# UP-DOWN SEPARATION OF OBS WAVEFIELD USING ICA TECHNIQUES

LUC T. IKELLE

*Imode Energy Research 9900 Westpark Drive, Houston, Texas 77063, U.S.A.*
*ikelle@imodeenergy.com*
*Texas A&M University, College Station, TX 77843, U.S.A.*

(Received March 7, 2016; revised version accepted June 22, 2016)

## ABSTRACT

Ikelle, L.T., 2016. Up-down separation of OBS wavefield using ICA techniques. *Journal of Seismic Exploration*, 25: 419-432.

We here describe an effective way of performing up-down separation of OBS wavefield. The key feature of the up-down separation is the decomposition matrices. We will also name them mixing matrices. These matrices describe how upgoing and downgoing wavefields have been combined to produce the multicomponent data that we record. The inverses of these matrices allow us to decompose multicomponent data. We describe statistical ways of estimating these mixing matrices. By taking advantage of the redundancy of seismic data, we here present a way to simultaneously reconstruct the mixing matrices, the upgoing and downgoing wavefields by essentially using the statistical methods. The attractive features of the statistical wavefield methods are that they are independent of the data dimension (i.e., they are applicable to 2D and 3D data without any modification of the computer code and without a sampling or interpolation requirement) and that they work for aliased data as well as for nonaliased data, and even for nonuniform sampled data. The only assumptions here are that (1) the mixing matrix is invertible and (2) the number of datapoints is large enough for statistical applications.

KEY WORDS: up-down separation, ocean bottom seismics, PZ summation, aliased data, nonuniform sampled data, 2D, 3D, independent component analysis, constrained independent component analysis.

## INTRODUCTION

In modern seismology, each sensor is actually a multicomponent recording system. In ocean-bottom acquisitions, we can record both pressure and particle velocity (i.e., four components or two components when only the vertical component of the particle velocity is recorded). These types of data allow us to perform up/down separation. The up/down separation consists of separating the

0963-0651/16/$5.00

vector wavefields into upgoing and downgoing wavefields at the receiver location. The key feature of these decompositions is the decomposition matrices. We will also name them mixing matrices. These matrices describe how upgoing and downgoing wavefields and P- and S-wavefields have been combined during the wave propagation to produce the multicomponent data that we record. The inverses of these matrices allow us to decompose multicomponent data.

In the previous solutions [see for example, Ikelle and Amundsen (2005)], we generally used deterministic ways of estimating the mixing matrices of wavefield decomposition (that is, the mixing matrices are derived from the wave-equation theory often under the assumption that the sea floor is flat). The use of these estimates in wavefield decomposition requires that data be uniformly sampled, especially when working in the frequency-wavenumber domain, and nonaliased to facilitate the computations of the Fourier transforms with respect to spatial coordinates. For wavefield decomposition in land data, the estimation of the mixing matrices also requires knowledge of near-surface physical properties. For wavefield decomposition in OBS data, the estimation of the mixing matrices requires knowledge of the physical properties just below the seafloor. These requirements may limit the use of the deterministic solutions for OBS data when the water-solid interface is not horizontally flat.

In this paper, we describe statistical ways of estimating these mixing matrices. The basic idea is that by taking advantage of the redundancy of seismic data, we can seek to simultaneously reconstruct the mixing matrices, the upgoing and downgoing wavefields by essentially using the statistical methods described in Ikelle (2010). The attractive features of the statistical wavefield methods are that they are independent of the data dimension (i.e., they are applicable to 2D and 3D data without any modification of the computer code and without a sampling or interpolation requirement) and that they work for aliased data as well as for nonaliased data, and even for nonuniform sampled data. The only assumptions here are that (1) the mixing matrix is invertible and (2) the number of datapoints is large enough for statistical applications.

## FUNDAMENTALS OF STATISTICAL WAVEFIELD SEPARATION

Up-down separation problems can be cast in one of the following forms:

$$D_i(\alpha_1,\alpha_2,\alpha_3) = \sum a_{i\,j} F_j(\alpha_1,\alpha_2,\alpha_3) \quad , \tag{1}$$

$$D_i(\alpha_1,\alpha_2,\alpha_3) = \sum a_{i\,j}(\alpha_1) F_j(\alpha_1,\alpha_2,\alpha_3) \quad , \tag{2}$$

$$D_i(\alpha_1,\alpha_2,\alpha_3) = \sum a_{i\,j}(\alpha_1,\alpha_2) F_j(\alpha_1,\alpha_2,\alpha_3) \quad , \tag{3}$$

