara and T. Parks, "Direction finding with an array of antennas having diverse polarization," *IEEE Trans. Antennas Propag.*, vol. AP-31, no. 3, pp. 231–236, Mar. 1983.

[13] K. T. Wong and M. D. Zoltowski, "Self-initiating MUSIC direction finding and polarization estimation in spatio-polarizational beamspace," *IEEE Trans. Antennas Propag.*, vol. 48, no. 8, pp. 1235–1245, Aug. 2000.

[14] J. Li and R. T. Compton, "Angle and polarization estimation using ESPRIT with a polarization sensitive array," *IEEE Trans. Antennas Propag.*, vol. 39, no. 9, pp. 1376–1383, Sep. 1991.

[15] K. T. Wong and M. D. Zoltowski, "Closed-form direction finding and polarization estimation with arbitrarily spaced electromagnetic vector-sensors at unknown locations," *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 671–681, May 2000.

[16] M. D. Zoltowski and K. T. Wong, "ESPRIT-based 2-D direction finding with a sparse array of electromagnetic vector-sensors," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2195–2204, Aug. 2000.

[17] K. T. Wong and M. D. Zoltowski, "Uni-vector-sensor ESPRIT for multi-source azimuth-elevation angle estimation," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1467–1474, Oct. 1997.

[18] H. L. van Trees, *Detection, Estimation, and Modulation Theory, Part IV*. New York: Wiley, 2001.

[19] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 8, pp. 823–831, Aug. 1985.

[20] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations—Part I: Far-field sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 286–299, Mar. 1987.

[21] J. Pierre and M. Kaveh, "Experimental performance of calibration and direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'91)*, Apr. 1991, vol. 2, pp. 1365–1368.

[22] S. Anderson and A. Nehorai, "Analysis of a polarized seismic wave model," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 379–386, Feb. 1996.

[23] W. H. Cho and T. W. Spencer, "Estimation of polarization and slowness in mixed wavefields," *Geophys.*, vol. 57, pp. 805–814, Jun. 1992.

[24] T. H. Keho, M. N. Toksos, C. H. Cheng, and R. M. Turpening, "Wave dynamics in a gulf coast VSP," in *Handbook of Geophysical Exploration*. St Gély du Fesc, France: Geophysical Press, 1984, vol. 14B, pp. 205–235.

[25] R. E. Sheriff, *Encyclopedic Dictionary of Applied Geophysics*. Tulsa, OK: Society of Exploration Geophysicists, 2002.

[26] P. L. Stoffa, M. K. Sen, R. K. Seifoullaev, R. C. Pestana, and J. T. Fokkema, "Plane-wave depth migration," *Geophys.*, vol. 71, pp. 261–272, Nov. 2006.

[27] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sep. 1978.

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

[29] S. L. Collier, "Fisher information for a complex gaussian random variable: beamforming applications for wave propagation in a random medium," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4236–4248, Nov. 2005.

[30] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[31] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameter via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[32] R. Moses and P. Stoica, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.
[33] G. H. Golub and C. van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
[34] L. J. Gleser, "Estimation in a multivariate 'errors in variables' regression model: Large sample results," *Ann. Statist.*, vol. 9, pp. 24–44, 1981.
[35] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 806–811, Aug. 1985.
[36] N. Ricker, "The form and laws of propagation of seismic wavelets," *Geophys.*, vol. 18, pp. 10–40, Jan. 1953.
[37] J. A. Hartigan and M. A. Wong, "A k-means clustering algorithm," *Appl. Statist.*, vol. 28, pp. 100–108, 1979.
[38] J. A. Hartigan, *Clustering Algorithms*. New York: Wiley, 1975.
[39] B. D. van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4–24, Apr. 1988.
[40] S. Shahbazpanahi, A. B. Gershman, Z. Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, pp. 2257–2269, Sep. 2003.
[41] M. H. Er and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 12, pp. 1378–1393, Dec. 1983.
[42] S. Vorobyov, A. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–323, Feb. 2003.
[43] S. Shahbazpanahi, S. Valaee, and A. B. Gershman, "A covariance fitting approach to parametric localization of multiple incoherently distributed sources," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 592–600, Mar. 2004.

Daniela Donno (S'06) received the double M.Sc. degrees in electrical and telecommunication engineering from the University of Illinois, Chicago, and the Politecnico di Milano, Milan, Italy, in 2005. She is currently pursuing the Ph.D. degree at the Politecnico di Milano.

From November 2007 to February 2008, she had a research internship with CGGVeritas in Massy, France. Her current research focuses on signal processing with applications to geophysics and, more specifically, parameter estimation, noise attenuation, and wavefield separation.

Ms. Donno received the IEEE ICASSP Best Student Paper Award in 2006.

Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion-Israel Institute of Technology, Haifa, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

From 1985 to 1995, he was a faculty member with the Department of Electrical Engineering, Yale University, New Haven, CT. In 1995, he joined the Department of Electrical Engineering and Computer Science at The University of Illinois at Chicago (UIC) as a Full Professor. From 2000 to 2001, he was Chair of the department's Electrical and Computer Engineering (ECE) Division, which then became a new department. In 2001, he was named University Scholar of the University of Illinois. In 2006, he became Chairman of the Department of Electrical and Systems Engineering at Washington University, St. Louis. He is the inaugural holder of the Eugene and Martha Lohman Professorship and the Director of the Center for Sensor Signal and Information Processing (CSSIP) at WUSTL since 2006.

Dr. Nehorai was Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2002. From 2003 to 2005, he was Vice President (Publications) of the IEEE Signal Processing Society, Chair of the Publications Board, member of the Board of Governors, and member of the Executive Committee of this Society. From 2003 to 2006, he was the founding editor of the special columns on Leadership Reflections in the *IEEE Signal Processing Magazine*. He was corecipient of the IEEE Signal Processing Society (SPS) 1989 Senior Award for Best Paper with P. Stoica, coauthor of the 2003 Young Author Best Paper Award and corecipient of the 2004 Magazine Paper Award with A. Dogandzic. He was elected Distinguished Lecturer of the IEEE SPS for the term 2004 to 2005 and received the 2006 IEEE SPS Technical Achievement Award. He is the Principal Investigator of the new multidisciplinary university research initiative (MURI) project entitled Adaptive Waveform Diversity for Full Spectral Dominance. He has been a Fellow of the Royal Statistical Society since 1996.

Umberto Spagnolini (SM'03) received the Dott.Ing. Elettronica degree (*cum laude*) from the Politecnico di Milano, Milan, Italy, in 1988.

Since 1988, he has been with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, where he is Full Professor in Telecommunications. He is the cofounder of WiSyTech (Wireless System Technology), a spinoff company of Politecnico di Milano on Software Defined Radio. His general interests are in the area of statistical signal processing. The specific areas of interest include channel estimation and space–time processing for wireless communication systems, parameter estimation/tracking, and wavefield interpolation applied to UWB radar, geophysics, and remote sensing.

Dr. Spagnolini served (1999–2006) as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.