



HAL
open science

Simultaneous Seismic Sources Separation Based on Matrioshka Orthogonal Matching Pursuit, Application in Oil and Gas Exploration

Ekaterina Shipilova, Michel Barret, Matthieu Bloch, Jean-Luc Boelle,
Jean-Luc Collette

► **To cite this version:**

Ekaterina Shipilova, Michel Barret, Matthieu Bloch, Jean-Luc Boelle, Jean-Luc Collette. Simultaneous Seismic Sources Separation Based on Matrioshka Orthogonal Matching Pursuit, Application in Oil and Gas Exploration. IEEE Transactions on Geoscience and Remote Sensing, 2020, pp.1-18. 10.1109/TGRS.2019.2959650 . hal-02625829

HAL Id: hal-02625829

<https://centralesupelec.hal.science/hal-02625829v1>

Submitted on 26 May 2020

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

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

Copyright

Simultaneous Seismic Sources Separation Based on Matrioshka Orthogonal Matching Pursuit, Application in Oil and Gas Exploration

Ekaterina Shipilova, Michel Barret¹, Matthieu Bloch², Jean-Luc Boelle, and Jean-Luc Collette

Abstract—We present Matrioshka orthogonal matching pursuit (OMP), a method consisting of two nested OMPs for separating seismic sources at an early stage of the signal processing chain. Matrioshka OMP is based on models of sensor signals that place nonrestrictive assumptions on the seismic survey using simultaneous sources. Our seismic event model is based on the spatial coherence of signals, which results in a straight or slightly curved feature in the trace representation of the data with a specific wavelet, whose magnitude can linearly vary according to the offset. We demonstrate the effectiveness of the approach on synthetic and real data.

Index Terms—Acquisition, matching pursuit, optimization methods, seismic signal processing, sources separation.

I. INTRODUCTION

A. Simultaneous-Source Seismic Acquisition

SEISMIC surveys are performed at all stages of oil and gas exploration and development, with the objective of constructing an image of the subsurface without actually penetrating into the Earth's crust. To obtain such an image, seismic sources generate a wavefield at or close to the surface, which then propagates into the subsurface where it is altered and reflected by the geological layers and bodies. The geological medium only absorbs some of the emitted energy and the remaining energy escapes and reaches the surface, where seismic receivers sensitive to minute vibrations record it. With some assumptions regarding the propagation velocities, the knowledge of the emission and detection time instants, as well as the spatial positions of the sources and receivers, provides information about the subsurface geometry and physical properties.

When simultaneous sources emit their signals, or when a single source emits a long signal or makes small pauses

between subsequent short shots, one must be able to separate the different sources and the different shots to identify the exact time of emission associated with each seismic event encountered. Since the crosstalk (effect of pollution of one signal by another) between shots significantly complicates the signal processing and eventually degrades image quality [1], conventional seismic surveys ensure that the time and location intervals between shots are large enough to avoid crosstalk. Nevertheless, simultaneous-source seismic data acquisition has recently attracted attention for its potential to acquire larger amounts of data in a reduced time [2], which might be beneficial in harsh meteorological environments [3] or because of environmental regulations.

The idea of allowing multiple seismic sources to fire simultaneously was first introduced in the seventies for marine and land seismic [4], [5], but the first simultaneous shooting was only implemented for land vibroseis acquisition in the late nineties [6] by controlling the pattern of the sources, also known as source sweeping. Since then, vibratory seismic source techniques have constantly improved, and sweep generation and management is still actively ongoing [7]–[9]. The first proposal of simultaneous shooting without constraints on the source pattern, i.e., without specific encoding or sweep management, dates back to the late nineties [10], but actual implementation of appropriate logistics, survey design, and processing has taken nearly a decade. The difficulty may be attributed in part to the dithering of shooting times required for best wavefield separation, which could result in complex real-time communication and synchronization of the sources in the field [11]. It was not until 2006 that BP proposed a new approach called independent simultaneous sourcing (ISS¹), in which no effort is made to synchronize the sources [12], and the burden is placed on the receiver to process a continuous recording. Subsequent published tests on synthetic and real data [13] have established the usefulness and potential of data acquired in simultaneous-source mode [14]; however, these early tests did not exploit any specific processing and only relied on the noise attenuation capacity of stacking for crosstalk suppression. Moreover, subsequent full scale surveys were only held at the exploration stage, in zones where structural interpretation was needed [15]–[18]. To speed up seismic campaigns, industry is now envisioning the use of simultaneous shooting at all exploration and development stages, including those having reservoir characterization [19],

Manuscript received February 8, 2019; revised July 25, 2019 and November 12, 2019; accepted November 21, 2019. (Corresponding author: Michel Barret.)

Ekaterina Shipilova and Jean-Luc Boelle are with TOTAL, CSTJF, 64018 Pau, France (e-mail: ekaterina.shipilova@total.com; jean-luc.boelle@total.com).

Michel Barret is with CentraleSupélec, 57070 Metz, France, and also with the UMI 2958 Georgia Tech-CNRS, 57070 Metz, France (e-mail: michel.barret@centralesupelec.fr).

Matthieu Bloch is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA, and also with the UMI 2958 Georgia Tech-CNRS, 57070 Metz, France (e-mail: matthieu.bloch@ece.gatech.edu).

Jean-Luc Collette is with CentraleSupélec, 57070 Metz, France (e-mail: jean-luc.collette@centralesupelec.fr).

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

Digital Object Identifier 10.1109/TGRS.2019.2959650

¹ISS is a registered trade mark of BP p.l.c.

[20] and monitoring [21]–[24] purposes. Consequently, more sophisticated processing is needed to achieve the high precision necessary at these stages. Seismic processing engineers could also benefit from simultaneous sources separation methods to suppress interference between neighboring seismic surveys [25].

The seismic data resulting from simultaneous-source shooting is now colloquially known as *blended data*. The methods proposed to process blended seismic data can be classified into the three following groups. Most of them impose constraints on data acquisition, as the firing times of different sources must be random enough, except for some completely different techniques, such as seismic apparition [26] or coherent simultaneous shooting [27].

- 1) *Random Noise Attenuation*: These methods consist in presenting the signal coming from each source as being coherent and removing the incoherent signals coming from the other sources using conventional denoising procedures such as median filtering [28]–[30], prediction error filtering [31], or a combination of these filters [32].
- 2) *Inversion-Based Source Separation*: These methods treat the signals of each source as signal and not as noise; they aim to explain all interpretable signals present in the data. The inversion approach has been the most successful so far and has been shown to render superior performance over the random noise attenuation [33]. The common mathematical formulation of such methods is given in [34]; this includes a coherency constraint usually expressed in a transform domain rather than the (t, x) domain, e.g., Fourier domain [35]–[37], Radon domain [38], curvelet domain [39], seislet domain [40], or in more sophisticated domain combinations [41]–[43]. The method presented in this article falls into this category.
- 3) *Direct Imaging of Blended Data*: These methods aim at processing blended data without explicit separation [44]. This approach has the potential to reduce computational complexity, since explicit source separation typically increases the amount of data volumes to process: a separate data set is created for each source.

Direct imaging methods might be the most promising in the future but would require a complete and costly overhaul of the currently used seismic signal processing chain. Since existing state-of-the-art industrial seismic processing algorithms are not compatible with blended data, there is still much interest in deblending the raw seismic signals to keep the subsequent processing unchanged.

Many of the currently proposed deblending methods need some preprocessing of the data, e.g., surface wave suppression [33]. In this article, we propose to use a data-driven seismic event model in a greedy decomposition to obtain a separation suitable for application at the earliest processing stages. We start by briefly recalling in Section II notions of seismicity from a perspective that facilitates the description of our method and identifies the conditions required for our method to apply. We then introduce in Section II-C the data-driven parametric model of a seismic event, which includes a curvature parameter and a magnitude attenuation factor depending on the position

of the source. We describe in Section III the main contributions of this article, which are the data-driven model and the decomposition method implementation. We also state in Sections II and III the assumptions that are necessary and sufficient to apply our method. Finally, we illustrate in Section IV the performance of our method on synthetic and real seismic data.

B. Signal Decomposition and Orthogonal Matching Pursuit

Let the signal to decompose be $d(t) \in \mathcal{H}$, where \mathcal{H} is a Hilbert space, with inner product and Euclidean norm, respectively, defined by $\langle d, g \rangle = \int_{-\infty}^{+\infty} d(t)\bar{g}(t) dt$ (where \bar{g} is the complex conjugate of g) and $\|d\| = \langle d, d \rangle^{1/2}$. The inner product becomes $\langle d, g \rangle = \sum_{t \in \Delta_t, \mathbb{Z}} d(t)\bar{g}(t)$ after sampling with a period Δ_t . A *dictionary* \mathcal{D} is a subset $\{g_\gamma(t)\}_{\gamma \in \Omega} \subset \mathcal{H}$ comprised of unit-norm vectors indexed by a set $\Omega \subset \mathbb{R}^\nu$, with $\nu \in \mathbb{N}$. The elements of a dictionary are called *atoms*.

When decomposing a signal d into a linear combination of L \mathcal{D} -atoms, we look for a subset $\Gamma \subset \Omega$ of L elements and complex numbers $\{c_\gamma\}_{\gamma \in \Gamma}$ that lead to the smallest approximation error $\min_{\{\Gamma \subset \Omega: |\Gamma|=L\}} \min_{\{c_\gamma\}} \|d - \sum_{\gamma \in \Gamma} c_\gamma g_\gamma\|$. If we knew L and Γ *a priori*, we could solve this problem with least-squares methods. However, for such complex dictionaries as ours, we cannot fix L and Γ beforehand; we first have to choose an optimal set of atoms and then find a linear combination that best approximates the signal.

Greedy algorithms, such as matching pursuit (MP) [45] and orthogonal matching pursuit (OMP) [46], provide an efficient solution to that problem. It consists in constructing successive approximations of d by making orthogonal projections on elements of \mathcal{D} . Let us set $R^0 d = d$ and suppose that the $(\ell - 1)$ th order residue $R^{\ell-1} d$ is computed for $\ell \geq 1$. Then, $R^{\ell-1} d$ is decomposed into $R^{\ell-1} d = \langle R^{\ell-1} d, g_{\gamma_\ell} \rangle g_{\gamma_\ell} + R^\ell d$. This leads to $\langle R^{\ell-1} d, g_{\gamma_\ell} \rangle = \langle R^{\ell-1} d, g_{\gamma_\ell} \rangle \langle g_{\gamma_\ell}, g_{\gamma_\ell} \rangle + \langle R^\ell d, g_{\gamma_\ell} \rangle$, which shows that the residue $R^\ell d$ is orthogonal to g_{γ_ℓ} , since the atoms are unit-norm vectors. Hence, $\|R^{\ell-1} d\|^2 = |\langle R^{\ell-1} d, g_{\gamma_\ell} \rangle|^2 + \|R^\ell d\|^2$, and to minimize the norm of the residue $\|R^\ell d\|$, one must choose $g_{\gamma_\ell} \in \mathcal{D}$ such that

$$|\langle R^{\ell-1} d, g_{\gamma_\ell} \rangle| = \max_{\gamma \in \Omega} |\langle R^{\ell-1} d, g_\gamma \rangle|. \quad (1)$$

If we now carry the decomposition up to the L th order, we obtain: $d = \sum_{\ell=1}^L \langle R^{\ell-1} d, g_{\gamma_\ell} \rangle g_{\gamma_\ell} + R^L d$ and $\|d\|^2 = \sum_{\ell=1}^L |\langle R^{\ell-1} d, g_{\gamma_\ell} \rangle|^2 + \|R^L d\|^2$, which proves that the residue norm is decreasing. At this stage, the signal d is modeled as a finite linear combination of L atoms with an error $R^L d$. However, this model can be improved using the same dictionary, since $R^L d$ is generally not orthogonal to \mathbf{V}_L , the linear span of $\{g_{\gamma_\ell}\}_{1 \leq \ell \leq L}$: it is only orthogonal to the last selected atom g_{γ_L} . The OMP algorithm corrects this shortcoming by computing orthogonal projections at each iteration. As in MP, at iteration ℓ the algorithm selects an atom g_{γ_ℓ} solving (1), builds an orthogonal basis of \mathbf{V}_L using the Gram-Schmidt orthogonalization, and uses it to compute the orthogonal projection of d on \mathbf{V}_L . With OMP, the residue norm is also proven to decrease monotonically.

Greedy algorithms have already been applied to seismic data for several different purposes, such as filtering [47], linear

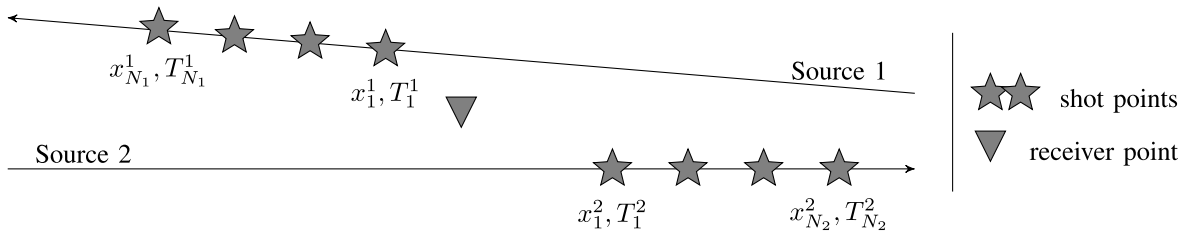


Fig. 1. Ocean Bottom Nodes (OBN) acquisition scheme for two seismic sources, where T_n^i and x_n^i denote the time and the coordinate of the n th shot of source i on the axis of its shooting line.