where $D_i$ are the multicomponent input data and $F_i$ are the separated data, where i and j vary from 1 to N. N can take the values 2, 3, ..., 9. We did not explicitly introduce the specific variables of the fields $D_i$ and $F_i$ because the forms in (1), (2), and (3) can be derived in the T-X domain, the F-X domain, the F-K domain, and even the $\tau$-p domain. Instead we use arbitrary variables $\alpha_1$,$\alpha_2$, and $\alpha_3$, which here represent the coordinate systems of the data in a given domain. Notice that (1) has a typical ICA form (see Ikelle, 2010). Eqs. (2) and (3) also have ICA forms if we process them as a series of ICAs for each $\alpha_3$ in the case of (2) and for each pair ($\alpha_1$,$\alpha_2$) in the case of (3). So we can then apply the ICA algorithms captured in Ikelle (2010) to recover $F_1$, ..., $F_N$ , along with the elements $a_{ij}$ of the mixing matrix - and therefore to perform up-down wavefield separation. These algorithms can be performed at any point in the subsurface because their application does not require any knowledge of the elastic parameters of the subsurface and because the ICA methods allow us to recover $a_{ij}$ and $F_i$ simultaneously. However, there are two fundamental requirements that must be fulfilled for using ICA methods to decompose wavefields. The first one is that the number samples must be large, more than 2,000. In other words, the number of triplets ($\alpha_1$,$\alpha_2$,$\alpha_3$) must be large in the case of the forms in (1), the number of pairs ($\alpha_1$,$\alpha_2$) in the case of the forms in (1) must be large, and the number of $\alpha_1$ must be large in the case of the forms in (1). As each of the coordinates in seismic data takes at least 2,000 values, this requirement will be generally fulfilled. The other requirement is that the fields $F_i$ must be statistically independent. Note that the variations of the mixing matrix with $\alpha_1$, as described in (2), allow us to effectively perform the data decomposition even when the upgoing and downgoing wave components are both not present at the hydrophone and the geophone at the same time, especially when working F-X or F-K domains.

When the receivers are very close to the sea surface, as in towed-streamer acquisition, the upgoing and downgoing wavefields are highly correlated and therefore not statistically independent. The ICA is not effective for such decomposition. However, for ocean-bottom acquisition, the ICA can still be effective in up-down separation because the upgoing and downgoing wavefields rarely interfere in this case.

## CONSTRAINED NEGENTROPY MAXIMIZATION

As in all ICA solutions, the wavefield decomposition based on ICA will suffer from an inherent indeterminacy on scale (dilation) and permutation. This indetermination can be reduced or even eliminated by using the constrained ICA (cICA), also known as ICA with reference. The cICA consists of reconstructing the independent components from their mixtures by using the objective function of the algorithms described in Ikelle (2010), with additional constraints and a priori information about the independent components and/or the demixing matrix. In other words, with additional conditions, the ICA problem becomes a constrained optimization problem (i.e., a constrained ICA problem). Moreover, with appropriate constraints the cICA can also be used to extract independent components from mixtures contaminated by non-Gaussian noise and to better deal with the ICA decomposition in which some independent components are very weak in amplitude compared to the other independent components in the mixtures. We here describe an cICA algorithms which is based on negentropy maximization.

Let us denote by $\mathbf{Y} = [\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_I]^T$ the mixed data, $\mathbf{X} = [\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_I]^T$ the independent component data. We start by whitening $\mathbf{Z}$; that is, we go from mixtures which are correlated and dependent to new mixtures which correspond to mixtures that are uncorrelated but remain statistically dependent. Mathematically, we can describe this process as finding a whitening matrix, $\mathbf{V}$, that allows us to transform the random vector, $\mathbf{Y}$, to another random vector, $\mathbf{Z} = [\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_I]^T$; i.e.,

$$\mathbf{z}_i = \sum_{k=1}^{I} v_{ik}\, \mathbf{y}_k \quad . \tag{4}$$

The whitening problem comes down to finding a $\mathbf{V}$ for which the covariance matrix of $\mathbf{Z}$ is the identity matrix; i.e.,

$$E[\mathbf{Z}\mathbf{Z}^T] = \mathbf{E}_Y \mathbf{L}_Y^{-1/2} \mathbf{L}_Y^{-1/2} \mathbf{E}_Y^T = \mathbf{I} \quad , \tag{5}$$

where $\mathbf{E}_Y$ is an orthogonal matrix and $\mathbf{L}_Y$ is a diagonal matrix with all nonnegative eigenvalues $\lambda_i$; that is, $\mathbf{L}_Y = \mathrm{Diag}(\lambda_1,\lambda_2,\ldots,\lambda_I)$. The columns of the matrix $\mathbf{E}_Y$ are the eigenvectors corresponding to the appropriate eigenvalues. So the matrix $\mathbf{V}$, which allows us to obtain the whiten random vector $\mathbf{Z}$, can be computed as follows:

$$\mathbf{V} = \mathbf{L}_Y^{-1/2}\mathbf{E}_Y \quad . \tag{6}$$

The central problem here is the reconstruction of $\mathbf{X}$ from $\mathbf{Z}$; that is, finding $\mathbf{W}$ such that $\mathbf{X}(\mathbf{X} = \mathbf{W}\mathbf{Z})$ is made of independent components.

The objective function of cICA can be written as follows:

$$\begin{cases} \text{maximize} \quad O(\mathbf{X}) \\ \text{subject to} \quad \mathbf{h}(\mathbf{W}) \le 0, \ \text{and/or} \ \mathbf{c}(\mathbf{W}) = 0 \end{cases} , \tag{7}$$

where $O(\mathbf{X})$ represents the ICA contrast function to be minimize or maximize,

