A CONVEX VARIATIONAL APPROACH FOR MULTIPLE REMOVAL IN SEISMIC DATA

D. Gragnaniello, C. Chaux, and J.-C. Pesquet

Univ. Paris-Est, LIGM UMR CNRS 8049
5 Bd Descartes, 77454 Marne-la-Vallée - France
e-mail: d.gragnaniello@gmail.com
{chaux,pesquet}@univ-mlv.fr

L. Duval

IFP Energies nouvelles, Dir. Technologie
1 et 4 Av de Bois-Préau
92852 Rueil-Malmaison Cedex - France
e-mail: laurent.duval@ifpen.fr

ABSTRACT

Due to complex subsurface structure properties, seismic records often suffer from coherent noises such as multiples. These undesired signals may hide the signal of interest, thus raising difficulties in interpretation. We propose a new variational framework based on Maximum A Posteriori (MAP) estimation. More precisely, the problem of multiple removal is formulated as a minimization problem involving time-varying filters, assuming that a disturbance signal template is available and the target signal is sparse in some orthonormal basis. We show that estimating multiples is equivalent to identifying filters and we propose to employ recently proposed convex optimization procedures based on proximity operators to solve the problem. The performance of the proposed approach as well as its robustness to noise is demonstrated on realistically simulated data.

Index Terms— convex optimization, wavelets, time-varying filters, regularization.

1. INTRODUCTION

Standard reflection seismology infers subsurface structure properties from seismic waves generated at their interfaces and recorded after propagation in the medium (see Fig. 1). Due to wave propagation physics and the complicated nature of subsurface layers, seismic data is subject to a variety of distortions and disturbances [1, 3] that hinder subsequent geophysical interpretation. The complexity of these data has contributed to the development of several efficient signal processing tools; for instance wavelets [2] or robust ℓ_1-based sparse restoration [4], which are now commonly used in signal and image analysis [5].

Multiples correspond to a specific type of unwanted coherent seismic events related to wave field reflection bounces inside layers [6]. These reverberations sometimes are strong enough to obscure deeper target reflectors of geological interest, which may become nearly invisible. Their attenuation thus represents one of the greatest challenges in present seismic processing. Since they relate to reflections on above-situated layers, they are highly correlated and possess spectral contents similar to deeper reflections. One of the most effective multiple subtraction is based on prior computation of a multiple model, denoted hereafter as “template”. A template is adapted to the observed data, traditionally using least-squares criteria, and subsequently subtracted. It is generally assumed that the matching should be performed in a non-stationary fashion, as the transfer function between the template and the actual multiple signal should account for time-dependent wave distortion. Moreover, the matched filter should compensate for inaccuracies in template modeling, especially in amplitude and delay. Recently, several works in geophysics have revisited the use of non-quadratic criteria, including Huber function [7] or ℓ_p $(p \in [1, 2])$ norms, due to the alleged non-Gaussianity of seismic data [8].

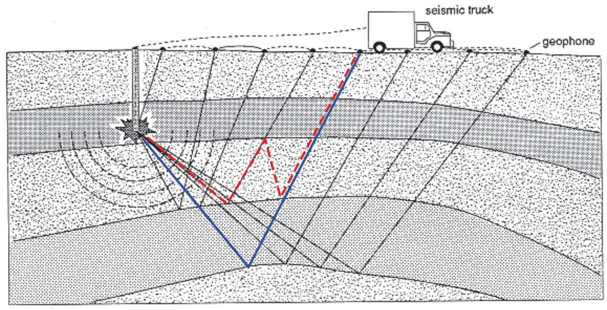

Fig. 1. Principles of seismic wave propagation, with reflections on different layers, and data acquisition. Solid blue: primary; dashed red: multiple.

The present work proposes an original approach to the multiple removal problem. In standard restoration [9], knowledge about the degradation kernel is often required. It is replaced here by the knowledge of a template, the degradation kernel being estimated. The resulting algorithm recursively estimate a Finite Impulse Response (FIR) filter, which is constrained to exhibit slow variations over time. This hypothesis is highly consistent with wave propagation assumptions. More precisely, the estimation problem is formulated as a convex variational problem involving a non-smooth cost function. The paper is structured as follows: Section 2 introduces the considered model, Section 3 presents the proposed MAP-based methodology, followed in Section 4 by reminders on proximal methods. The proposed approach is validated in Section 5 on realistically simulated multiple contaminations.