189 noise suppression [48], [49], seismic data interpolation and
 190 regularization [50], [51], seismic data compression and sparse
 191 storage [52], [53], or reflectivity inversion [54], but few, if any,
 192 contributions have considered their application to the problem
 193 of separating signals from different sources.

194 II. MODELING SENSOR SIGNALS IN SIMULTANEOUS 195 SOURCES SEISMIC SURVEY

196 We introduce our model of simultaneous-source seismic
 197 surveying and highlight the assumptions that justify our data-
 198 driven model for simple geometries of the Earth's subsurface.

199 A. Earth's Transfer Function

200 We consider an ocean bottom seismic acquisition with
 201 $K \geq 1$ sources $\{S_k\}_{1 \leq k \leq K}$ and a single sensor D located at
 202 fixed positions in the Earth's space-time referential. We denote
 203 the k th source excitation by $s_k(t)$, and the measured signal
 204 by $d(t)$, both time-dependent. The Earth acts as a filter
 205 (i.e., a linear, time-shift invariant, and continuous system) for
 206 the emitted signals $s_k(t)$, which enables us to represent the
 207 recorded signal as a convolution product $d(t) = \sum_{k=1}^K (r_k \star$
 208 $s_k)(t)$, with the Earth's response coefficients r_k depending on
 209 the positions of all the sources and the detector. Each source S_k
 210 makes N_k shots at times $\{T_n^k\}_{n \in \llbracket 1; N_k \rrbracket}$ and in the corresponding
 211 positions $\{\mathbf{x}_n^k\}_{n \in \llbracket 1; N_k \rrbracket}$. We also make the following hypothesis.

212 *Hypothesis 1:* Source S_k emits the same short excitation s_k
 213 for each of its shots.

214 Consequently, the signal $d(t)$ is given by

$$215 \quad d(\mathbf{x}_D, t) = \sum_{k=1}^K \sum_{n=1}^{N_k} (r_k(\mathbf{x}_n^k, \mathbf{x}_D) \star s_k)(t - T_n^k) + b(t) \quad (2)$$

216 where $b(t)$ is the additive noise capturing the unavoidable
 217 imperfections of real seismic acquisitions. Note that $r_k(\mathbf{x}_n^k, \mathbf{x}_D)$
 218 from (2) does not correspond to the true Earth reflectivity
 219 between \mathbf{x}_n^k and \mathbf{x}_D (the detector position) but acts as a
 220 transfer function between the source and receiver locations
 221 that accumulates Earth's entire response. Since the posi-
 222 tion of the detector is constant, we will write $d(t)$ instead
 223 of $d(\mathbf{x}_D, t)$.

224 B. Simultaneous Sources for Classical Seismic Survey Design

225 1) *Experimental Conditions for Simultaneous-Source Sur-*
 226 *veys:* We assume that each receiver continuously records all
 227 the seismic signals produced during the acquisition, which

228 requires that all the survey equipment be kept synchronized.
 229 Time ranges from 0 to T_{glob} , the global acquisition time.
 230 We make the following hypothesis.

231 *Hypothesis 2:* The sources fire along straight lines, which
 232 may differ for different sources.

233 As illustrated in Fig. 1, the moveout of each seismic event
 234 depends on the source location along its shooting line. After
 235 sampling with period Δ_t , the recorded data have the shape
 236 of a column matrix $d(k) = d(k \Delta_t)$. This type of recording
 237 is specific to simultaneous sources surveys. We further make
 238 two realistic hypotheses to simplify our analysis.

239 *Hypothesis 3:* Each source makes pauses between consec-
 240 utive shots, during which its emitted signal is null.

241 *Hypothesis 4:* Shooting times of different sources are asyn-
 242 chronous and shooting intervals of each source are random.

243 The benefit of Hypotheses 3 and 4 is illustrated in Fig. 2,
 244 in which we align the data according to the shooting times of
 245 different sources. Each shot of the same source can be distin-
 246 guished from the others following the time axis and the shots
 247 of different sources can be separated using a spatial coherence
 248 criterion—as detailed in Section II-C—which consists in a
 249 straight or slightly curved feature in the representation space
 250 (t, x) of the data.

251 2) (t, x) *Representation Spaces of the Data* : We define a
 252 linear operator, called pseudo-deblending, to align the sensor
 253 signal by the source i to form the traces (see Fig. 3). For $d(t)$
 254 from (2), it is written as $\mathcal{A}_i : L^2(\mathbb{R}) \rightarrow L^2([0, \max_n (T_{n+1}^i -$
 255 $T_n^i)]) \times [x_{\min}^i, x_{\max}^i]$, $d(t) \mapsto \mathbf{D}_i(t', x) = d(t' + T_n^i)$, if $x =$
 256 $x_n^i - x_0^i$ and $t' \in [0; T_{n+1}^i - T_n^i]$; $\mathbf{D}_i(t', x) = 0$, otherwise.
 257 Pseudo-deblending creates as many data representation spaces
 258 (t, x) as there are sources. To simplify, the (t, x) representation
 259 space is called in the following a (t, x) *trace domain*.

260 In conventional single source seismic, the operation \mathcal{A}_i
 261 is done implicitly: the data are cut into traces according to
 262 the shooting times T_n^i , which do not play any further role
 263 in the processing. In contrast, for *simultaneous-source* data
 264 processing, it is crucial to preserve the shooting times, as they
 265 contain critical information to separate the signals coming
 266 from different sources.

267 We introduce these notions to clarify the concept of a
 268 *seismic event*, used in our data-driven model and related to
 269 the notion of travelttime curve, the graph of the time that
 270 a seismic wave spends to travel from the shot point to the
 271 receiver point. Note that the knowledge of the firing times
 272 and positions of each source makes it easy to switch from a
 273 1-D representation to any 2-D trace domain representation of

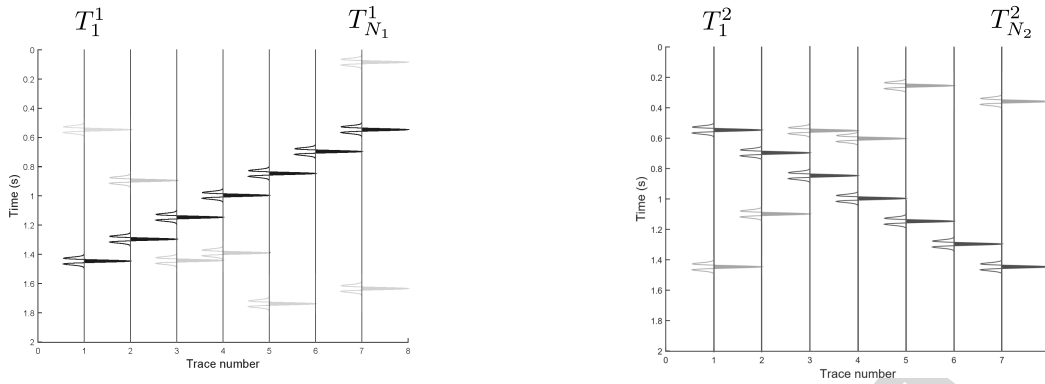


Fig. 2. Separation by random shot-times. Data aligned according to the (Left) first and (Right) second source shot-times. Dark (resp. light) wiggles correspond to the aligned (resp. nonaligned) source.

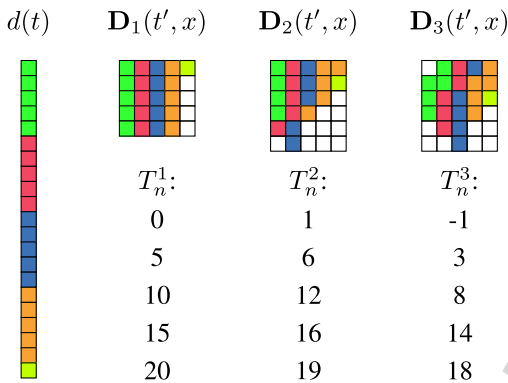


Fig. 3. Example of (t, x) representations of a continuous signal cut into traces with regular shooting times T_n^1 for the first source and irregular ones T_n^2 and T_n^3 for the second and the third sources, respectively. Zero padding (the white squares correspond to zeros) is applied in order to keep the matrices rectangular.

274 the signals and vice versa. Therefore, with a slight abuse of
275 language, we use the same terminology for events and patterns
276 in 1-D and 2-D representations, even though such events or
277 patterns are only clearly visible in 2-D representations.

278 *Hypothesis 5:* Traveltimes of coherent seismic waves
279 (e.g., direct waves, surface waves, and reflected waves) are
280 identifiable in one (and only one) seismic traces domain.

281 Traveltimes are usually close to straight lines, parabolas,
282 or hyperbolas in synthetic and real seismic data [55]. This
283 observation and Hypothesis 5 imply the possibility of decom-
284 posing $d(t)$ into a sum of a finite number of coherent features
285 that have a reasonably simple mathematical representation,
286 as we shall see in Section II-C.

287 C. Data-Driven Seismic Event Model

288 We now introduce our parametric model of a seismic
289 event that may either carry information about the Earth's
290 subsurface geometry or correspond to a direct arrival. This
291 model, which includes a curvature parameter, a magnitude
292 attenuation factor depending on the source positions and the
293 wavelet's decomposition into a sum of simple signals, and its
294 implementation are the main contributions of this article.

295 1) *Decomposition Into a Sum of Seismic Events:* Actual
296 seismic data usually have a significant size: one gather can

297 contain hundreds of traces acquired with maximal offsets
298 of 6 km or more. In complex geological environments with
299 lateral velocity and density variability, it is difficult to establish
300 a data-driven seismic model that would directly apply to the
301 whole gather. Therefore, we choose to restrict our area of
302 search to N seismic traces in the (t, x) domain, with N
303 typically between 10 and 30 depending on the data complexity.
304 This allows us to make the following reasonable hypothesis.

305 *Hypothesis 6:* The wavelet $w(t)$ found in the data does not
306 vary significantly from one seismic trace to another within
307 some constrained spatial window of N seismic traces.

308 When dealing with multiple sources recorded by the same
309 receiver, which results in multiple (t, x) trace domains to
310 consider, one must adopt a consistent decomposition strategy.

311 There may be several relevant ones, such as fully explaining
312 all coherent features in the first source before passing on to
313 the second one. If we were to follow this strategy, we would be
314 able to cut our data into traces once, using the \mathcal{A}_i operator for
315 each source i , and continue with the 2-D (t, x) trace domain
316 representation common for a geophysicist. This approach has
317 the following disadvantage: the algorithm aims at retrieving
318 the low-amplitude signal hidden by the high-amplitude blending
319 noise originating from the other sources. We propose to
320 *simultaneously* work in all the (t, x) trace domains in order to
321 first identify and subtract the globally most energetic features
322 and then continue with less energetic ones. The less energetic
323 features are initially hidden under the crosstalk but are revealed
324 by the first iterations of the algorithm. This is the main
325 reason why we stick to the 1-D representation of the data.
326 The decomposition is therefore simultaneously performed in
327 all the sources (t, x) trace domains, in which we look for
328 particular identifiable features that we call *seismic events*.
329 To do so, we first have to find in the column matrix $d(k\Delta t)$
330 the N -traces part of the signal corresponding to each of the
331 sources. We then represent the data $d(t)$ as a finite sum of
332 seismic events $h_\ell \star w_\ell(t)$

$$333 \quad d(t) = \sum_{\ell=1}^L h_\ell \star w_\ell(t) + R^L d(t). \quad (3)$$

334 Our model consists of two parts: $h(t)$, called *traveltimes curve*
335 (we call it curve because of the trace representation (t, x)
336 of 1-D signals), contains all the parameters related to the wave

337 propagation time (medium characteristics), distance between
 338 the sources and the receiver, the sources firing times and the
 339 linear amplitude variation from one trace to another; $w(t)$,
 340 called *signature* or *wavelet*, is associated with the excitations
 341 emitted by the sources and distorted by propagation and
 342 reflection. Note that even if (2) and (3) are similar, there
 343 is a significant difference between the reflectivity $r_i(\mathbf{x}_n^i, \mathbf{x}_D)$,
 344 which is Earth's transfer function between the locations of the
 345 detector and of the n th shot of source i , and the traveltimes
 346 curve $h(t)$, which indicates the position of a seismic event
 347 in the traces domain and is driven by the data. Moreover,
 348 the residue $R^L d(t)$ generally differs from the noise $b(t)$. In
 349 the case of two sources, we rewrite (3) as

$$350 \quad d(t) = \sum_{\ell=1}^{K_1} h_{\ell}^{(1)} \star w_{\ell}^{(1)}(t) + \sum_{\ell=1}^{K_2} h_{\ell}^{(2)} \star w_{\ell}^{(2)}(t) + R^L d(t) \quad (4)$$

351 with $K_1 + K_2 = L$ and where the first (resp. second) sum
 352 corresponds to the seismic events identifiable in the (t, x)
 353 traces domain of the first (resp. second) source. Thus, a perfect
 354 deblending would consist in reducing the residue $R^L d(t)$ to
 355 the ambient noise, as in this case, each sum would correspond
 356 to the isolated signal of the corresponding source. Before
 357 developing this point in Section III, we clarify the concepts of
 358 traveltimes curve and wavelet in Sections II-C2 and II-C3.