$\mathbf{W} = [\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_I]^T$ represents the ICA demixing matrix, $\mathbf{h}(\mathbf{W}) = [h_1(\mathbf{W}),h_2(\mathbf{W}),\ldots,h_M(\mathbf{W})]^T$ represent a set of M inequality constraints, and $\mathbf{c}(\mathbf{W}) = [c_1(\mathbf{W}),c_2(\mathbf{W}),\ldots,c_N(\mathbf{W})]^T$ represent a set of N equality constraints. These constraints allow us to introduce some *a priori* information about the mixing matrix and the desired independent components. For example, we can introduce *a priori* information about the desired independent component as a reference component, $\mathbf{r}_i$, to obtain an output which is a desired independent component. The reference signal must carry some information about the desired independent component. It does not need to be a prefect match, but it should be close enough to point the algorithm in the direction of a particular independent component. The closeness constraint can be written as

$$h_i(\mathbf{x}_i,\mathbf{r}_i) = \epsilon(\mathbf{x}_i,\mathbf{r}_i) - \zeta \leq 0 \ , \tag{8}$$

where $h_i$ is an inequality constraint function, $\epsilon$ is a closeness measure, and $\zeta$ is a closeness threshold parameter. Note that $\mathbf{r}_i$ can be a reference for $\mathbf{w}_i$. The constraint in this case is

$$h_i(\mathbf{w}_i,\mathbf{r}_i) = \epsilon(\mathbf{w}_i,\mathbf{r}_i) - \zeta \leq 0 \ . \tag{9}$$

Similar measures can also be used for the row of the demxing matrix. The measure of closeness can take any form, such as a mean squared-error, $\epsilon(\mathbf{w}_i,\mathbf{r}_i) = E[(\mathbf{w}_i - \mathbf{r}_i)^2]$ (with the requirement that both $\mathbf{x}_i$ and $\mathbf{r}_i$ are normalized to have the same mean and variance) and correlation, $\epsilon(\mathbf{x}_i, \mathbf{r}_i) = 1/(E[(\mathbf{x}_i\mathbf{r}_i)])^2$ (both $\mathbf{x}_i$ and $\mathbf{r}_i$ need to be normalized to ensure that the value of correlation is bounded). Also the inner product can be used to measure the distance between $\mathbf{r}_i$ and either $\mathbf{x}_i$ or $\mathbf{w}_i$:

$$\epsilon(\mathbf{x}_i, \mathbf{r}_i) = \begin{cases} |\mathbf{x}_i^T\mathbf{r}_i| & \text{for } \mathbf{x}_i \\ \\ |\mathbf{w}_i^T\mathbf{r}_i| & \text{for } \mathbf{w}_i \end{cases} \ . \tag{10}$$

In our numeral implementation, we will use the mean-squared-error and scalar product as measures of closeness.

For a case in which constraints are simply used for ordering and normalizing independent components, we can use $h_i(\mathbf{W}) = I(\mathbf{x}_{i+1}) - I(\mathbf{x}_i)$ to define the descending order of independent components, where $I(\mathbf{x}_i)$ is the index of some statistical measure of the recovered independent components - for example, variance, $E[\mathbf{x}_i^2]$, or normalized kurtosis, $E[\mathbf{x}_i^4]/E[\mathbf{x}_i^2] - 3$. The set of N equality constraints are defined as $c_i(\mathbf{W}) = (\mathbf{w}_i^T\mathbf{w}_i) - 1$, $i = 1, \ldots, I$.

Here we will focus on the particular case in which $O(\mathbf{X})$ represents the negentropy maximization; i.e.,

$$O(\mathbf{X}) = \sum_{\mathrm{i}} O_{\mathrm{i}}(\mathbf{x}_{\mathrm{i}})$$

$$= \sum_{\mathrm{i}} [\mathrm{E}\{\mathrm{G}(\mathbf{w}_{\mathrm{i}}\mathbf{Z})\} - \mathrm{E}\{\mathrm{G}(\mathbf{x}_{\mathrm{G}})\}]^2 \ , \qquad (11)$$

with

$$O_{\mathrm{i}}(\mathbf{x}_{\mathrm{i}}) = [\mathrm{E}\{\mathrm{G}(\mathbf{w}_{\mathrm{i}}\mathbf{Z})\} - \mathrm{E}\{\mathrm{G}(\mathbf{x}_{\mathrm{G}})\}]^2 \ , \qquad (12)$$

where $\mathbf{x}_{\mathrm{G}}$ is a set of Gaussian random variables with a zero mean and a unit variance, and where G(.) is one of the nonquadratic functions, such as

$$\mathrm{G}(\mathrm{x}) = \log[\cosh(a_1\mathrm{x})/a_1] \ , \qquad (13)$$

$$\mathrm{G}(\mathrm{x}) = \exp(-\mathrm{x}^2/2) \ , \qquad (14)$$

and

$$\mathrm{G}(\mathrm{x}) = \mathrm{x}^4/4 \ , \qquad (15)$$

where $1 \leq a_1 \leq 2$.

We here present a technique for solving the constrained optimization problem in (7) for reconstructing one row of **W** at a time. This technique is generally known as the one-unit ICA. We will use the method of Lagrange multipliers to solve this optimization problem. The augmented Lagrangian function of (7) can be written as follows (Bertsekas, 1996):

$$L(\mathbf{w}_{\mathrm{i}},\mu_{\mathrm{i}},\lambda_{\mathrm{i}},\upsilon_{\mathrm{i}}) = O_{\mathrm{i}}(\mathbf{x}_{\mathrm{i}}) + \mu_{\mathrm{i}}[\mathrm{h}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \upsilon_{\mathrm{i}}^2] + \tfrac{1}{2}\gamma[\mathrm{h}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \upsilon_{\mathrm{i}}^2]^2$$
$$+ \lambda_{\mathrm{i}}\mathrm{c}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \tfrac{1}{2}\gamma[\mathrm{c}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}})]^2 \ , \qquad (16)$$