2. MODELING MULTIPLE REFLECTIONS

This section aims at describing the employed model, accounting for multiple reflections in seismic data. More precisely, we assume that an array of sensors delivers data

$$z^{(n)} = s^{(n)} + y^{(n)} \tag{1}$$

where $n \in \{0, \cdots, N-1\}$ is the time index, $z = (z^{(n)})_{0 \le n < N}$ are the observed data, combining the primary $y = (y^{(n)})_{0 \le n < N}$ (signal of interest, unknown), and the multiples $(s^{(n)})_{0 \le n < N}$ (sum of undesired reflected signals), depicted in solid blue and dashed red in Fig. 1.

We assume that a template $(r^{(n)})_{0 \le n < N}$ for the disturbance signal is available, which is related to $(s^{(n)})_{0 \le n < N}$ through an FIR non-causal convolutive model

$$s^{(n)} = \sum_{p=p'}^{p'+P-1} h^{(n)}(p) r^{(n-p)} \tag{2}$$

where $h^{(n)}$ is an unknown impulse response corresponding to time n and where $p' \in \{-P+1, \cdots, 0\}$ ($p' = 0$ corresponds to the causal case). It must be emphasized that the dependence w.r.t. the time index n of the impulse response implies that the filtering process is not time invariant, although it can be assumed slowly varying in practice. Eq. (2) can be expressed more concisely as

$$s = Rh \tag{3}$$

by appropriately defining vectors s, h and matrix R. More precisely,

$$s = \begin{bmatrix} s^{(0)} & \cdots & s^{(N-1)} \end{bmatrix}^\top, \tag{4}$$

$$h = \Big[h^{(0)}(p') \cdots h^{(0)}(p'+P-1) \cdots \\ \cdots h^{(N-1)}(p') \cdots h^{(N-1)}(p'+P-1)\Big]^\top \tag{5}$$

and

$$R = \begin{bmatrix} R^{(0)} & 0 & \cdots & 0 \\ 0 & R^{(1)} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & R^{(N-1)} \end{bmatrix} \tag{6}$$

where

$$\Big[(R^{(0)})^\top (R^{(1)})^\top \cdots (R^{(N-1)})^\top\Big]^\top = \\ \begin{bmatrix} r^{(-p')} & \cdots & r^{(0)} & 0 & \cdots & 0 \\ r^{(-p'+1)} & \cdots & & r^{(0)} & 0 \cdots & 0 \\ \vdots & & & & & \\ r^{(N-1)} & r^{(N-2)} & & \cdots & & r^{(N-P)} \\ 0 & r^{(N-1)} & & \cdots & & r^{(N-P+1)} \\ \vdots & & & & & \\ 0 & \cdots & 0 & r^{(N-1)} & \cdots & r^{(N-P-p')} \end{bmatrix}. \tag{7}$$

One can note that the matrix R is a block diagonal matrix and that the concatenation of its block diagonal elements is a Toeplitz matrix of size $N \times P$.

With this formulation, the problem of providing an estimate $\widehat{y}$ of the primary turns out to be equivalent to computing an estimate $\widehat{h}$ of the impulse response and to recovering $\widehat{y} = z - R\widehat{h}$.

3. PROPOSED APPROACH

3.1. Maximum A Posteriori estimation

Let us now show how the problem can be addressed from a Bayesian perspective. We assume that the characteristics of the primary are appropriately described through a prior statistical model in a basis, e.g. a wavelet one [10]. For instance, if $F \in \mathbb{R}^{N \times N}$ designates the analysis operator and x the associated coefficients, we have [11]

$$y = F^{-1} x \tag{8}$$

where $F^{-1} \in \mathbb{R}^{N \times N}$ is the synthesis operator. In addition, we assume that x is a realization of a random vector, whose probability density function (pdf) is given by