359 2) *Traveltimes Curve Model*: If we omit the amplitude vari-
 360 ation, a traveltimes curve is a graph of arrival time depending
 361 on the coordinates of the detector and the source shots. One
 362 can prove [55] that, for a simple case of a single horizontal
 363 reflector with a constant velocity above it, the traveltimes curve
 364 is a hyperbola. Furthermore, with reasonable accuracy, one can
 365 model the arrival time function of a coherent seismic wave as a
 366 straight or slightly curved line in the (t, x) trace domain within
 367 some lateral processing window (the closer the shot is to the
 368 receiver, the more curvature is observed). This assumption
 369 holds if the acoustic and elastic properties of the subsurface
 370 do not abruptly change in the horizontal direction within the
 371 chosen lateral processing window. The ‘‘pure’’ traveltimes part
 372 $\tilde{h}(t)$ of the seismic event takes the form

$$373 \quad \tilde{h}^{(i)}(x_n^i, t) = \delta \left(t - \tau - p(x_n^i - x_0^i) - q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 \right) \quad (5)$$

374 or, for the convenience of our computation, in 1-D

$$375 \quad \tilde{h}^{(i)}(t) = \sum_{n=1}^N \delta \left(t - \tau - p(x_n^i - x_0^i) - q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 - T_n^i \right). \quad (6)$$

377 Equations (5) and (6) are equivalent, but we stick to the
 378 1-D representation to highlight the specific nature of the
 379 simultaneous-source data. Note that we omit the index ℓ
 380 present in (3) to alleviate notation. Here, i is the index of
 381 the source associated with the event; N the number of shots
 382 taken into account to construct the event; x_0^i , x_{\min}^i and x_{\max}^i
 383 are the reference coordinates of the i th source; $\delta(t)$ is the Dirac
 384 distribution; τ , p , and q are the parameters that define the
 385 seismic event: the reference time, the slope and the curvature.

386 Finally, to obtain the full traveltimes curve, we add a linear
 387 amplitude variation parameter α to this representation and
 388 obtain

$$389 \quad h^{(i)}(t) = \sum_{n=1}^N [1 + \alpha(x_n^i - x_0^i)] \times \delta \left(t - \tau - p(x_n^i - x_0^i) - q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 - T_n^i \right). \quad (7)$$

391 Note that the attenuation factor $1 + \alpha(x_n^i - x_0^i)$ cannot van-
 392 ish when $x_n^i = x_0^i$ in (7). We shall see in the following
 393 (Section III-B and criterion (13)) how one can address such
 394 a case. Strictly speaking, (7) defines an amplitude-variation-
 395 preserving traveltimes curve, but for brevity we use the term
 396 *traveltimes curve* in the following. It is worth noting that differ-
 397 ent sources illuminating the same area in the subsurface, e.g.,
 398 an interface between two geological layers approximately at
 399 the same location, correspond to a single *physical* (geological)
 400 event; however, with our model (4), we obtain at least one
 401 separate *seismic* event per source. Moreover, even though our
 402 seismic event atoms correspond to simple cases of physical
 403 events, their linear combinations allow us to model complex
 404 physical situations (see Section IV).

405 3) *Wavelet Model*: Wavelet estimation has been a long-
 406 standing issue in seismic prospecting and different methods
 407 have been suggested in the literature. We focus on methods
 408 based on coherence; in other words, from Hypothesis 6,
 409 we assume that the wavelet does not abruptly change from
 410 trace to trace in a seismic event. This is intuitively justified by
 411 Hypothesis 1, and the fact that Earth's response to excitations
 412 varies slowly with respect to the source displacement. Never-
 413 theless, we take into account the eventual presence of low and
 414 high energy noise that may perturb the wavelet originating
 415 from a single source by averaging the wavelet encountered
 416 in neighboring traces after getting rid of eventual outliers.
 417 Since propagation and reflection distort the source signals,
 418 the wavelet encountered in seismic gathers differs from the
 419 signal emitted by the source, and we suppose that the wavelet
 420 differs from one seismic event to another, even if it originates
 421 from the same seismic source.

422 As already mentioned, a single physical event may be
 423 captured by a sum of several seismic events, so we look for
 424 a new seismic event within a limited time interval that we
 425 denote $[-M\Delta_t, M\Delta_t]$ and call ‘‘corridor’’ (M is an integer
 426 meta-parameter). This corridor is defined along a traveltimes
 427 curve of the form (7), which is assumed known for now.
 428 It must be large enough not to change the wavelet spec-
 429 trum. The fact that the traveltimes curve $h^{(i)}$ is not perfectly
 430 known is addressed in Sections III-B and III-D. After get-
 431 ting a first estimation $\hat{w}^{(i)}$ of the wavelet $w^{(i)}$, we refine
 432 the estimation thanks to optimization stages described in
 433 Section III-C2. To reduce the dimension of the optimization
 434 problem, we choose to decompose the estimated wavelet into
 435 a linear combination of a small number of wavelet atoms.
 436 We also choose wavelet atoms that can be represented analyt-
 437 ically, e.g., Ricker and Ormsby wavelets, which are elementary
 438 wavelets widely used in seismic exploration and for which

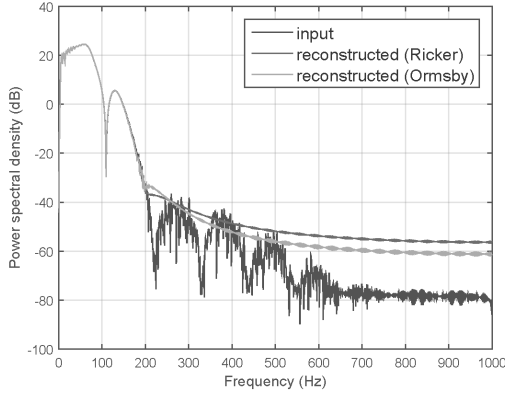


Fig. 4. Power spectrum of the signal of length 500 ms in black, reconstructed using 178 Ricker (in dark gray), or 212 Ormsby (in light gray), wavelets.

we can explicitly compute temporal derivatives. This allows us to reduce the computational complexity of our algorithm optimization stages.

We choose to decompose the estimated wavelet into a linear combination of a small number of wavelet atoms using the OMP algorithm, which requires the identification of an adapted dictionary. We shall see in Section III-C1, how we construct a finite number S (of several units) of classical wavelet shapes from a preliminary spectral analysis of the data. The index s denotes the shape of the wavelet $w_s(t)$, and the dictionary consists of atoms (before normalization) $\{w_s(t - \tau) : 1 \leq s \leq S, \tau \in [0, T]\}$ where $T > 0$ is a meta-parameter. Thus, we obtain the following parametric wavelet estimation:

$$\begin{aligned} \hat{w}^{(i)}(t) &= \sum_{k=1}^K a_k w_{s_k}(t - \tau_k) + R^K \hat{w}^{(i)}(t) \quad \text{and} \\ w^{(i)}(t) &= \sum_{k=1}^K a_k w_{s_k}(t - \tau_k). \end{aligned} \quad (8)$$

Fig. 4 shows the power spectrum of a modeled marine seismic source signature. We observe that the parametric model of the form (8) is accurate enough in the useful part of the spectrum both with Ricker ($K = 178$ for this example) and Ormsby ($K = 212$) wavelets. Note that these numbers are significantly larger than those used in our deblending algorithm because here the whole length of the source signal is taken into account (0.5 s) with a very dense sampling (0.5 ms of period). For subsequent simulations we use narrower corridors for wavelet estimation, typically 0.1 s, with $\Delta_t = 2$ ms.

III. MATRIOSHKHA OMP IMPLEMENTATION

We now present the implementation of our algorithm. In Section III-A, we present the fundamentals of our method, which we call *Matrioshka OMP* and which relies on detailed parameter optimization. To obtain suitable parameter values, we use iterative optimizations, which require a sufficiently accurate prior knowledge of the parameters, i.e., satisfactory initial conditions. The initial condition computation is described in Section III-B. Sections III-C and III-D provide an overview of the different parts of the algorithm.

A. Deblending Using Data-Driven Model and OMP

An iterative method that performs a decomposition as in (4) automatically results in a partial deblending of the data. Moreover, if the first terms in this decomposition correspond to the seismic events having the most energy, then only the lowest energy cross-talks are left in the residue $R^L d(t)$, which can then be handled by classical seismic processing techniques as if no other sources had been firing simultaneously. Consequently, we look for a decomposition (4) in which the most energetic features of the deblended signal associated with the i th source are found in the sum

$$\sum_{\ell=1}^{K_i} h_{\ell}^{(i)} \star w_{\ell}^{(i)} \quad (9)$$

and the most energetic cross-talks due to the other sources are captured in the sums

$$\sum_{\ell=1}^{K_j} h_{\ell}^{(j)} \star w_{\ell}^{(j)} \quad \text{with } j \neq i \quad (10)$$

to allow classical processing of the deblended data $\sum_{\ell=1}^{K_i} h_{\ell}^{(i)} \star w_{\ell}^{(i)} + R^L d(t)$. To successfully deblend with this approach, it is crucial that the sum (9) contains the most energetic features of the deblended signal associated with the source i and not any coherent seismic events originating from a source $j \neq i$. Hypotheses 2, 3, and 4 justify the fact that we can expect to capture in the sum (9) seismic events originating from the source i alone. Moreover, if the most energetic features are identified at the first iterations of the decomposition, then the most energetic cross-talks from other sources are captured in other sums (10), and thus, do not pollute the residue $R^L d(t)$ any more. Fortunately, this is exactly how OMP proceeds provided that we choose a well-adapted dictionary. Now, if the atoms are expressed, before normalization, as $\mathcal{G}_{\gamma} = h^{(i)} \star w^{(i)}$, and $h^{(i)}$ and $w^{(i)}$ given, respectively, by (7) and (8), then

$$\begin{aligned} \mathcal{G}_{\gamma}(t) &= \sum_{n=1}^N [1 + \alpha(x_n^i - x_0^i)] \sum_{k=1}^K a_k \\ &\quad \times w_{s_k} \left(t - \tau - p(x_n^i - x_0^i) - q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 - T_n^i - \tau_k \right) \end{aligned} \quad (11)$$

with $\gamma = \{i, \tau, p, q, \alpha, K, \{s_k, a_k, \tau_k\}_{1 \leq k \leq K}\}$ the complete set of parameters. Hence, we can construct a decomposition (4) that fulfills the aforementioned conditions required for deblending. Examples of such atoms, given in Fig. 5, show the ability of the algorithm to handle curvature and amplitude variation.

Note that, when L tends to infinity, the residue $R^L d$ is not necessarily white noise or any other type of noise. Indeed, it corresponds to the last residue (part of the signal) not explained by the dictionary, i.e., orthogonal to the dictionary used for decomposition. A “good” decomposition, though, would leave the noise in the residue.

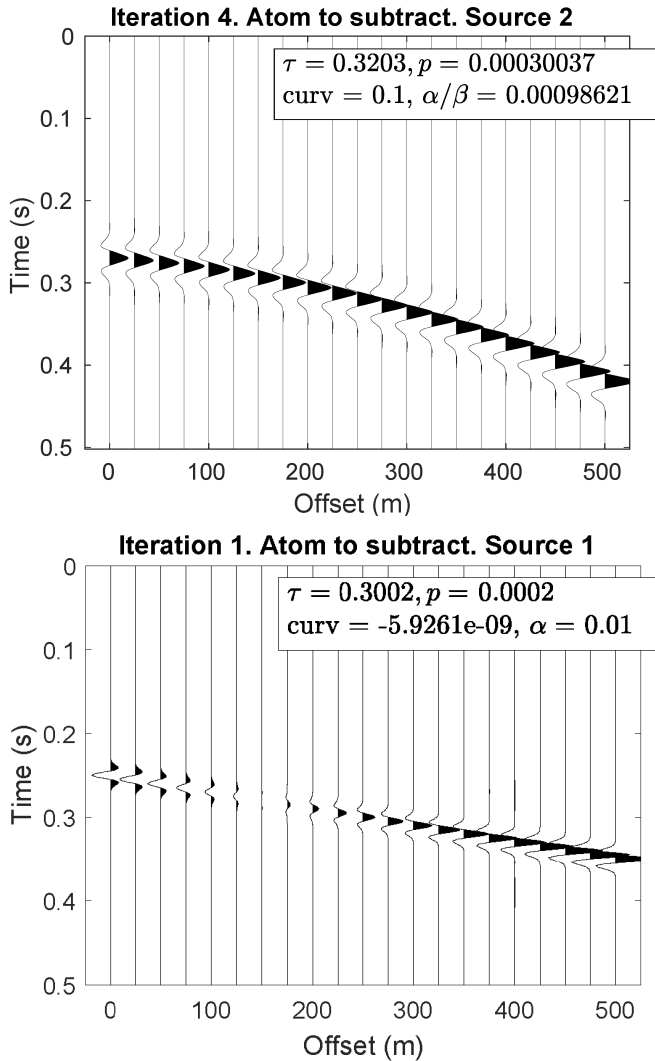


Fig. 5. Examples of atoms $\mathcal{G}_\gamma(t)$ (before normalization) of the seismic events dictionary.