where $\mu_{\mathrm{i}}$ and $\lambda_{\mathrm{i}}$ are the positive Lagrange multipliers corresponding to inequality and equality constraints, respectively, and $\gamma$ is the penalty parameter. Notice that the inequality constraint in (7) is incorporated into the ICA as an equality constraint using the slack variables, $\upsilon_{\mathrm{i}}$. In other words, we have transformed inequality constraints $\mathrm{h}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) \leq 0$ into equality constraints $\mathrm{h}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \upsilon_{\mathrm{i}}^2 = 0$ by using slack variables. The equality constraint is added to the cost function using the method of multipliers. The quadratic penalty terms ensure that the local convexity assumption holds. To find the optimal $\upsilon_{\mathrm{i}}$ that maximizes $L_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}},\mu_{\mathrm{i}},\lambda_{\mathrm{i}},\upsilon_{\mathrm{i}})$, we first need to rewrite (16) as follows:

$$L_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}},\mu_{\mathrm{i}},\lambda_{\mathrm{i}},\upsilon_{\mathrm{i}}) = O_{\mathrm{i}}(\mathbf{x}_{\mathrm{i}}) + (\gamma/2)[\mathrm{h}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \upsilon_{\mathrm{i}}^2 + (\mu_{\mathrm{i}}/\gamma)]^2 - (\mu_{\mathrm{i}}^2/2\gamma)$$
$$+ \lambda_{\mathrm{i}}\mathrm{c}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}}) + \tfrac{1}{2}\gamma[\mathrm{c}_{\mathrm{i}}(\mathbf{w}_{\mathrm{i}})]^2 \ . \qquad (17)$$

To maximize $L_i(\mathbf{w}_i,\mu_i,\lambda_i,\upsilon_i)$, $\upsilon_i$ must satisfy the following condition:

$$\upsilon_i^2 = \begin{cases} -h_i(\mathbf{w}_i) + \mu_i/\gamma \ , & a_i \leq 0 \\ 0 \ , & a_i > 0 \end{cases} \ , \tag{18}$$

or

$$\upsilon_i^2 = (1/\gamma)\max\{0, -[\gamma h_i(\mathbf{w}_i) + \mu_i]\} \ , \tag{19}$$

where $a_i = h_i(\mathbf{w}_i) + \mu_i$. By substituting (19) in (17), we finally obtain the following simplified augmented Lagrangian-optimization function

$$L_i(\mathbf{w}_i,\mu_i,\lambda_i) = O_i(\mathbf{x}_i) + [(1/2\gamma)(\max\{0,[\gamma h_i(\mathbf{w}_i) + \mu_i]\})^2 - \mu_i^2] + \lambda_i c_i(\mathbf{w}_i) + \tfrac{1}{2}\gamma[c_i(\mathbf{w}_i)]^2 \ . \tag{20}$$

To find the maximum of $L_i(\mathbf{w}_i,\mu_i,\lambda_i)$, we iteratively update $\mathbf{w}_i$ in the Newton-like method. We go from one iteration (described by $\mathbf{w}_i$) to another iteration (described by $\mathbf{w}_i'$) as follows:

$$\mathbf{w}_i' = \mathbf{w}_i - \eta[\partial^2 L_i(\mathbf{w}_i,\mu_i,\lambda_i)/\partial\mathbf{w}_i^2]^{-1}\, \partial 2L_i(\mathbf{w}_i,\mu_i,\lambda_i)/\partial\mathbf{w}_i \ . \tag{21}$$

$\eta$ is a positive stepsize added to avoid the uncertainty in the convergence. After some algebra, we arrive at

$$\mathbf{w}_i' = \mathbf{w}_i - \eta\{\alpha E[\mathbf{Z}G'(\mathbf{w}_i\mathbf{Z})] - \mu_i E[\mathbf{Z}h_i'(\mathbf{w}_i)] - 2\lambda_i E[\mathbf{Z}^T\mathbf{x}_i]\} / \mathbf{C}_z\{\alpha E[G''(\mathbf{Z}\mathbf{x}_i^T)] - \mu_i E[h_i''(\mathbf{w}_i)] - 2\lambda_i\} \ , \tag{22}$$

with

$$\alpha = 2\text{sign}(E[G(\mathbf{w}_i\mathbf{Z})] - E[G(\mathbf{x}_G)]) \ , \tag{23}$$

where $G'(u)$ and $G''(u)$ are the first and second derivatives, respectively, of $G(u)$ with respect to u, $h_i'(u)$ and $h_i''(u)$ are the first and second derivatives, respectively, of $h_i(u)$ with respect to u, and $\mathbf{C}_Z$ is the covariance of $\mathbf{Z}$. The estimated vector must be normalized as follows

$$\mathbf{w}_i = \mathbf{w}_i'/\|\mathbf{w}_i'\| \ . \tag{24}$$

There are several methods for updating Lagrange multipliers. Here, we use the gradient ascent method (Bertsekas, 1996):

$$\mu_i' = \mu_i + \max\{-\mu_i, \gamma h_i(\mathbf{w}_i)\} \ , \tag{25}$$

$$\lambda_i' = \lambda_i + \gamma c_i(\mathbf{w}_i) \ . \tag{26}$$

Table 1. The key steps of the one-unit negentropy-based cICA based on the negentropy-maximization criterion.

