

SEISMIC MULTIPLE REMOVAL WITH A PRIMAL-DUAL PROXIMAL ALGORITHM

Mai Quyen Pham^{1,3}, Caroline Chaux², Laurent Duval¹, and Jean-Christophe Pesquet³

¹ IFP Energies nouvelles
1 et 4 av. de Bois-Préau
92852 Rueil-Malmaison, France
{mai-quyen.pham, laurent.duval}@ifpen.fr

² Aix-Marseille Univ.
LATP UMR CNRS 7353
39 rue F. Joliot-Curie,
13453 Marseille, France
caroline.chaux@latp.univ-mrs.fr

³ Univ. Paris-Est
LIGM UMR-CNRS 8049
5 bd Descartes,
77454 Marne-la-Vallée, France
pesquet@univ-mlv.fr

ABSTRACT

Both random and structured perturbations affect seismic data. Their removal, to unveil meaningful geophysical information, requires additional priors. Seismic multiples are one form of structured perturbations related to wave-field bouncing. In this paper, we model these undesired signals through a time-varying filtering process accounting for inaccuracies in amplitude, time-shift and average frequency of available templates. We recast the problem of jointly estimating the filters and the signal of interest (primary) in a new convex variational formulation, allowing the incorporation of knowledge about the noise statistics. By making some physically plausible assumptions about the slow time variations of the filters, and by adopting a potential promoting the sparsity of the primary in a wavelet frame, we design a primal-dual algorithm which yields good performance in the provided simulation examples.

Index Terms— Optimization methods, Wavelet transforms, Adaptive filters, Geophysical signal processing, Signal restoration.

1. INTRODUCTION

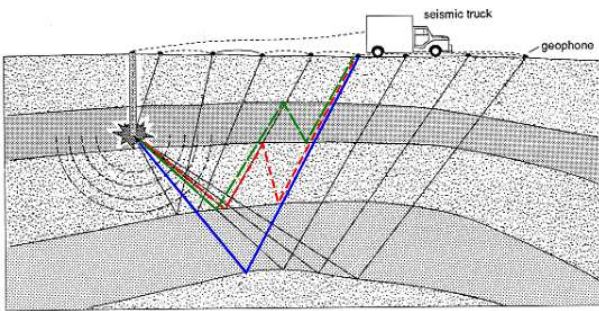


Fig. 1. Principles of seismic wave propagation, with reflections on different layers, and data acquisition. Solid blue: primary; dashed red and green: multiple reflection disturbances.

Geophysical signal processing [1, 2] addresses the extraction of relevant information present in seismic data. In reflec-

tion seismology, seismic waves propagate through the subsurface medium. The portion of seismic wave fields recorded at the surface forms the seismic traces whose reflections at geological interfaces and propagation-related distortions inform about the subsurface structure (see Fig. 1). An idealization would consist in inferring the relative distances and velocity contrasts between layers through an impulsive seismic source signal traveling first downwards, then upwards, toward the seismic sensors. Many types of unpredicted disturbances affect seismic signals. Consequently, geophysics has nurtured several tools central to potent signal processing trends, including robust, ℓ_1 -promoted deconvolution [3], or complex, continuous wavelet transforms [4]. One of the most severe types of interferences, hence still requiring mitigation, are multiple reflections, and correspond to seismic waves bouncing betwixt layers [5]. Such reverberations, from the point of view of geological information interpretation, imitate and even bedim genuine target reflectors, since they possess similar waveform and frequency content. Model-based multiple removal is one of the industry standard techniques. It consists of estimating a realistic template of the multiples, which is subsequently adapted in amplitude, delay and frequency by time-varying matched filtering techniques, for instance in a wavelet or curvelet domain, see [6, 7] and references therein. When highly complicated propagation paths occur (dashed lines in Fig. 1), several multiple templates are devised and adaptively weighted depending on the time and space location of seismic traces. Inaccuracies in template modeling as well as the complexity of the time-varying adaptation combined with additional unmodeled disturbances require additional constraints to obtain geophysically sound solutions. We use prior knowledge on seismic data distribution (sparsity in wavelet frames) and assume that the time-varying filters, adapting each template, possess a finite impulse response (FIR) that smoothly varies in time. We assume that a seismic trace is composed as follows:

$$z^{(n)} = \bar{s}^{(n)} + \bar{y}^{(n)} + b^{(n)}, \quad (1)$$

where $n \in \{0, \dots, N-1\}$ denotes the time index and $z = (z^{(n)})_{0 \leq n < N}$ corresponds to the observed data (seismic trace) combining: the primary $\bar{y} = (\bar{y}^{(n)})_{0 \leq n < N}$ (signal

of interest, unknown) depicted in solid blue (Fig. 1), the multiples $(\bar{s}^{(n)})_{0 \leq n < N}$ (sum of undesired reflected signals) depicted in dashed red and green (Fig. 1) and the noise $b = (b^{(n)})_{0 \leq n < N}$ which is assumed to be additive zero-mean white Gaussian. The paper improves upon [8] by: taking into account several multiple templates, incorporating additional random noise into the generic model, and introducing the nuclear norm as a potential multiple selection objective criterion. Section 2 formulates the generic variational form of the problem. Section 3 recalls the primal-dual proximal formulation, followed in Section 4 by precisions concerning the chosen optimization criterion. The performance of the proposed method is assessed in Section 5.

2. VARIATIONAL FORMULATION

2.1. Model description

We assume that J templates $(r_j^{(n)})_{0 \leq n < N, 0 \leq j < J}$ for the disturbance signal are available, which are related to $(\bar{s}^{(n)})_{0 \leq n < N}$ through an FIR, possibly non-causal, linear model

$$\bar{s}^{(n)} = \sum_{j=0}^{J-1} \sum_{p=p'}^{p'+P_j-1} \bar{h}_j^{(n)}(p) r_j^{(n-p)} \quad (2)$$

where $\bar{h}_j^{(n)}$ is an unknown impulse response (P_j tap coefficients) corresponding to template j and time n and where $p' \in \{-P_j + 1, \dots, 0\}$ ($p' = 0$ corresponds to the causal case). It must be emphasized that the dependence w.r.t. the time index n of the impulse responses implies that the filtering process is not time invariant, although it can be assumed slowly varying in practice. Templates are generated with standard geophysical modeling based on primaries. The adaptive FIR assumption is common practice. The observation that adapted filters are ill-behaved, due to the band-pass nature of seismic data is well known, although rarely documented, motivating the need for filter coefficient control.

Eq. (2) can be expressed more concisely as

$$\bar{s} = \sum_{j=0}^{J-1} R_j \bar{h}_j \quad (3)$$

by appropriately defining vectors \bar{s} , $(\bar{h}_j)_{0 \leq j < J}$ and matrices $(R_j)_{0 \leq j < J}$. More precisely,

$$\bar{s} = [\bar{s}^{(0)} \quad \dots \quad \bar{s}^{(N-1)}]^\top, \quad (4)$$

$$\bar{h}_j = [\bar{h}_j^{(0)}(p') \quad \dots \quad \bar{h}_j^{(0)}(p' + P_j - 1) \quad \dots \quad \bar{h}_j^{(N-1)}(p') \quad \dots \quad \bar{h}_j^{(N-1)}(p' + P_j - 1)]^\top. \quad (5)$$

Each matrix R_j is defined as a block diagonal matrix of size

$N \times NP_j$:

$$R_j = \begin{bmatrix} R_j^{(0)} & 0 & \dots & 0 \\ 0 & R_j^{(1)} & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & R_j^{(N-1)} \end{bmatrix} \quad (6)$$

where

$$\left[(R_j^{(0)})^\top (R_j^{(1)})^\top \dots (R_j^{(N-1)})^\top \right]^\top = \begin{bmatrix} r_j^{(-p')} & \dots & r_j^{(0)} & 0 & \dots & 0 \\ r_j^{(-p'+1)} & \dots & r_j^{(0)} & r_j^{(0)} & 0 & \dots & 0 \\ \vdots & & & & & & \\ r_j^{(N-1)} & r_j^{(N-2)} & \dots & \dots & \dots & r_j^{(N-P_j)} \\ 0 & r_j^{(N-1)} & \dots & \dots & \dots & r_j^{(N-P_j+1)} \\ \vdots & & & & & \\ 0 & \dots & 0 & r_j^{(N-1)} & \dots & r_j^{(N-P_j-p')} \end{bmatrix}.$$