To simplify the computation of vector norms $\|\mathcal{G}_\gamma\|$, we make the following hypothesis.

Hypothesis 7: For each source, the pauses between two consecutive shots are significantly longer than the emission time of each shot of the same source.

Note that Hypothesis 7 does not forbid crosstalk between consecutive shots of the same source, i.e., the delay between consecutive shots can be smaller than the listening time implying auto-pollution or self-simultaneous sourcing.

Now, the problem is to find an approximate solution of (1). For this, we must overcome two major difficulties: 1) the objective function to maximize is not concave and 2) the number of parameters describing an atom is too large for sampling the dictionary into a finite subset of atoms $\Gamma \subset \Omega$. To overcome the first difficulty, we use iterative optimization algorithms that converge to a local maximum whose position depends on the initial conditions. It is therefore crucial to accurately choose the initial conditions. To overcome the second difficulty, we gradually build atoms of the dictionary close to the desired maximum.

B. Initial Conditions of the OMP Optimization Step

In this section, we present our approach to find the initial conditions of the iterative algorithm. We construct an atom \mathcal{G}_γ (before normalization) given by (11) in several steps. We start by building the traveltime curve $h^{(i)}$ given in (7), first looking for parameters i, τ, p, q that maximize the objective function

$$C(i, \tau, p, q) = \left| \sum_{n=1}^N R^{\ell-1} d \left(\tau + p(x_n^i - x_0^i) + q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 + T_n^i \right) \right|. \quad (12)$$

In other words, noting that $C(i, \tau, p, q) = |\tilde{h}^{(i)} \star \widetilde{R^{\ell-1}d(0)}|$, for $\tilde{h}^{(i)}$ introduced in (6), and $R^{\ell-1}d(t) = R^{\ell-1}d(-t)$, we are looking for a traveltime curve $\tilde{h}^{(i)}$ that maximizes the magnitude of its correlation at time $t = 0$ with the residue $R^{\ell-1}d(t)$ at the ℓ th OMP iteration. Here too, the objective function is not concave, and the parameters i, τ, p, q that maximize (12) are found using an iterative optimization algorithm starting from suitable initial conditions and converging to a local maximum. To do so, we introduce the following hypothesis.

Hypothesis 8: To maximize the objective function in (12), good initial conditions are $q = 0$ and the values of i, τ, p that maximize the slant stack magnitude of the residue $R^{\ell-1}d$.

Various successful applications of slant stack (or Linear Radon Transform) to seismic data processing justify this hypothesis, for example, coherent noise suppression, such as multiples and direct arrivals removal [56]; plane-wave decomposition for velocity picking [57]. In our case, if one can pick the absolute maximum in the (τ, p) domain, this maximum identifies a real seismic event with nearly the most energy. Once we have identified a traveltime curve $\tilde{h}^{(i)}$ that maximizes its correlation with the residue, we compute the coefficients α' and β' of the linear regression between the term $R^{\ell-1}d(\tau + p(x_n^i - x_0^i) + \dots)$ appearing in (12) and $x_n^i - x_0^i$ for $1 \leq n \leq N$ that minimize

$$C(\alpha', \beta') = \sum_{n=1}^N \left[R^{\ell-1} d \left(\tau + p(x_n^i - x_0^i) + q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 + T_n^i \right) - [\beta' + \alpha'(x_n^i - x_0^i)] \right]^2 \quad (13)$$

in order to obtain a first estimation of the complete traveltime curve $h^{(i)}$ given in (7). In addition, when $|\beta'| > \varepsilon$ (in our implementation we took $\varepsilon = 10^{-7}$), we set $\alpha = \alpha'/\beta'$.

We observed that criterion (12) does not give the best initial conditions to the OMP optimization when the factor $\beta' + \alpha'(x_n^i - x_0^i)$ changes its sign between the extreme values of x_n^i and we shall see in Section III-E how to modify (12) to obtain better initial conditions.

Next, we define a ‘‘corridor’’ in the representation space (t, x) associated with the i th source. This corridor has a width of $(2M+1)\Delta_\tau$, it is centered around the maximal values of $\tilde{h}^{(i)}$ and passes through the N considered traces. We thus obtain a

nonparametric estimation $\hat{w}^{(i)}$ of the wavelet $w^{(i)}$, associated with the atom \mathcal{G}_γ introduced in (11). The estimation is locally made from the current residue, within the corridor and after making the following hypothesis.

Hypothesis 9: A wavelet estimation can be statistically derived from the N traces by stacking along curves parallel to the traveltime curve maxima weighted by attenuation factors.

We then apply a Tukey window to this nonparametric wavelet estimation to avoid discontinuities at the corridor edges. Finally, we obtain a parametric estimation $w^{(i)}$ of the wavelet having the form (8) by applying the OMP algorithm to the windowed wavelet estimation, and we compute the nonnormalized atom \mathcal{G}_γ , mentioned at the beginning of this section, as $h^{(i)} \star w^{(i)}$.

We can summarize the computation of the initial conditions \mathcal{G}_γ into the following stages.²

- 1) Find the values i , $\hat{\tau}$, and \hat{p} that maximize the slant stack magnitude of the residue $R^{\ell-1}d$.
- 2) From the initial conditions obtained at the previous stage and $q = 0$, find a traveltime curve $\tilde{h}^{(i)}$ maximizing its correlation magnitude with $R^{\ell-1}d$ at time $t = 0$, within the N traces.
- 3) Find the coefficients α' and β' of the regression (13) to obtain $\alpha = \alpha'/\beta'$ and the amplitude-variation-preserving traveltime curve $h^{(i)}(t)$, with the attenuation factor $1 + \alpha(x_n^i - x_0^i)$.
- 4) In the (t, x) trace domain associated with source i , identify a $(2M+1)\Delta_t$ -seconds high corridor centered around the traveltime curve maxima found at the previous stage; then make a nonparametric wavelet estimation using a weighted stacking by reverse attenuation factors along the curves parallel to the $\tilde{h}^{(i)}$ maxima within the corridor.
- 5) Window the nonparametric wavelet estimation obtained at the previous stage and apply OMP to get a parametric estimation $w^{(i)}$ given by (8).
- 6) Find the initial conditions atom which, before normalization, equals to $\mathcal{G}_\gamma = h^{(i)} \star w^{(i)}$.

In this way, we propose to perform deblending by means of OMP, and we use the OMP algorithm twice. To distinguish them, we denote by *outer* OMP, the one which has a dictionary of atoms of the form (11) before normalization, and by *inner* OMP, the one performing the parametric wavelet estimation. In Section III-C, we present the *inner* OMP algorithm.

C. Inner OMP Overview

Before starting the iterations of Matrioshka OMP, we perform a spectral analysis of the data to determine the shapes of the wavelets to use in the inner OMP dictionary. For this, we compute the power spectrum of $d(t)$ and pick the frequency values at its maximum and 3 and 6 dB lower. This procedure provides a set of frequencies that we use to build the wavelet dictionary (for example, the five frequencies above give the dominant frequencies of the Ricker wavelets).

²We shall see in Section III-E that, when the factor $\beta' + \alpha'(x_n^i - x_0^i)$ changes its sign between the extreme values of x_n^i , the stages 2 and 3 can be iterated, modifying the criterion (12). For simplicity, we do not present this procedure here.

1) *Wavelet Dictionary:* We choose a finite number S of classical wavelet shapes. The shape index s ($1 \leq s \leq S$) corresponds to either a Ricker wavelet with a given dominant frequency or an Ormsby wavelet with a given set of cut-off frequencies. If we need Ricker wavelets of different dominant frequencies, we use as many Ricker shapes as needed and likewise for Ormsby wavelets. The predefined shapes can be extended to any other kind of wavelets.

The dictionary is composed of time-shifted unit-norm elementary wavelets of the predefined shapes. Since the estimated wavelet must be inside the abovementioned corridor, we limit the time shifts so that an atom is represented as $w_\gamma(t) = (w_s(t - \nu\Delta'_\tau - \tau')/\|w_s(t)\|)$, where $\nu \in \llbracket -\mu M, \mu M \rrbracket$ is an integer, $\tau' \in (-\Delta'_\tau/2, \Delta'_\tau/2[$ with $\Delta'_\tau = \Delta_t/\mu$, and $\mu^{-1} \in \mathbb{N}$ divides M . Thus, the dictionary is $\mathcal{D} = \{w_\gamma\}_{\gamma \in \Omega}$ with $\Omega = \{(s, \nu, \tau') : s \in \llbracket 1, S \rrbracket, \nu \in \llbracket -\mu M, \mu M \rrbracket \text{ and } \tau' \in (-\Delta'_\tau/2, \Delta'_\tau/2[)$. We also use a discrete version of the dictionary, with vanishing τ' : $\{w_\gamma\}_{\gamma \in \Gamma}$ with $\Gamma = \{(s, \nu, 0) \in \Omega\}$.

2) *Inner OMP:* For simplicity, in this paragraph, we omit the superscript (i) of an estimated wavelet $w^{(i)}$, and we consider wavelets as continuous-time signals.

The inner OMP is initialized with the windowed nonparametric estimation $R^0 w(t) = \tilde{w}(t)$. Let $R^{k-1}w$ be the residue after $(k-1)$ iterations of the inner OMP. At iteration k , first we look for a solution $\hat{\gamma}_k = (s_k, \nu_k, 0)$ to $|\langle R^{k-1}w, w_{\hat{\gamma}_k} \rangle| = \max_{\gamma \in \Gamma} |\langle R^{k-1}w, w_\gamma \rangle|$, which gives initial conditions for the iterative optimization algorithm converging to a local maximum, approximate solution to $|\langle R^{k-1}w, w_{\gamma_k} \rangle| = \max_{\gamma \in \Omega} |\langle R^{k-1}w, w_\gamma \rangle|$. Thus, we obtain w_{γ_k} , the atom of the inner OMP chosen at the iteration k . In the following step, we update the coefficients of the orthogonal projection of \tilde{w} on the vector subspace of the first k atoms obtained via the inner OMP. After K iterations, we obtain the decomposition $\tilde{w}(t) = \sum_{k=1}^K a_k w_{s_k}(t - \nu_k \Delta'_\tau - \tau'_k) + R^K \tilde{w}(t)$, which gives the parametric estimation of the stage 5 above: $w^{(i)} = \sum_{k=1}^K a_k w_{s_k}(t - \nu_k \Delta'_\tau - \tau'_k)$.

Section III-D presents a complete view of the deblending algorithm Matrioshka OMP.

D. Matrioshka OMP Overview

Matrioshka OMP [58] stands for two OMP algorithms embedded into one another. The algorithm is illustrated in Fig. 6, where the *outer* OMP consists of the whole algorithmic loop with the *inner* OMP embedded into it and highlighted in orange. We now describe each step individually.