1. Center the observed signals **Y** to make them mean zero and whiten them to **Z**.
2. Set e to a small value to control the learning precision.
3. Set the number I of independent components to be recovered.
4. Choose penalty parameter $\gamma$.
5. Set i = 1 for the first independent components to be recovered.
6. Choose an initial value for the Lagrange multiplier, $\mu_i$.
7. Choose an initial value for the Lagrange multiplier, $\lambda_i$.
8. Initialize $\mathbf{w}_i$ with a random vector and set its norm to 1.
9. Update $\mathbf{w}_i$ by using (22).
10. Normalize $\mathbf{w}_i$ by using (24).
11. Update $\mu_i$ by using (25).
12. Update $\lambda_i$ by using (26).
13. Go back to step 9 until convergence is achieved (i.e., until $|1 - \|\mathbf{w}_i^T\mathbf{w}_i'\| \, | \leq \epsilon$).
14. Set i = i + 1 and go back step 6, until i > I.

Eqs. (22), (23), (24), (25), and (26) together constitute a negentropy-based cICA method. The key steps in this algorithm are described in Table 1.

To numerically validate the cICA algorithm, let us first discuss the standard ICA results. We consider the single-shot gathers in Fig. 1 and the corresponding multishot data in Fig. 2. Notice also that the decoded results in Fig. 3 are almost identical to the original single-shot gathers shown in Fig. 1. However, we have to adjust the amplitudes.

Let us now look at some numerical illustrations of cICA. We consider the decoding of the mixtures in Fig. 2. The single-shot gathers in these mixtures are responses to the same subsurface models for the same receiver array but different source points (i.e., $x_s$ = 0 m for the single-shot gather in Fig. 1a, $x_s$ = 875 m for the single-shot gather in Fig. 1b, and $x_s$ = 1700 m for the single-shot gather in Fig. 1c). We have chosen as a reference single-shot gather

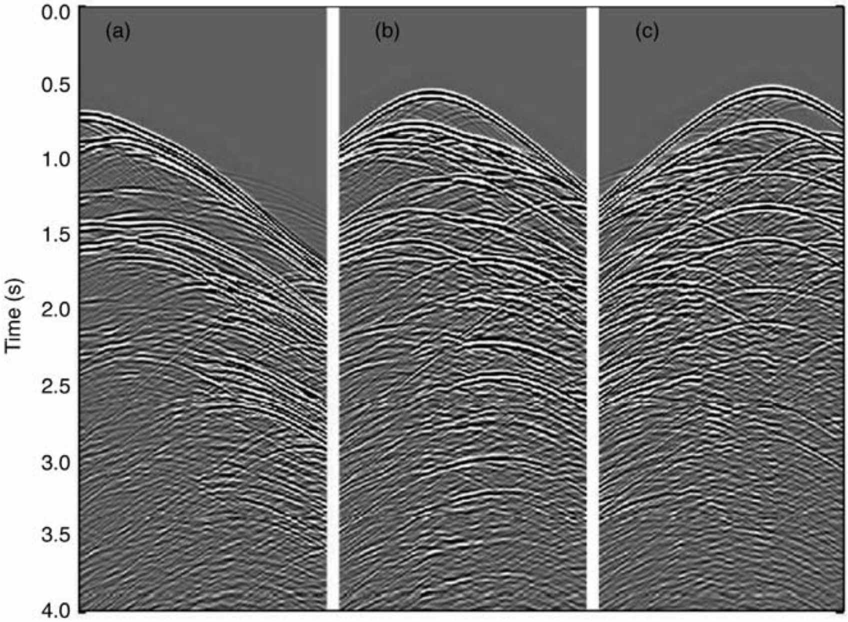

Fig. 1. Three single-shot gathers representing three responses to the same subsurface models for the same receiver array but different source points (i.e., $x_s = 0$ m for the single-shot gathers in Fig. 1a, $x_s = 875$ m for the single-shot gather in Fig. 1b, and $x_s = 1700$ m for the single-shot gather in Fig. 1c).

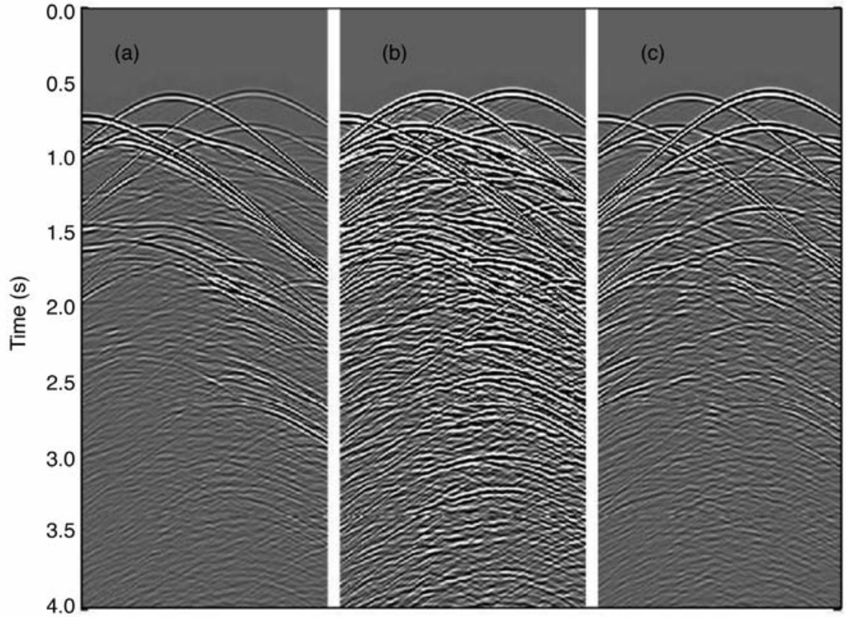

Fig. 2. Three mixtures of the single-shot gathers in Fig. 1.