$$(\forall x \in \mathbb{R}^N) \qquad f_X(x) \propto \exp(-\varphi(x)) \tag{9}$$

where φ is the associated potential. For simplicity, φ can be chosen to be separable, which corresponds to an independence assumption on the basis coefficients:

$$\big(\forall x = (x_k)_{1 \le k \le N} \in \mathbb{R}^N\big) \qquad \varphi(x) = \sum_{k=1}^{N} \varphi_k(x_k). \tag{10}$$

where, for every $k \in \{1, \ldots, N\}$, $\varphi_k \colon \mathbb{R} \to]-\infty, +\infty]$. In order to promote the sparsity of the decomposition, a standard choice for this potential is $\varphi_k = \kappa_k |\cdot|$ where $\kappa_k > 0$. On the other hand, to take into account the available information on the unknown filter, especially its regular variations along the time dimension, it can be assumed that h is a realization of a random vector, whose pdf is expressed as

$$(\forall h \in \mathbb{R}^{NP}) \qquad f_H(h) \propto \exp(-\rho(h)), \tag{11}$$

and which is independent of x. By resorting to an estimation of h in the sense of the MAP, the problem can thus be formulated under the following variational form:

$$\underset{h\in\mathbb{R}^{NP}}{\text{minimize}}\ \varphi\big(F(z-Rh)\big)+\rho(h). \tag{12}$$

In this approach, φ represents some data fidelity term taking into account the statistical properties of the basis coefficients and ρ models prior informations that are available on h.

3.2. Prior information about the filters

The filters are here assumed to vary along the time index n. To additionally take into account a prior knowledge on slow filter variations, the following bounded variation constraint can be introduced

$$\big(\forall(n,p)\big) \qquad |h^{(n+1)}(p)-h^{(n)}(p)|\leq\varepsilon_p. \tag{13}$$

This inequality restricts corresponding FIR coefficient variations between estimations at two consecutive time samples. The bound ε_p may depend on the shape of the expected filter. For example, its dependence on the coefficient index p may enable a larger (resp. smaller) difference for filter coefficients taking larger (resp. smaller) values.

The associated closed convex set is defined as

$$C=\Big\{h\in\mathbb{R}^{NP}\mid\forall(n,p)\ \ |h^{(n+1)}(p)-h^{(n)}(p)|\leq\varepsilon_p\Big\}. \tag{14}$$

We subsequently assume that the pdf f_H is compactly supported on C, so yielding the following criterion to be minimized, with $\widetilde{\rho}\colon\mathbb{R}^{NP}\to\,]-\infty,+\infty]$

$$\underset{h\in\mathbb{R}^{NP}}{\text{minimize}}\ \varphi\big(F(z-Rh)\big)+\widetilde{\rho}(h)+\iota_C(h). \tag{15}$$

For computational issues, the convex set C can be expressed as the intersection of two convex subsets C_1 and C_2:

$$C_1=\Big\{h\mid\forall p,\,\forall n\in\Big\{0,\ldots,\Big\lfloor\frac{N}{2}\Big\rfloor-1\Big\}\ \ \Big|h^{(2n+1)}(p)-h^{(2n)}(p)\Big|\leq\varepsilon_p\Big\} \tag{16}$$

$$C_2=\Big\{h\mid\forall p,\,\forall n\in\Big\{1,\ldots,\Big\lfloor\frac{N-1}{2}\Big\rfloor\Big\}\ \ \Big|h^{(2n)}(p)-h^{(2n-1)}(p)\Big|\leq\varepsilon_p\Big\}. \tag{17}$$

Note that in each subset, the involved variables are decoupled. Using these two subsets, Criterion (15) becomes:

$$\underset{h\in\mathbb{R}^{NP}}{\text{minimize}}\ \varphi\big(F(z-Rh)\big)+\widetilde{\rho}(h)+\iota_{C_1}(h)+\iota_{C_2}(h). \tag{18}$$

For tractability, in the following, the functions φ and $\widetilde{\rho}$ are assumed to be convex.