After the spectral analysis of the data, the second stage of the processing is to split the continuously recorded signal $d(t)$ into temporal frames suitable for deblending. It is worth noting that the definition of windows width using a number of traces is no longer compatible with the data. Indeed, the number of traces (whole or parts) does not necessarily match for the different sources. To overcome this ambiguity, we chose to define window width in terms of time. When a window break occurs between shooting times of a source, we use the knowledge of the previous shooting time to exploit all the

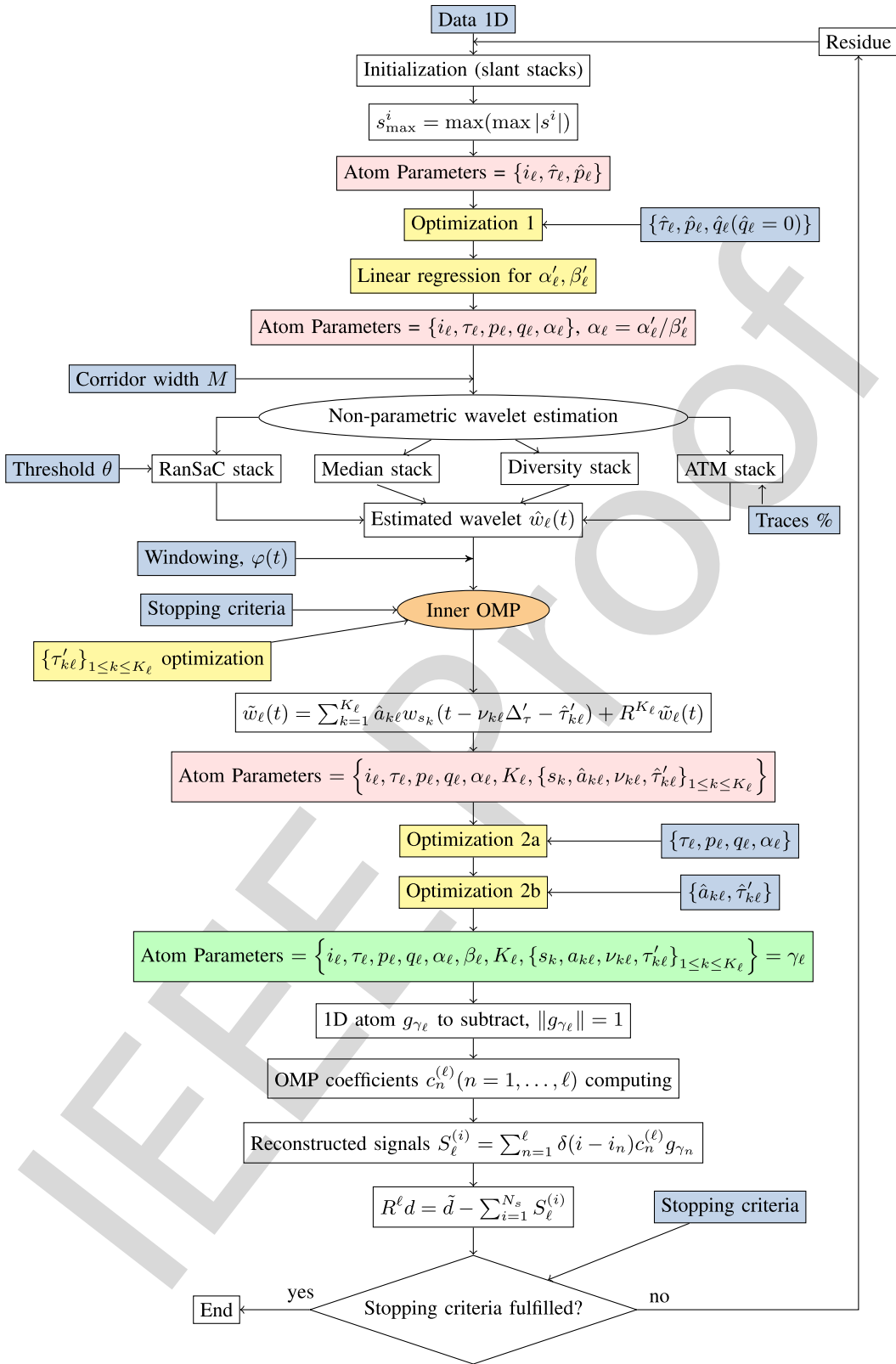


Fig. 6. Matrioshka OMP algorithm, with **inputs**, **optimization steps**, **intermediate parameters**, **final parameters**.

696 information available in the data. Thus, the outer OMP is
 697 initialized from the input data $d(t)$ windowed by a rectangular
 698 time window strictly included in the interval $[0, T_{\text{glob}}]$ and
 699 corresponding to N seismic traces for one source. We denote

by $\tilde{d}(t)$ the windowed signal $d(t)$ and take it as the first
 residue: $R^0 d = \tilde{d}$.

Let $R^{\ell-1} d$ be the residue after $(\ell - 1)$ iterations of
 the outer OMP. At the ℓ th iteration, we have seen in

700
 701
 702
 703

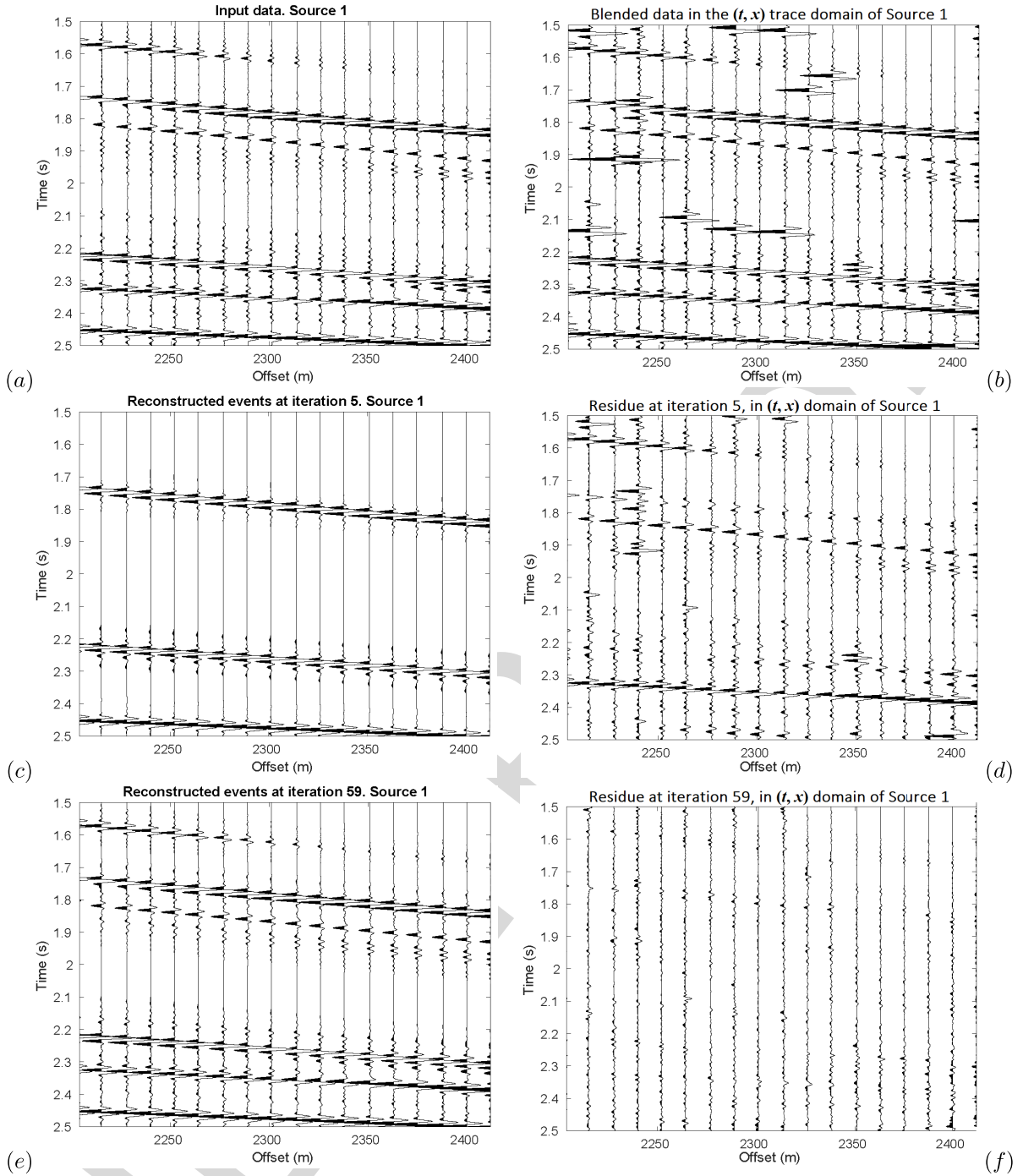


Fig. 7. (a) Input unblended data for the first source and (b) same data after blending. (c) Reconstruction results after $L = 5$ iterations of the outer OMP: events attributed to the first source, i.e., $S_\ell^{(1)}$ of (14) with $\ell = L$ and (d) residue $R^L d$ of (4). (e) and (f) Idem after $L = 59$ iterations. All the signals are represented in the (t, x) trace domain of Source 1. In our method, the debled signal associated with Source 1 after L iterations is the sum of signals appearing in the graphs (c) and (d) for $L = 5$ and the sum of signals in the graphs (e) and (f) for $L = 59$.

704 Sections III-B and III-C how to obtain the initial condi- 712
 705 tions (11) before normalization, which allow an iterative 713
 706 optimization algorithm to converge to a local maximum. 714
 707 Relationships allowing a fast computation of the norm of 715
 708 seismic events of the form (11) under Hypothesis 7 can be 716
 709 found in [59, Appendix C].

710 In order to separate travel-path-related parameters from the 717
 711 wavelet-defining ones, so that they do not intercompensate 718
 719

each other, we first optimize the τ, p, q and α parameters 712
 and then the $(a_k, \tau'_k)_{1 \leq k \leq K}$ parameters. Note that the whole 713
 dictionary is never created or stored due to computational 714
 costs: a new element of the dictionary is estimated at each iter- 715
 ation. We obtain after these optimization stages the atom g_{γ_ℓ} , 716
 approximate solution of (1), with $\gamma_\ell = (i_\ell, \tau_\ell, p_\ell, q_\ell, \alpha_\ell, K_\ell,$ 717
 $(s_{p,\ell}, a_{p,\ell}, v_{p,\ell}, \tau'_{p,\ell})_{1 \leq p \leq K_\ell})$. We then update the coefficients 718
 $(c_p^{(\ell)})_{1 \leq p \leq \ell}$ of the orthogonal projection of \tilde{d} on the linear 719

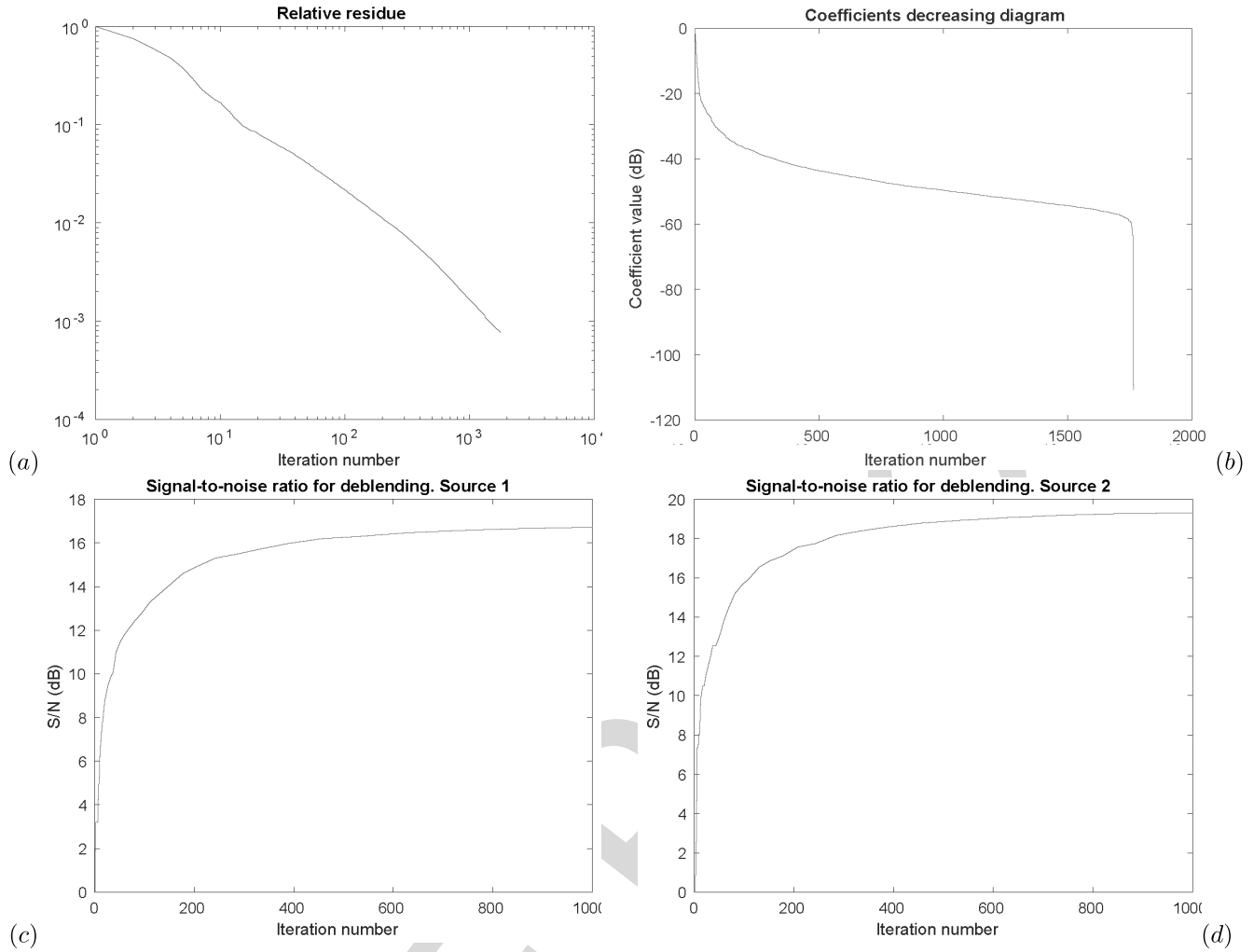


Fig. 8. (a) Results obtained on synthetic data with real seismic noise added. Residue energy decreasing in bilogarithmic scale. (b) Coefficients magnitude decreasing diagram. (c) and (d) Signal-to-noise ratio increasing with the outer OMP iterations for the two sources. The signal-to-noise ratio is computed as $S/N = 10 \log_{10}(\|d_s\|^2 / \|d_s - d_d\|^2)$, where d_s denotes the initial single source data, and d_d the debleded data for the same source.