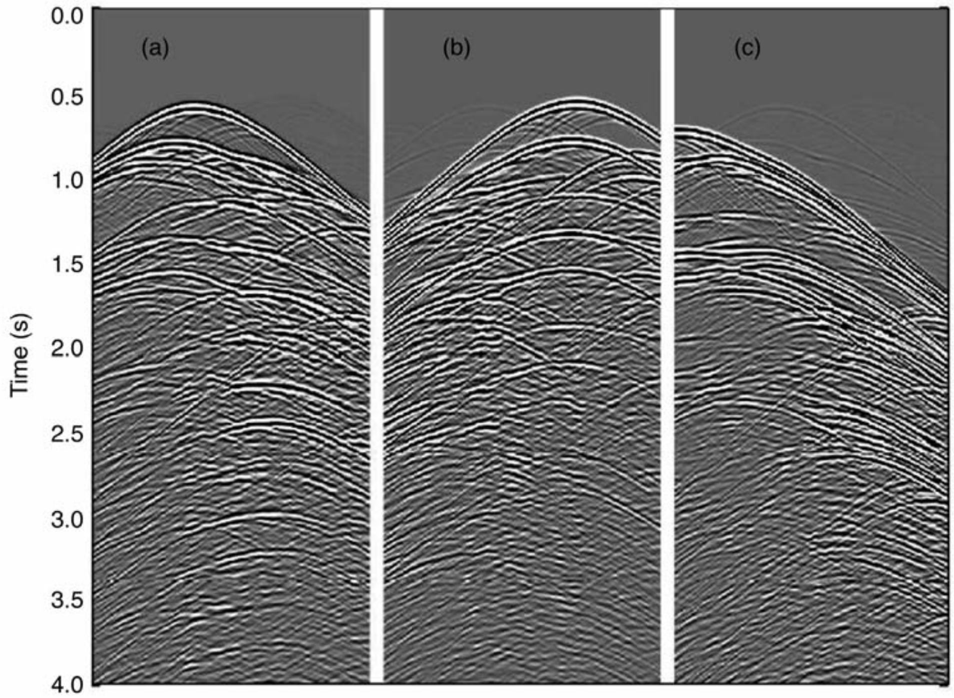

Fig. 3. The SOBI decoded results of the mixtures in Fig. 2. When compared to the original single-shot gathers in Fig. 1, these results are quite satisfactory.

a single-shot gather which is a response to the same subsurface models for the same receiver array as the actual single-shot gathers but with a different source point which is at $x_s = 250$ m. Therefore we expect this reference to allow us to recover the single-shot gather at $x_s = 0$ m. One can think of this reference as representing a gather extracted from an survey several years old. As illustrated in Fig. 4, we can recover quite well the desired single-shot gather (i.e., the gather with the shot point at $x = 0$) by using cICA without the permutation and even the scale issues.

## UPGOING-DOWNGOING SEPARATION

One cICA feature allows us to decode some non-instantaneous mixtures if we can build a good reference model. Consider the problem of up-down separation [see Box 2.7, Ikelle and Amundsen (2005), Chapter 9)]. Suppose that we recorded the pressure and the vertical component of the particle velocity. For a shot point at $\mathbf{x}_s$, we could then write the pressure and the vertical component of the particle velocity as mixtures; i.e.,

$$\begin{cases} P(\mathbf{x},t) = P_U(\mathbf{x},t) + P_D(\mathbf{x},t) \\ V(\mathbf{x},t) = V_U(\mathbf{x},t) + V_D(\mathbf{x},t) \end{cases} , \tag{27}$$

or

$$\begin{cases} P(\mathbf{x},t) = P_U(\mathbf{x},t) + P_D(\mathbf{x},t) \\ V(\mathbf{x},t) = \gamma(\mathbf{x},t)P_U(\mathbf{x},t) - \gamma(\mathbf{x},t)P_D(\mathbf{x},t) \end{cases} , \tag{28}$$

with

$$V_U(\mathbf{x},t) = \gamma(\mathbf{x},t)P_U(\mathbf{x},t) \text{ and } V_D(\mathbf{x},t) = -\gamma(\mathbf{x},t)P_D(\mathbf{x},t) \ , \tag{29}$$

where $P(\mathbf{x},t)$ is the pressure field, $V(\mathbf{x},t)$ is the vertical component of the particle velocity, $P_U(\mathbf{x},t)$ is the upgoing pressure field, $P_D(\mathbf{x},t)$ is the downgoing pressure field, $V_U(\mathbf{x},t)$ is the upgoing particle-velocity field, $V_D(\mathbf{x},t)$ is the downgoing particle-velocity field, and $\gamma(\mathbf{x},t)$ is a scaling factor which causes V to be a non-instantaneous mixture. We now consider the pressure in Fig. 3a and the vertical component of the particle velocity in Fig. 4a as input data. Using the standard ICA decoding, we will end up with one of the following four permutations $(\alpha_1P_U,\beta_1P_D)$, $(\alpha_2P_U,\beta_2V_U)$, $(\alpha_3P_U,\beta_3V_D)$, $(\alpha_4P_D,\beta_4V_U)$, and $(\alpha_5P_D,\beta_5V_D)$ - where $\alpha_i$ and $\beta_i$ are unknown constants. Because the direct wave is not well separated from the rest of the data, it is difficult to blindly identify the permutation that ICA will output, as all the permutations have an equal chance of being an output of ICA. Let us now look at how cICA can help us solve this problem. The physics of this problem suggests that for receiver points $\mathbf{x}$, located near the source point $\mathbf{x}_s$, we can assume that $\gamma(\mathbf{x},t) \approx 1/Z$, where Z is the P-wave impedance of the water column; the P-wave impedance of the water column is constant, and a well known quantity. In other words, (28) can be approximated as