4. PROXIMAL ALGORITHM

To perform the minimization in (18), we employ the PPXA+ method recently developed in [12]. This algorithm constitutes a generalization of the Parallel Proximal Algorithm (PPXA) proposed in [13]. It is also closely related to Augmented Lagrangian methods [14]. This algorithm requires to compute the proximity operators of each of the terms in (18). Some basic facts about proximity operators are recalled next.

4.1. Proximity operators

Let $\mathcal{H}$ be a separable real Hilbert space with scalar product $\langle\cdot\mid\cdot\rangle$ and norm $\|\cdot\|$. $\Gamma_0(\mathcal{H})$ denotes the class of proper lower semi-continuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$. The proximity operator [15] of $\phi\in\Gamma_0(\mathcal{H})$ is defined as

$$\text{prox}_\phi\colon\mathcal{H}\to\mathcal{H}\colon u\mapsto\arg\min_{v\in\mathcal{H}}\frac{1}{2}\|v-u\|^2+\phi(v). \tag{19}$$

Hence, if C is a nonempty closed convex set of $\mathcal{H}$, and ι_C denotes the indicator function of C, i.e., $(\forall u\in\mathcal{H})\ \iota_C(u)=0$ if $u\in C$, $+\infty$ otherwise, then, prox_{ι_C} reduces to the projection Π_C onto C.

This operator possesses numerous properties [11, 16]. Some of them, which are useful to derive explicit forms of proximity operators, are stated below:

- Let $\psi=\phi(\cdot-v)$, where $v\in\mathcal{H}$. Then

$$(\forall u\in\mathcal{H}) \qquad \text{prox}_\psi u=v+\text{prox}_\phi(u-v). \tag{20}$$

- Let $\psi\colon v\mapsto\phi(-v)$. Then

$$(\forall u\in\mathcal{H}) \qquad \text{prox}_\psi u=-\text{prox}_\phi(-u). \tag{21}$$

Furthermore, when dealing with a composition of a linear operator and a convex function, the following property can be used:

Proposition 4.1 [16] *Let $\mathcal{G}$ be a real Hilbert space. Let $\phi\in\Gamma_0(\mathcal{G})$ and let $L\colon\mathcal{G}\to\mathcal{H}$ denote a bounded linear operator. Suppose that $LL^\top=\chi\,\mathrm{I}$, for some $\chi\in\,]0,+\infty[$. Then, $\phi\circ L\in\Gamma_0(\mathcal{H})$ and* $\text{prox}_{\phi\circ L}=\mathrm{I}+\chi^{-1}L^\top(\text{prox}_{\chi\phi}-\mathrm{I})L$.

Hereabove, I denotes the identity operator.

4.2. Choice of the data fidelity and regularization terms and associated proximity operators

Concerning the data fidelity term, it can be observed that it is equal to $h\mapsto\Phi(Rh)$, where $\Phi\triangleq\varphi(F(z-\cdot))$. If F is a decomposition onto an orthonormal basis, by using (20), (21) and Proposition 4.1, it can be easily derived that

$$(\forall x\in\mathbb{R}^N) \qquad \text{prox}_\Phi x=z-F^{-1}\,\text{prox}_\varphi(F(z-x)). \tag{22}$$

As already mentioned in Section 3.1, the separable form (10) can be adopted for φ, where we can set, for every $k\in$

$\{1,\ldots,N\}$, $\varphi_k = \kappa_k|\cdot|^{p_k}$ with $p_k \in [1,+\infty[$ and $\kappa_k \in]0,+\infty[$. Closed form expressions of the considered power functions are indeed available for certain values of the exponents [11]. A similar parametric form has been adopted for the regularization function $\widetilde{\rho}$.