720 subspace spanned by the first ℓ outer OMP atoms, and the
 721 weighted sums—called explained signals in the following—
 722 assuming we have N_s sources:

$$723 \quad S_\ell^{(i)}(t) = \sum_{p=1}^{\ell} \delta_{i,p} c_p^{(\ell)} g_{\gamma_p} \quad (\text{for source } i = 1, \dots, N_s) \quad (14)$$

724 where $\delta_{i,j}$ is the Kronecker delta function. After L iterations,
 725 $\tilde{d}(t) = \sum_{\ell=1}^L c_\ell g_{\gamma_\ell}(t) + R^L d(t) = \sum_{i=1}^{N_s} S_L^{(i)}(t) + R^L d(t)$, and
 726 the debleded signal associated with the i th source is equal
 727 to $S_L^{(i)}(t) + R^L d(t)$. To reduce the computational complexity
 728 of the method, the optimization stages must be efficiently
 729 implemented. An asymptotic complexity analysis of the algo-
 730 rithm is given in [59]. After processing each temporal window,
 731 the debleded data are merged. To increase the debleding
 732 quality and avoid high-frequency residual noise, windows
 733 overlap. We end the section by presenting the initial condition
 734 computation when the maximum magnitude of the wavelet
 735 changes sign from one end of the seismic event to the other.

E. Seismic Events With a Phase Rotation

To find the initial conditions of \mathcal{G}_γ , the approach described
 in stages 2 and 3 of Section III-B works perfectly for
 seismic events which have the same polarity all along the
 processing window. However, it is common to encounter
 a “phase rotation” corresponding to events whose maxima
 have different signs on the left and on the right edge of
 the processing window (see Fig. 5). In this case, (12) no
 longer represents a good objective function to maximize
 because the algorithm tends to favour (to follow) amplitudes
 of the same sign. To solve this problem, we modified the
 criterion (12) to

$$748 \quad C(i, \tau, p, q) = \left| \sum_{n=1}^N R^{\ell-1} d \left(\tau + p(x_n^i - x_0^i) + q \left(\frac{x_n^i - x_0^i}{x_{\max}^i - x_{\min}^i} \right)^2 + T_n^i \right) \right. \\ \left. \times \operatorname{sgn} [\beta' + a'(x_n^i - x_0^i)] \right| \quad (15) \quad 750$$

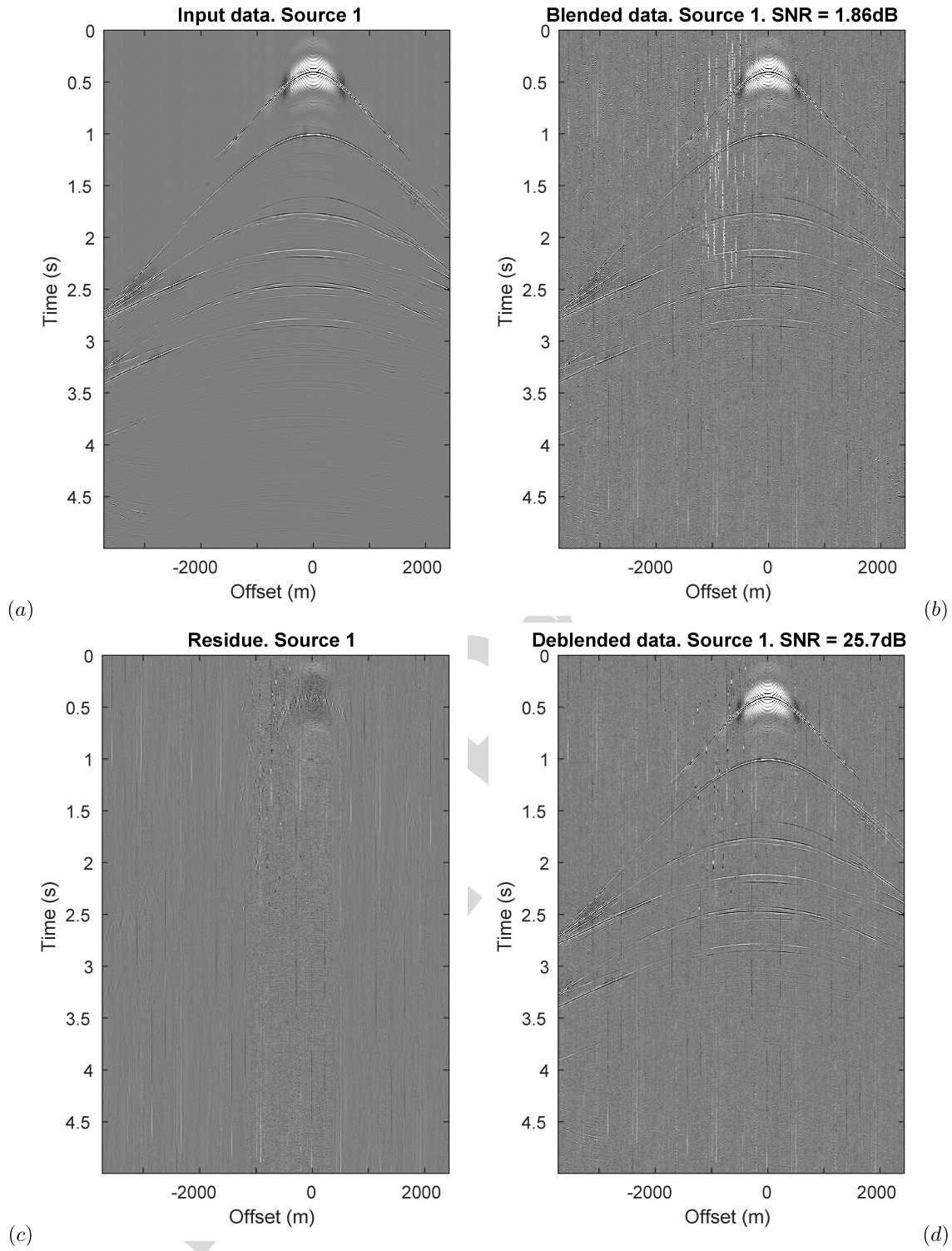


Fig. 9. (a) Full gather tests for the synthetic Marmousi data with added real seismic noise, the first source. Input nonblended signal and (b) same data after blending. (c) Residue after decomposition and (d) data after debleding (explained signal with the residue added).

751 and iterated twice the stages 2 and 3; this proved to be effective
 752 in our simulations.

753 F. Stopping Criteria

754 Due to the significant complexity of seismic data with
 755 respect to our dictionary, it is very difficult to define a single

stopping criterion applicable everywhere. Moreover, the stop-
 ping criterion must be adapted to the downstream processing.
 For this reason, we propose setting multiple stopping criteria
 for each simulation to achieve more accurate results and,
 at the same time, avoid wasting machine time on unnecessary
 precision seeking.

756
 757
 758
 759
 760
 761

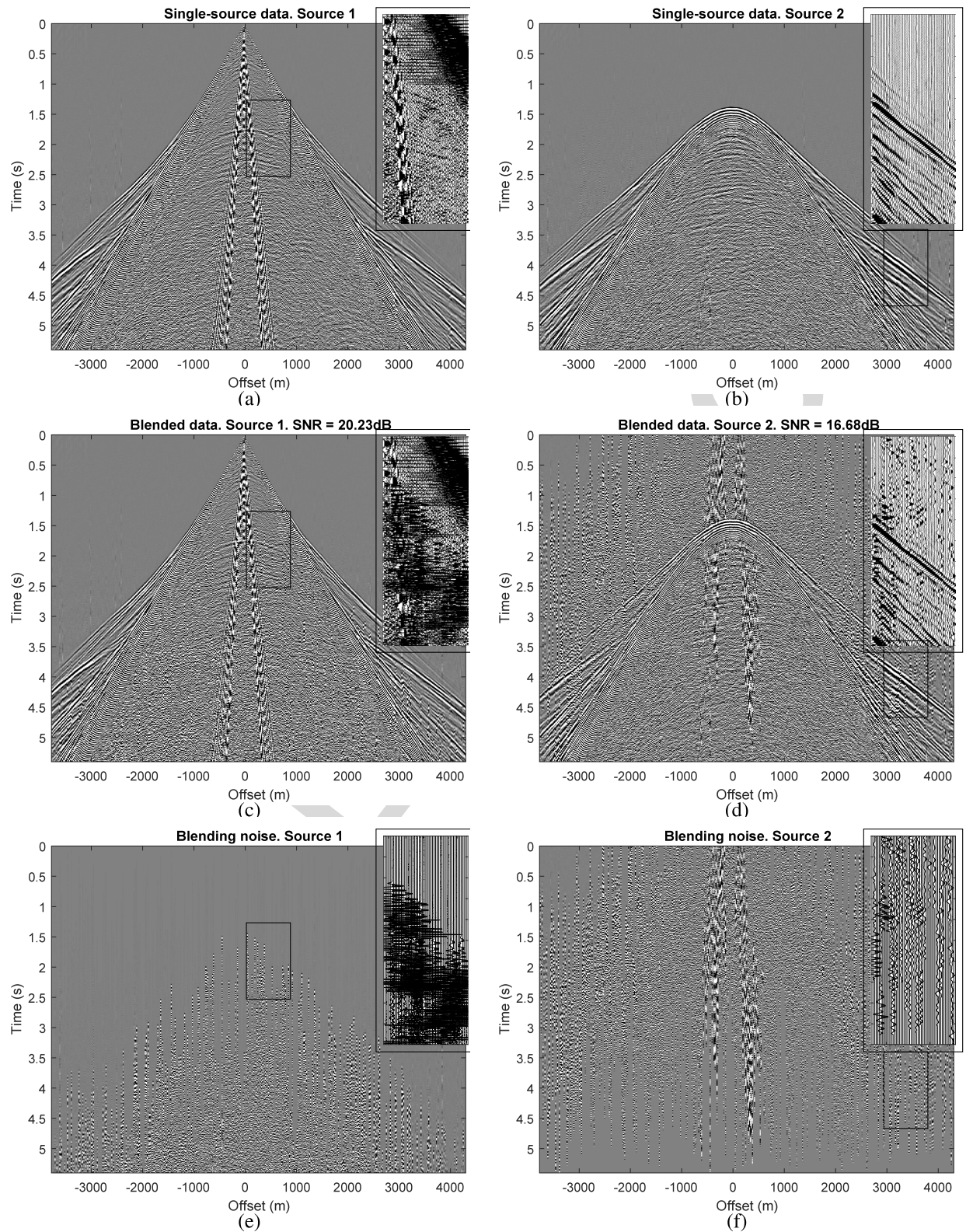


Fig. 10. (a) and (b) Real seismic data example: Torpille data. Input clean signal. (c) and (d) Artificially blended signal. (e) and (f) Isolated blending noise. For each image, its zoomed-in part highlighted by a rectangle is given at its top-right corner.

762
763
764

1) The OMP stopping criterion proposed in [45] is the achievement of a null, or at least of a sufficiently small ℓ_2 -norm of the residue $R^L d : \|R^L d\| < N_R$.