$$\begin{cases} P(\mathbf{x},t) \approx P_U'(\mathbf{x},t) + P_D'(\mathbf{x},t) \\ ZV(\mathbf{x},t) \approx P_U'(\mathbf{x},t) - P_D'(\mathbf{x},t) \end{cases} , \tag{30}$$

which lead to

$$\begin{cases} P_U'(\mathbf{x},t) \approx ½[P(\mathbf{x},t) + ZV_z(\mathbf{x},t)] \\ P_D'(\mathbf{x},t) \approx ½[P(\mathbf{x},t) - ZV_z(\mathbf{x},t)] \end{cases} . \tag{31}$$

By using the solution $P_U'$ (Fig. 5b) as the reference component, for example, we can use cICA to recover the actual upgoing pressure from P (Fig. 5a) and V (Fig. 6a), as illustrated in Fig. 5c. Note that we do not have permutation and scaling issues with this solution. By scaling the pressure field in (30) by 1/Z, we can set up the system in (30) to reconstruct $V_U'$ and $V_D'$ instead of $P_U'$ and $P_D'$, respectively. Hence by using the solution $V_U'$ (Fig. 6b), for example, as the reference component, we can run cICA again to recover the actual upgoing particle-velocity field from P (Fig. 5a) and V (Fig. 6a), as illustrated in Fig. 6c. Note that the three plots shown in Fig. 5 are plotted at the same scale. This observation is also valid for plots shown in Fig. 6.

We have indicated in raw data in Figs. 5a and 6a some examples of downgoing events which are well attenuated in the upgoing wavefields obtained by cICA as shown in Figs. 5c and 6c. In other words, the cICA has recovered the independent components which are closer to the reference components (i.e., the upgoing wavefields) instead of some arbitrary component as the standard ICA. Moreover, cICA has recovered the scales of upgoing wavefields because their scales are directly comparable to those of raw data in these figures.

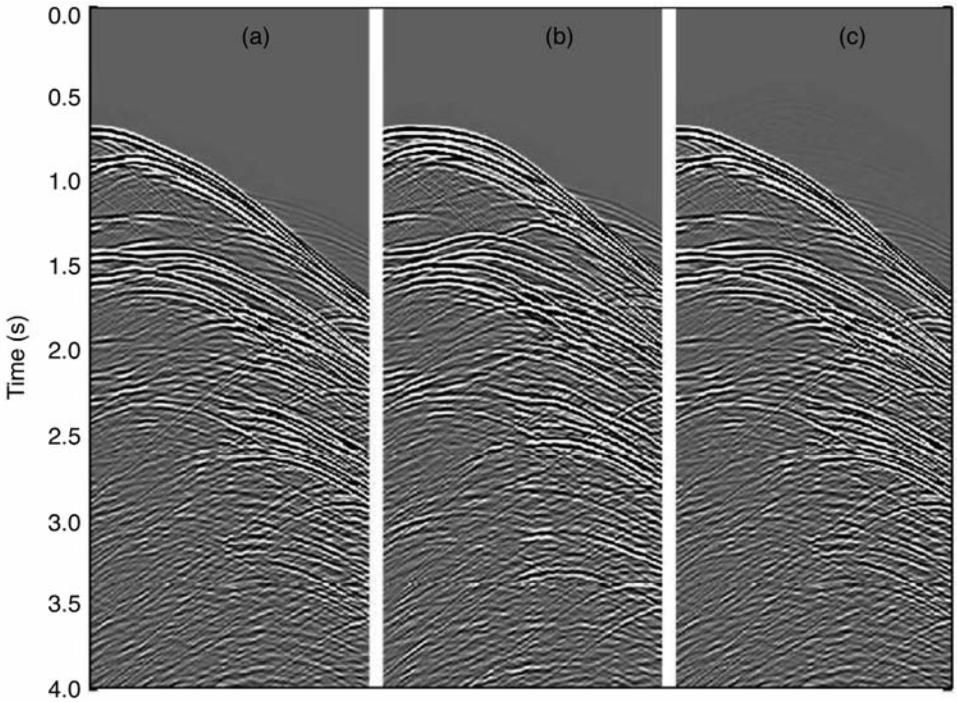

Fig. 4. The cICA decoded results of the mixtures in Fig. 2 with as reference single-shot gather which is a response to the same subsurface model for the same receiver array as the actual single-shot gathers but with a different source point, which is at 250 m. The desired single-shot gather is a gather with the shot point at x = 0 (Fig. 1a). (a) The original single-shot gather, (b) the reference single-shot gather, and (c) the decoded single-shot gather by using cICA without the permutation and even without scale issues. When compared to the original single-shot gather, these decoding results are quite satisfactory.