Concerning the constraints modeled by the closed convex sets C_1 and C_2, the proximity operators of the associated indicator functions are given by the projections onto these sets. These projections reduce to projections onto a set of hyperslabs of $\mathbb{R}^2$. More precisely, the projection onto C_1 is calculated as follows: let $h \in \mathbb{R}^{NP}$ and let $g_1 = \Pi_{C_1}(h)$; then for every $p \in \{p',\cdots,p'+P-1\}$ and for every $n \in \{0,\ldots,\lfloor\frac{N}{2}\rfloor - 1\}$,

1. if $|h^{(2n+1)}(p) - h^{(2n)}(p)| \leq \varepsilon_p$, then

$$g_1^{(2n)}(p) = h^{(2n)}(p), \quad g_1^{(2n+1)}(p) = h^{(2n+1)}(p);$$

2. if $h^{(2n+1)}(p) - h^{(2n)}(p) > \varepsilon_p$, then

$$g_1^{(2n)}(p) = \frac{h^{(2n+1)}(p) + h^{(2n)}(p)}{2} - \frac{\varepsilon_p}{2}$$
$$g_1^{(2n+1)}(p) = \frac{h^{(2n+1)}(p) + h^{(2n)}(p)}{2} + \frac{\varepsilon_p}{2};$$

3. if $h^{(2n+1)}(p) - h^{(2n)}(p) < -\varepsilon_p$, then

$$g_1^{(2n)}(p) = \frac{h^{(2n+1)}(p) + h^{(2n)}(p)}{2} + \frac{\varepsilon_p}{2}$$
$$g_1^{(2n+1)}(p) = \frac{h^{(2n+1)}(p) + h^{(2n)}(p)}{2} - \frac{\varepsilon_p}{2}.$$

Similar expressions hold for the projection onto C_2.

4.3. Proposed algorithm

In the considered application, we propose to employ the iterative algorithm in [12] (see Algorithm 1). This algorithm mainly consists of alternately computing proximity operators and projections (see Section 4.2). Additionally, it can be noticed that it requires to compute the inverse of matrix $Q \in \mathbb{R}^{NP\times NP}$ given by $\omega_1 R^\top R + (\omega_2+\omega_3+\omega_4)\,\mathrm{I}$. Since $R^\top R$ has a block diagonal structure, the inversion can be performed in a very efficient manner.

Algorithm 1 PPXA+

Set $(\omega_1,\omega_2,\omega_3,\omega_4) \in]0,+\infty[^4$ and $t_{1,0} \in \mathbb{R}^N, t_{2,0} \in \mathbb{R}^{NP}, t_{3,0} \in \mathbb{R}^{NP}, t_{4,0} \in \mathbb{R}^{NP}$
$Q = \omega_1 R^\top R + (\omega_2+\omega_3+\omega_4)\,\mathrm{I}$
$h_0 = Q^{-1}\Big(\omega_1 R^\top t_{1,0} + \omega_2 t_{2,0} + \omega_3 t_{3,0} + \omega_4 t_{4,0}\Big)$
for $i = 0, 1, \ldots$ **do**
$w_{1,i} = \mathrm{prox}_{\frac{\Phi}{\omega_1}}(t_{1,i})$ and $w_{2,i} = \mathrm{prox}_{\frac{\widetilde{\rho}}{\omega_2}}(t_{2,i})$
$w_{3,i} = \Pi_{C_1}(t_{3,i})$ and $w_{4,i} = \Pi_{C_2}(t_{4,i})$
$c_i = Q^{-1}\Big(\omega_1 R^\top w_{1,i} + \omega_2 w_{2,i} + \omega_3 w_{3,i} + \omega_4 w_{4,i}\Big)$
$t_{1,i+1} = t_{1,i} + \lambda_i\,(R(2c_i - h_i) - w_{1,i})$
$t_{2,i+1} = t_{2,i} + \lambda_i\,(2c_i - h_i - w_{2,i})$
$t_{3,i+1} = t_{3,i} + \lambda_i\,(2c_i - h_i - w_{3,i})$
$t_{4,i+1} = t_{4,i} + \lambda_i\,(2c_i - h_i - w_{4,i})$
$h_{i+1} = h_i + \lambda_i\,(c_i - h_i)$
end for

5. NUMERICAL EXPERIMENTS

In this section, we aim at showing the good performance of the proposed approach. Data $y = (y^{(n)})_{0\leq n<N}$ and $r = (r^{(n)})_{0\leq n<N}$ considered here were generated from actual seismic data primaries and multiples. In the following, we have $N = 2048$. The inner parameters of PPXA+ have been chosen in an empirical manner: $\lambda_k \equiv 1.5$, $\omega_1 = 10000/N, \omega_2 = \omega_1/P, \omega_3 = \omega_4 = 10\,\omega_2$; the algorithm is initialized by randomly generating N positive-valued vectors of size P summing up to one. The algorithm is launched on 10000 iterations and may stop earlier at iteration i if $\|h_{i+1} - h_i\| < 10^{-5}$.