This approach is intuitive, but not easy to implement, as different seismic data sets do not have the same amplification, nor do they have the same level of ambient

765
766
767

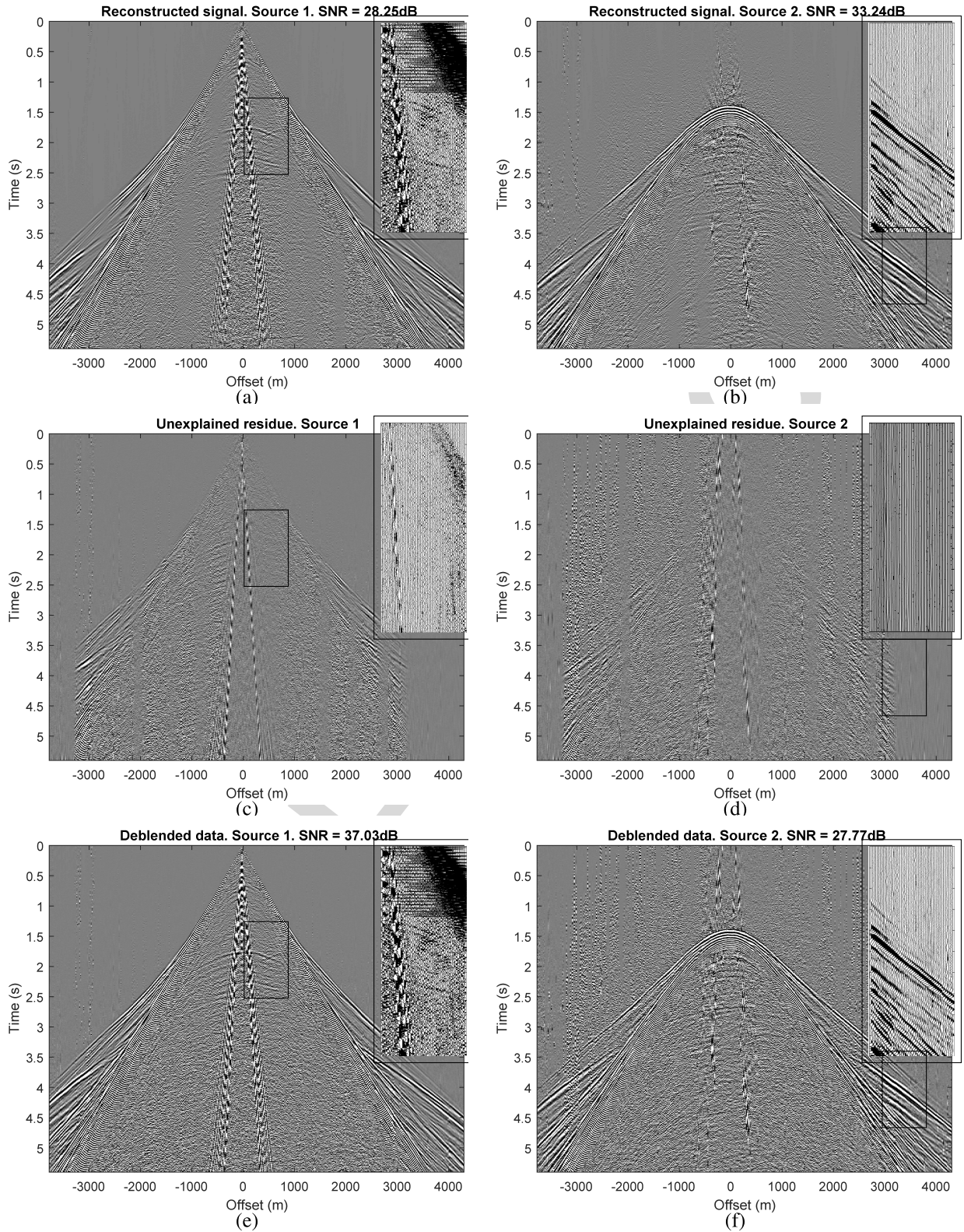


Fig. 11. (a) and (b) Explained signal. (c) and (d) Unexplained residue for the two sources after approximately 1000 iterations of OMP per lateral window. (e) and (f) Deblended data (explained signal of the source with the residue added) for the two sources. For each image, its zoomed-in part highlighted by a rectangle is given at its top-right corner.

768 noise or other noises which we would not want to
 769 reconstruct. In other words, the meta-parameter N_R is
 770 difficult to choose as it is highly data dependent.

2) One relative value related to the residue energy is the
 relative residual energy $\|R^L d\|_2^2 / \|\tilde{d}\|_2^2 < E_R$. The meta-
 parameter E_R can easily be set to some very small value

771

772

773

(of the order of computation error) in the absence of noise or can be derived from the pre-estimated signal-to-noise ratio in the case of noisy data.

- 3) In some cases the parameters N_R and even E_R are difficult to define. If, in addition, the user only requires a low reconstruction precision (only wants to reconstruct and separate the most energetic events), it could be helpful to set L_{\max} —the maximal number of iterations to perform—to a low value.
- 4) The reciprocal condition number is used to measure whether a matrix is well or badly conditioned (if this number is small). The condition number of a matrix affects the solutions of similar linear systems of equations: if the values of the matrix are slightly perturbed, this leads to big differences in the solution; thus, we stop the decomposition when this number is too small.

IV. RESULTS

This section shows results obtained with the Matrioshka OMP algorithm, applied to a complex synthetic data set issued from the Marmousi geological model [60], with real seismic noise (see Section IV-A) and to real ocean bottom node (OBN) seismic data acquired in Torpille (Offshore Gabon) (see Section IV-B). More results on simple synthetic data examples that demonstrate the performance in “laboratory” conditions of the method can be found in [59].

A. Complex Synthetics—Realistic Case Study

We tested our method on a realistic data set, generated by the Allied Geophysical Laboratories of the University of Houston from the Marmousi geological model. Martin *et al.* [61] performed a highly precise elastic modeling to provide as many of the seismic features usually present in real seismic data as possible. Namely, the data contain not only primary and multiple reflections, but also diffractions, head waves, surface waves, scattering effects, and other realistic particularities. The acquisition geometry adopted for this simulation is that of a source vessel towing an airgun source at a depth of 10 m and performing a shot every 25 m. The source signature is a zero-phase 5–10–60–80 Hz Ormsby wavelet with frequencies up to 80 Hz. The Ocean Bottom Cable is situated at the water bottom at a depth of 450 m. We performed an artificial blending of the data by attributing different parts of the data to two different sources and creating overlapping shooting time patterns for each source: the first source shoots regularly, with an interval equal to 5 s, and the second source shoots with irregular time intervals around 7 ± 2 s.

The first test, illustrated in Figs. 7 and 8, contains 20 traces for each source, which corresponds to a 500-m-wide lateral window. Fig. 7 shows decomposition residue and explained signal for the first source in the upper part of the section, where the signal is quite strong since it contains direct arrival and surface waves. Fig. 7(c) and (d) shows the decomposition result after only five iterations of the outer OMP: several of the most energetic seismic events have already been reconstructed, and the residue energy has significantly decreased. After 59 iterations of the outer OMP [see Fig. 7(e) and (f)], the useful

signal present in the section is almost perfectly explained. The leakage of Source 1 remaining in the residue [see Fig. 7(f)] is present in the deblended signal associated with Source 2, but with a sufficiently low energy to be eliminated as acquisition noise by the classical downstream processing. However, because of the presence of significantly weaker signals in other parts of the studied sections, we continued the decomposition up to 1750 iterations of the outer OMP, getting a perfectly explained useful signal. The energy of the residue decreases almost linearly in logarithmic scale, as shown in Fig. 8(a). Fig. 8(b) shows the magnitude of the coefficients found during the decomposition. Note the rapid decrease in the beginning of the curve, indicating the sparsity of the transform. Fig. 8(c) and (d) shows for the two sources the increasing signal-to-noise ratio computed as $S/N = 10 \log_{10}(\|d_s\|^2 / \|d_s - d_d\|^2)$, where d_s denotes the initial single source data, and d_d the deblended data for the same source.

Fig. 9 shows a test on the same data, with the entire shot lines processed using sliding windows and in the presence of real seismic noise. Note that most of the noise is left in the residue, moreover to avoid any signal loss, the residue can be added back to the explained coherent events, if there is any signal left in it.

B. Real Seismic Data Example

In this section, we present test results on real data extracted from a 3-D OBN seismic survey acquired in Torpille. The acquisition used a conventional single source mode, with an airgun seismic source towed at a 7-m depth with a shot-point interval of 50 m. The water depth in this area varies from 25 to 35 m, which implies the presence of Scholte waves making the data almost as difficult to process as onshore. The sampling period was of 3 ms, and the listening time for each shot was of 5.4 s. We blended them artificially as for the Marmousi data. The clean and blended input data are shown in Fig. 10(a)–(d). Note that the shooting line of the first source is significantly closer to the receiver, since the useful signal in Fig. 10(b) and (d) is located deeper (i.e., later in time) than that in Fig. 10(a) and (c). Obviously, the further away the source is from the receiver, the weaker its recorded signal is. Therefore, the blending appears more aggressive for the second source than for the first one, as shown in Fig. 10(e) and (f). The first source, however, is also significantly contaminated, especially in the part where useful signals, as the primary reflections, are present (below 2 s). The decomposition allows us to reconstruct the most energetic physical events, such as the direct arrivals, the surface waves and the guided waves. A significant part of the reflections is also reconstructed, which is well seen in the zoomed-in parts of Fig. 11(a) and (b). However, part of the coherent signals stays in the residue [see Fig. 11(c) and (d)]. Nevertheless, in order to avoid leakage, the residue can be added back to the reconstructed events for each source, as shown in Fig. 11(e) and (f). Note that the decomposition and deblending results for the real seismic data have inferior quality compared to the synthetic data with real noise added. This can be explained as follows. First, the Torpille data contain a significant part of incoherent noise—which

our algorithm is not trying to capture—and when we blend data, we sum up the ambient noise recorded at different times, leading to a less favorable situation than a true simultaneous sources acquisition. Second, the big difference in the energy of the two sources is also difficult to handle, as sometimes the energetic noise tends to be reconstructed as coherent signal. Nevertheless, we were able to achieve a significant improvement of the signal-to-noise ratio for the deblending results shown in Fig. 11(e) and (f): around 10–15 dB for both sources. Taking into account that the deblending takes place in the very beginning of the processing sequence, the residual blending noise is likely to be handled by further conventional denoising or other processing. Limitations of our method include potential high computational complexity when big data sets need to be processed with a high level of precision. We addressed this issue by implementing analytical derivation in the optimization routines and fast norm calculation, but further code optimizations may be needed to industrialize the algorithm. Our method provides both deblended signals and a sparse representation of seismic data with a given precision, which is beneficial for diverse seismic data processing problems.

V. CONCLUSION

In this article, we have proposed a new source-separation method applied to seismic data acquired in simultaneous-source mode. This method consists of two nested OMPs and is called Matrioshka OMP. We have proposed two mathematical models of sensor signals in simultaneous-source seismic surveys. These models are justified by nonrestrictive assumptions on the seismic survey and the simultaneous sources, which we have stated as hypotheses. Our data-driven seismic event model is based on features which are characterized by spatial coherence of wavelet signals. Precisely, a seismic event is a straight or slightly curved feature in the trace representation of the data with a specific wavelet sufficiently stable within a local spatial window, whose magnitude can linearly vary according to the offset. We have deduced from this model specific dictionaries adapted to raw seismic data without preprocessing, and we have implemented two nested OMPs with these dictionaries. For this, we have efficiently solved a nonconvex optimization problem thanks to the gradual construction of the initial conditions close to the globally optimal solution. Finally, we have tested our method on complex synthetic seismic data with real noise and on real data. The synthetic data examples presented show excellent deblending results: the algorithm is capable of explaining almost all of the coherent seismic events present in the data. The real data example was more difficult to process, but the final results are acceptable in terms of further processing.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments that helped improve the presentation of this article.

REFERENCES

- [1] W. Lynn, M. Doyle, K. Lerner, and R. Marschall, "Experimental investigation of interference from other seismic crews," *Geophysics*, vol. 52, pp. 1501–1524, 1987.
- [2] P. I. Pecholcs *et al.*, "Over 40,000 vibrator points per day with real-time quality control: Opportunities and challenges," in *Proc. SEG Tech. Program Expanded Abstr.*, 2010, pp. 111–115.
- [3] J. Kommedal, G. Alexander, L. Wyman, and S. Wagner, "ISS on ice: Seismic acquisition in the arctic," in *Proc. SEG Tech. Program Expanded Abstr.*, 2016, pp. 6–10.
- [4] M. G. Barbier and J. R. Viallix, "SOSIE: A new tool for marine seismology," *Geophysics*, vol. 38, no. 4, pp. 673–683, 1973.
- [5] D. Silverman, "Method of three dimensional seismic prospecting," U.S. Patent 4 159 463 A, Jun. 26, 1979.
- [6] H. J. Rozemond, "Slip-sweep acquisition," in *Proc. SEG Tech. Program Expanded Abstr.*, 1996, pp. 64–67.
- [7] H. Liu and R. Abma, "Simultaneous sources and deblending using multiple sweeps," in *Proc. SEG Tech. Program Expanded Abstr.*, 2017, pp. 141–145.
- [8] N. Moldoveanu, P. Jones, S. Totten, and E. Rosso, "Vibro-seis simultaneous shooting using encoded sweeps: A field experiment," in *Proc. SEG Tech. Program Expanded Abstr.*, 2017, pp. 146–150.
- [9] A. Zhukov, I. Korotkov, E. Sidenko, I. Nekrasov, P. Gridin, and T. Galikeev, "Simultaneous pseudo-random shuffle-sweep generation and increased seismic data acquisition productivity," in *Proc. SEG Tech. Program Expanded Abstr.*, 2017, pp. 151–155.
- [10] C. J. Beasley, R. E. Chambers, and Z. Jiang, "A new look at simultaneous sources," in *Proc. SEG Tech. Program Expanded Abstr.*, 1998, pp. 133–135.
- [11] S. T. Vaage, "Method and system for acquiring marine seismic data using multiple seismic sources," U.S. Patent 6906981 B2, Jun. 14, 2005.
- [12] D. Howe, "Independent simultaneous sweeping—A method to increase productivity of land seismic crews," in *Proc. SEG Tech. Program Expanded Abstr.*, 2008, pp. 2826–2830.
- [13] G. Hampson, J. Stefani, and F. Herkenhoff, "Acquisition using simultaneous sources," *Lead. Edge*, vol. 27, no. 7, pp. 918–923, 2008.
- [14] A. J. Berkhou, G. Blaquiè, and D. J. Verschuur, "From simultaneous shooting to blended acquisition," in *Proc. SEG Tech. Program Expanded Abstr.*, 2008, pp. 2831–2838.
- [15] R. Abma, Q. Zhang, A. Arogunmati, and G. Beaudoin, "An overview of BP's marine independent simultaneous source field trials," in *Proc. SEG Tech. Program Expanded Abstr.*, 2012, pp. 1–5.
- [16] W. Dai, X. Wang, and G. T. Schuster, "Least-squares migration of multisource data with a deblurring filter," *Geophysics*, vol. 76, no. 5, pp. R135–R146, 2011.
- [17] D. J. Verschuur and A. J. Berkhou, "Seismic migration of blended shot records with surface-related multiple scattering," *Geophysics*, vol. 76, no. 1, pp. A7–A13, 2011.
- [18] G. Henin *et al.*, "Deblending 4-component simultaneous-source data—A 2D OBC case study in Malaysia," in *Proc. SEG Tech. Program Expanded Abstr.*, 2015, pp. 43–47.
- [19] P. Paramo, K. Vincent, A. Cegna, J. Kommedal, P. Napier, and S. Cardinez, "AVO analysis of independent simultaneous source OBC data from trinidad," in *Proc. SEG Tech. Program Expanded Abstr.*, 2013, pp. 368–372.
- [20] E. Shipilova *et al.*, "Simultaneous-source seismic acquisitions: Do they allow reservoir characterization? A feasibility study with blended onshore real data," in *Proc. SEG Tech. Program Expanded Abstr.*, 2016, pp. 107–112.
- [21] T. Krupovnickas, K. Matson, C. Corcoran, and R. Pascual, "Marine simultaneous source OBS survey suitability for 4D analysis," in *Proc. SEG Tech. Program Expanded Abstr.*, 2012, pp. 1–5.
- [22] D. M. Davies and M. Ibram, "Evaluating the impact of ISS HD-OBC acquisition on 4D data," in *Proc. 77th EAGE Conf. Exhib.*, 2015, pp. 1–5.
- [23] R. R. Haacke, G. Hampson, and B. Golebiowski, "Simultaneous shooting for sparse OBN 4D surveys and deblending using modified Radon operators," in *Proc. 77th EAGE Conf. Exhib.*, 2015.
- [24] K. Eggenberger *et al.*, "Signal apparition-enabled parallel-source acquisition of 4D-grade seismic data: Results from a field test in the North Sea," in *Proc. SEG Tech. Program Expanded Abstr.*, 2017, pp. 68–73.