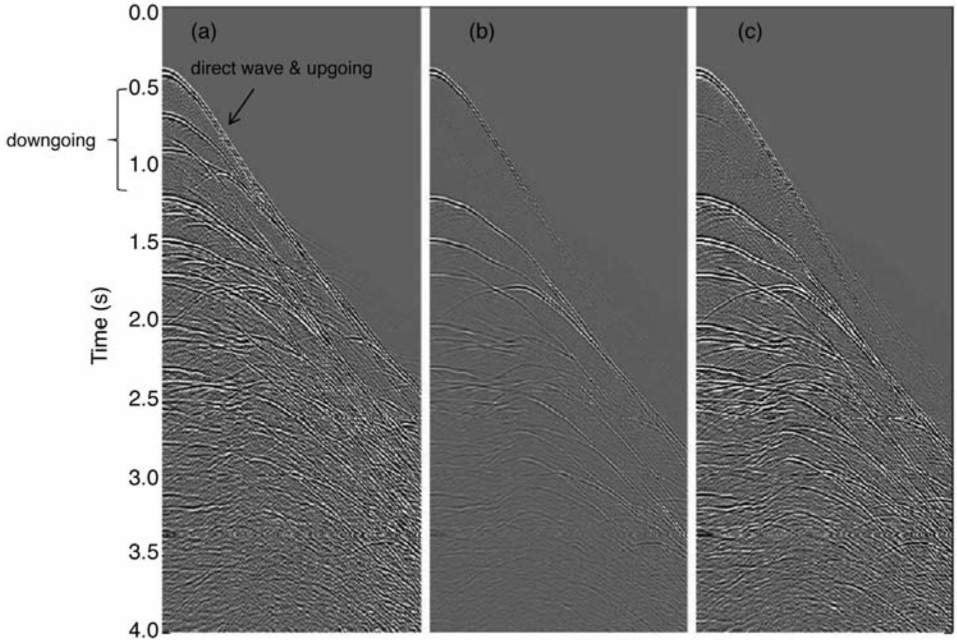

Fig. 5. (a) The pressure data. They are here treated, along with the vertical component of the particle-velocity data, as two mixtures for the upgoing and downgoing separation. The offset range is from 0 to 4.5 km. (b) The reference upgoing pressure data. (c) The upgoing wavefield obtained by using cICA. All three plots are at the same scale.

Note also that the polarity of the upgoing pressure field (Fig. 5c) is different from the polarity of the upgoing particle-velocity field (Fig. 6c), despite the fact that we have used the same inputs (i.e., Figs. 5a and 6a) in the two runs of cICA. This observation indicates that cICA has recovered the desired component in each of the cICA runs.

## CONCLUSION

We have derived and implemented a new method of performing up-down separation of OBS wavefield. This method is based on independent component analysis. It is applicable to 2D and 3D data, to aliased data as well as aliased data, and to nonuniform sampled data.

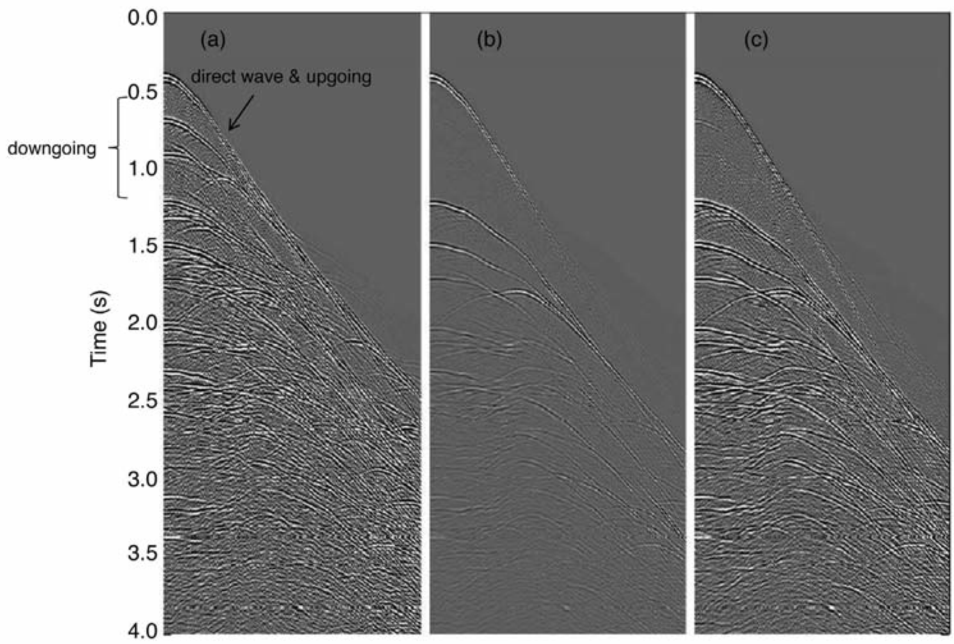

Fig. 6. (a) The vertical component of the particle-velocity data. It is here treated, along with the pressure data as two mixtures for the purpose of the upgoing and downgoing separation. The offset range is from 0 to 4.5 km. (b) The reference upgoing pressure data. (c) The upgoing wavefield obtained by using cICA. All three plots are at the same scale.

## REFERENCES

Bertsekas, D.P., 1996. Constrained Optimization and Lagrange Multiplier Methods. Athena Scientific, Belmont, MA.

Bertsekas, D.P., 1998. The auction algorithm: A distributed relaxation method for the assignment problem. Ann. Oper. Res., 14: 103-123.

Ikelle, L.T., 2010. Coding and Decoding: Seismic Data. Elsevier Science Publishers, Amsterdam.

Ikelle, L.T. and Amundsen, L., 2005. Introduction to Petroleum Seismology. SEG, Tulsa, OK.