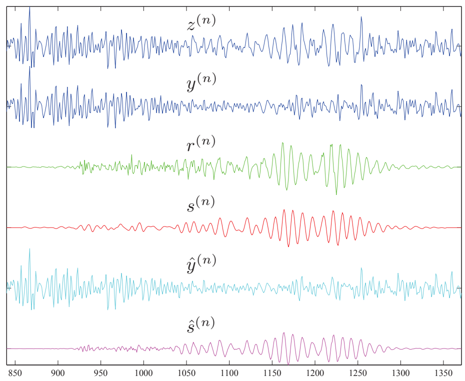

Fig. 2. Cropped version of the results in the ideal case. From up to down: original data z, reference primary signal y, template r, multiples s, estimated signal $\hat{y}$ and multiples $\hat{s}$.

The numerical results below have been obtained with $\varphi_k \equiv |\cdot|$, $\widetilde{\rho} = \mu\|\cdot\|^2$ and $\mu = 0.01$. Symlet wavelets of length 8 over 3 resolution levels are used.

5.1. Ideal case

In a first set of experiments, the observations being generated according to (1), we assume P, p' and $\epsilon = \max_{p'\leq p\leq p'+P-1, 0\leq n<N-1} |\overline{h}^{(n+1)}(p) - \overline{h}^{(n)}(p)|$ to be known, where $\overline{h}$ denotes the "true" impulse response. We subsequently set $\epsilon_p \equiv \epsilon$ for defining the convex constraints on the filters.

Results obtained considering $P = 14$, $p' = -7$ and $\epsilon = 1.46\times 10^{-3}$ are displayed in Fig. 2. In this case, as the filter length is quite large, the multiples are of low amplitude and so, they are difficult to detect. However, the signal is satisfactorily recovered.

5.2. Objective performance evaluation

To emulate actual geophysical configurations, data and templates obtained from actual seismic surveys have been combined with filters of varying length (from four to twenty taps), starting time (p') and shape (from fully symmetric to asymmetric). Due to amplitude and spectrum variations in actual multiples, an absolute statistical analysis is not straightforward. Instead, we focus on relative errors with respect to reference data, before and after multiple removal. Namely, Fig. 3 represents pairs (with identical colors) of ratios of ℓ_1 and ℓ_2 norms for the initial versus final error with respect to y. Three key observations are derived:
• pairs of error metrics (identical colors) behave consistently,
• reported errors stand below the main diagonal (dashed line), indicating generic improvements,
• relative errors exhibit a reduction by a factor of four in average (dotted line).

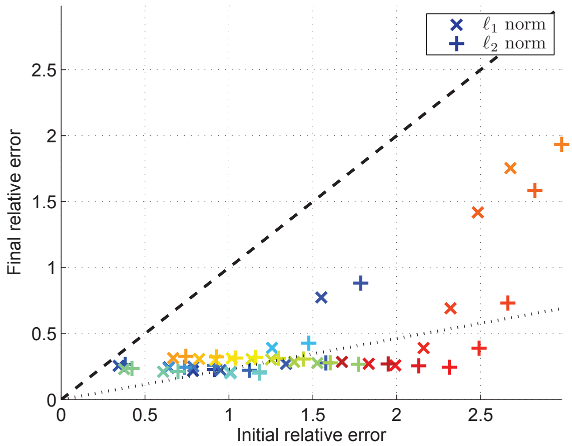

Fig. 3. Initial vs final relative errors.