- [25] I. Moore, "Removing seismic interference using simultaneous or near simultaneous source separation," U.S. Patent 0097885 A1, Apr. 22, 2010.
- [26] J. Robertsson, L. Amundsen, and A. Pedersen, "Signal apparition for simultaneous source wavefield separation," *Geophys. J. Int.*, vol. 206, no. 2, pp. 1301–1305, 2016.
- [27] Z. Tang and X. Campman, "A coherent simultaneous shooting scheme and its source separation," in *Proc. 78th EAGE Conf. Exhib.*, 2016.
- [28] S. Huo, Y. Luo, and P. G. Kelamis, "Simultaneous sources separation via multidirectional vector-median filtering," *Geophysics*, vol. 77, no. 4, pp. V123–V131, 2012.
- [29] C. Peng, B. Liu, A. Khalil, and G. Poole, "Deblending of simulated simultaneous sources using an iterative approach: An experiment with variable-depth streamer data," in *Proc. SEG Tech. Program Expanded Abstr.*, 2013, pp. 4278–4282.
- [30] S. Gan, Y. Wang, S. Chen, and X. Chen, "Deblending using a structural-oriented median filter," in *Proc. SEG Tech. Program Expanded Abstr.*, 2015.
- [31] S. Spitz, G. Hampson, and A. Pica, "Simultaneous source separation: A prediction-subtraction approach," in *Proc. SEG Tech. Program Expanded Abstr.*, 2008, pp. 2811–2815.
- [32] Z. Zhang, Q. Liu, Y. Xuan, H. Sun, Y. Hu, and L. Han, "The direct arrival in blended data," in *Proc. SEG Tech. Program Expanded Abstr.*, 2016, pp. 275–279.
- [33] C. Bagaini, M. Daly, and I. Moore, "The acquisition and processing of dithered slip-sweep vibroseis data," *Geophys. Prospecting*, vol. 60, no. 4, pp. 618–639, 2012.
- [34] R. Abma and J. Yan, "Separating simultaneous sources by inversion," in *Proc. 71st EAGE Conf. Exhib.*, 2009.
- [35] R. Abma, T. Manning, M. Tanis, J. Yu, and M. Foster, "High-quality separation of simultaneous sources by sparse inversion," in *Proc. 72nd EAGE Conf. Exhib.*, 2010.
- [36] K. Wapenaar, J. van der Neut, and J. Thorbecke, "On the relation between seismic interferometry and the simultaneous-source method," *Geophys. Prospecting*, vol. 60, no. 4, pp. 802–823, 2012.
- [37] R. Abma *et al.*, "Independent simultaneous source acquisition and processing," *Geophysics*, vol. 80, no. 6, pp. WD37–WD44, 2015.
- [38] I. Moore, "Simultaneous sources—Processing and applications," in *Proc. 72nd EAGE Conf. Exhib.*, 2010.
- [39] H. Mansour, H. Wason, T. Lin, and F. J. Herrmann, "Randomized marine acquisition with compressive sampling matrices," *Geophys. Prospecting*, vol. 60, no. 4, pp. 648–662, 2012.
- [40] Y. Chen, "Deblending by iterative orthogonalization and seislet thresholding," in *Proc. SEG Tech. Program Expanded Abstr.*, 2015, pp. 53–58.
- [41] P. Doulgeris, A. Mahdad, and G. Blacquièrè, "Iterative separation of blended marine data: Discussion on the coherency-pass filter," in *Proc. SEG Tech. Program Expanded Abstr.*, 2011, pp. 26–31.
- [42] A. Mahdad, P. Doulgeris, and G. Blacquièrè, "Separation of blended data by iterative estimation and subtraction of blending interference noise," *Geophysics*, vol. 76, no. 3, pp. Q9–Q17, 2011.
- [43] P. Doulgeris, "Inversion methods for the separation of blended data," Ph.D. dissertation, Dept. Geophys., Delft Univ. Technol., Delft, The Netherlands, 2013.
- [44] A. J. Berkhout, "Changing the mindset in seismic data acquisition," *Lead. Edge*, vol. 27, pp. 924–938, Jul. 2008.
- [45] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [46] Y. C. Pati, R. Rezaïifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. Signals, Syst. Comput.*, Nov. 1993, pp. 40–44.
- [47] T. Nguyen and J. Castagna, "Matching pursuit of two dimensional seismic data and its filtering application," in *Proc. SEG Tech. Program Expanded Abstr.*, 2000.
- [48] P. Hugonnet and J.-L. Boelle, "Beyond aliasing regularisation by plane event extraction," in *Proc. 69th EAGE Conf. Exhib.*, 2007.
- [49] P. Hugonnet, J.-L. Boelle, and F. Prat, "Local linear events extraction and filtering in the presence of time-shifts," in *Proc. 74th EAGE Conf. Exhib.*, 2012.
- [50] J. Wang, M. Ng, and M. Perz, "Seismic data interpolation by greedy local Radon transform," *Geophysics*, vol. 75, no. 6, pp. WB225–WB234, 2010.
- [51] A. Adamo, P. Mazzucchelli, and N. Bienati, "Weak orthogonal matching pursuit with geophysical atom selection," in *Proc. 76th EAGE Conf. Exhib.*, 2014.
- [52] F. Boßmann and J. Ma, "Asymmetric chirplet transform for sparse representation of seismic data," *Geophysics*, vol. 80, no. 6, pp. WD89–WD100, 2015.
- [53] H. Hu, Y. Liu, A. Osen, and Y. Zheng, "Compression of local slant stacks by the estimation of multiple local slopes and the matching pursuit decomposition," *Geophysics*, vol. 80, no. 6, pp. WD175–WD187, 2015.
- [54] R. Zhang and J. Castagna, "Seismic sparse-layer reflectivity inversion using basis pursuit decomposition," *Geophysics*, vol. 76, no. 6, pp. R147–R158, 2011.
- [55] R. E. Sheriff and L. P. Geldart, *Exploration Seismology*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1995.
- [56] D. Hampson, "Inverse velocity stacking for multiple elimination," *J. Can. Soc. Explor. Geophys.*, vol. 22, no. 1, pp. 44–55, 1986.
- [57] P. L. Stoffa, J. B. Diebold, and P. Buhl, "Inversion of seismic data in the τ -p plane," *Geophys. Res. Lett.*, vol. 8, no. 8, pp. 869–872, 1981.
- [58] E. Shipilova, J.-L. Boelle, M. Bloch, M. Barret, and J.-L. Collette, "Matrioshka orthogonal matching pursuit for blended seismic source separation," in *Proc. SEG Tech. Program Expanded Abstr.*, 2017, pp. 4919–4924.
- [59] E. Shipilova, "Separation of signals originating from simultaneous seismic sources by greedy signal decomposition methods," Ph.D. dissertation, Dept. Signal Process., CentraleSupélec, France, 2018.
- [60] R. Versteeg, "The Marmousi experience: Velocity model determination on a synthetic complex data set," *Lead. Edge*, vol. 13, no. 9, pp. 927–936, Sep. 1994.
- [61] G. S. Martin, R. Wiley, and K. J. Marfurt, "Marmousi2: An elastic upgrade for Marmousi," *Lead. Edge*, vol. 25, no. 2, pp. 156–166, Jan. 2006.



Ekaterina Shipilova received the bachelor's and master's degrees in geology from Lomonosov Moscow State University, Moscow, Russia, in 2011 and 2013, respectively, and the Ph.D. degree in signal processing from CentraleSupélec, Metz, France, in 2018.

In 2013, she joined TOTAL, Pau, France, to work as a Research Geophysicist in seismic acquisition and processing. She moved to Paris in 2018 to join the Total's Petroleum Basins Evaluation Department, where she is currently working as a Synthesis Geoscientist. Her research interests include seismic acquisition and processing, seismic attributes, and seismic characterization of source rocks.



Michel Barret received the Engineering degree from Ecole Supélec, Paris, France, the Ph.D. and Habilitation degrees in signal processing from the University of Paris-Sud, Orsay, France, in 1993 and 2010, respectively.

In 1986, he joined the Automatics and Signal Processing Department, Ecole Supélec, where he became an Associate Professor in 1997 and a Professor in 2011. He is currently a Professor with CentraleSupélec, Metz, France, where he teaches statistical signal processing. He is also with the UMI 2958 Georgia Tech-CNRS, Metz. His research interests include stability of multidimensional digital recursive filters, adapted filter banks, multicomponent image compression, statistical processes, and signal representations.

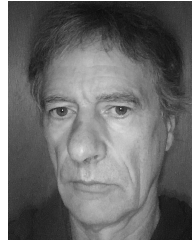
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172



Matthieu Bloch received the Engineering degree from Supélec, Gif-sur-Yvette, France, in 2003, the M.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2003, the Ph.D. degree in engineering science from the Université de Franche-Comté, Besançon, France, in 2006, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology (Georgia Tech.) in 2008.

From 2008 to 2009, he was a Post-Doctoral Research Associate with the University of Notre Dame, South Bend, IN, USA. Since July 2009, he has been with the Faculty of the School of Electrical and Computer Engineering. From 2009 to 2013, he was based at Georgia Tech Lorraine, Metz, France. He is currently an Associate Professor with the School of Electrical and Computer Engineering, Georgia Tech. He is the coauthor of the textbook *Physical-Layer Security: From Information Theory to Security Engineering* (Cambridge University Press). His research interests are in the areas of information theory, error-control coding, wireless communications, and cryptography.

Dr. Bloch was a co-recipient of the IEEE Communications Society and the IEEE Information Theory Society 2011 Joint Paper Award. He has served on the organizing committee of several international conferences. He was the Chair of the Online Committee of the IEEE Information Theory Society from 2011 to 2014, an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION from 2016 to 2019, and he has been on the Board of Governors of the IEEE Information Theory Society. He has been an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY since 2016.



Jean-Luc Boelle received the Engineering degree from École Centrale des Arts et Manufactures, Paris, France, in 1977, and the Ph.D. degree from the Soils and Structures Laboratory, École Centrale Paris, Paris, in 1983.

He joined the French Research Institute for Ocean Science, Brest, France, in 1983 and TOTAL, Pau, France, Oil and Gas company, in 1988. He worked and headed several research projects dealing with seismic wave propagation modeling, and seismic data acquisition and processing.

1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183



Jean-Luc Collette received the Engineering degree from Ecole Supélec, Paris, France, in 1985.

In 1986, he joined the Automatics and Signal Processing Department, Ecole Supélec. He is currently a Professor with CentraleSupélec, Metz, France, where he teaches automatic control and image processing. His research interests include hybrid filter banks and biomedical imaging.

1184
1185
1186
1187
1188
1189
1190
1191