5.3. Noisy case

In a second time, we generated again observations according to (1) where $P = 10$, $p' = 0$ and $\epsilon = 1.28 \times 10^{-3}$ are still assumed to be known, but the observed data are corrupted with noise. Model (1) becomes then

$$(\forall n \in \{0, \cdots, N-1\}) \qquad z^{(n)} = s^{(n)} + y^{(n)} + b^{(n)} \quad (23)$$

where b is a realization of an additive white Gaussian noise with zero-mean. The noise standard deviation is here chosen equal to $\sigma = 8.35 \times 10^{-2}$ (SNR = 0.95 dB). The results shown in Fig. 4 demonstrate the robustness of the proposed method with respect to noise.

6. CONCLUSION

A new variational framework for multiple removal in seismic data, when a disturbance signal template is available, has been developed. The proposed algorithm, based on recent advances in the theory of proximal operator, allows us to estimate FIR filters that vary along the time dimension. The

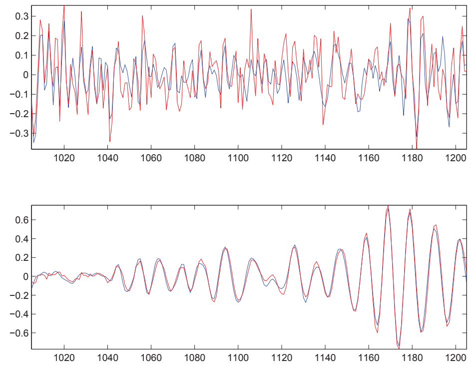

Fig. 4. Cropped version of the results in the noisy case. Top: reference signal y (blue line) and estimated signal $\widehat{y}$ (red line); Bottom: multiples s (blue line) and estimated multiples $\widehat{s}$ (red line).

results provided by this approach appear to be very promising. In our future work, we plan to extend this method to the 2D case, so as to exploit prior information along the sensor dimension.

7. REFERENCES

[1] S. Ventosa, H. Rabeson, and L. Duval, "Coherent noise removal in seismic data with redundant multiscale directional filters," in *Proc. Eur. Sig. and Image Proc. Conference*, Spain, Barcelona, Aug. 29-Sep. 2, 2011, 5 p.

[2] P. Goupillaud, A. Grossmann, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," *Geoexploration*, vol. 23, no. 1, pp. 85–102, 1984.

[3] S. Ventosa, S. Le Roy, I. Huard, A. Pica, H. Rabeson, P. Ricarte, and L. Duval, "Adaptive multiple subtraction with wavelet-based complex unary Wiener filters," *Geophysics*, in press, 2012.

[4] J. F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, vol. 38, no. 5, pp. 826–844, 1973.

[5] I. W. Selesnick, R. Van Slyke, and O. G. Guleryuz, "Pixel recovery via ℓ_1 minimization in the wavelet domain," in *Proc. Int. Conf. Image Process.*, Singapore, Oct. 24-27, 2004, vol. 3, pp. 1819–1822.

[6] R. Essenreiter, M. Karrenbach, and S. Treitel, "Multiple reflection attenuation in seismic data using backpropagation," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 2001–2011, Jul. 1998.

[7] A. Guitton and W. Symes, "Robust inversion of seismic data using the Huber norm," *Geophysics*, vol. 68, no. 4, pp. 1310–1319, 2003.

[8] A. T. Walden and J. W. J. Hosken, "The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution," *Geophys. Prospect.*, vol. 34, no. 7, pp. 1038–1066, 1986.

[9] J.-C. Pesquet, A. Benazza-Benyahia, and C. Chaux, "A SURE approach for digital signal/image deconvolution problems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4616–4632, Dec. 2009.

[10] S. Mallat, *A wavelet tour of signal processing: the sparse way*, Academic Press, San Diego, CA, USA, 3rd edition, 2009.

[11] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, "A variational formulation for frame based inverse problems," *Inverse Probl.*, vol. 23, no. 4, pp. 1495–1518, Aug. 2007.

[12] J.-C. Pesquet and N. Pustelnik, "A parallel inertial proximal optimization method," *Pacific J. Opt.*, vol. 8, no. 2, pp. 273–305, Apr. 2012.

[13] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse Probl.*, vol. 24, no. 6, 27 p., Dec. 2008.

[14] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[15] J. J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.

[16] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer Verlag, 2010.