For more conciseness, one can write $\bar{s} = \mathbf{R}\bar{\mathbf{h}}$ by defining $\mathbf{R} = [R_0 \dots R_{J-1}] \in \mathbb{R}^{N \times Q}$ where $Q = NP$ with $P = \sum_{j=0}^{J-1} P_j$ and $\bar{\mathbf{h}} = [\bar{h}_0^\top \dots \bar{h}_{J-1}^\top]^\top \in \mathbb{R}^Q$.

2.2. Problem formulation

We define the following objective function, for every $y \in \mathbb{R}^N$ and $\mathbf{h} \in \mathbb{R}^Q$,

$$\theta(y, \mathbf{h}) = \psi(z - y - \mathbf{R}\mathbf{h}) + \varphi(Fy) + \tilde{\rho}(\mathbf{L}\mathbf{h}) + \sum_{m=1}^M \iota_{C_m}(\mathbf{h}), \quad (7)$$

where $\psi : \mathbb{R}^N \rightarrow]-\infty, +\infty]$, $\varphi : \mathbb{R}^K \rightarrow]-\infty, +\infty]$, $\tilde{\rho} : \mathcal{G} \rightarrow]-\infty, +\infty]$, \mathcal{G} is a real Hilbert space that we will specify later on, and, for every $m \in \{1, \dots, M\}$, ι_{C_m} designates the indicator function¹ of a closed convex set C_m (M constraints are considered here). $F \in \mathbb{R}^{K \times N}$ designates an analysis frame operator [9] and \mathbf{L} is a nonzero bounded linear operator from \mathbb{R}^Q to \mathcal{G} . For tractability, in the following, functions ψ , φ and $\tilde{\rho}$ are assumed to be convex. Our approach consists of solving the following optimization problem:

$$\underset{y \in \mathbb{R}^N, \mathbf{h} \in \mathbb{R}^Q}{\text{minimize}} \theta(y, \mathbf{h}). \quad (8)$$

The minimization of this criterion can be interpreted as an estimation of $(\bar{y}, \bar{\mathbf{h}})$ in the sense of the Maximum A Posteriori (MAP), where the first term represents a data fidelity term accounting for the noise, φ represents a regularization term taking into account the statistical properties of the frame coefficients of the primary, and $\tilde{\rho}$ models prior, available information on \mathbf{h} . To further account for smooth variations of

¹The indicator function ι_C of a set C is such that $\iota_C(u) = 0$ if $u \in C$, $+\infty$ otherwise.

filters \mathbf{h} along the time dimension, hard constraints on the filters can be added through the convex sets $(C_m)_{1 \leq m \leq M}$.

It can be observed that the data fidelity term defines a function $\Psi \triangleq \psi(z - \Omega_1 \cdot)$ where $\Omega_1 = [\mathbf{I} \ \mathbf{R}]$. Concerning the regularization terms, we will introduce a second function namely $\Phi(\Omega_2 \cdot)$, where $\Phi: \mathbb{R}^K \times \mathcal{G} \rightarrow]-\infty, +\infty]: (x, u) \mapsto \varphi(x) + \tilde{\rho}(u)$ and Ω_2 is the block diagonal linear operator $\begin{bmatrix} y \\ \mathbf{h} \end{bmatrix} \mapsto (Fy, \mathbf{Lh})$. With these notations, (7) can be reexpressed as

$$\theta(y, \mathbf{h}) = \Psi \left(\begin{bmatrix} y \\ \mathbf{h} \end{bmatrix} \right) + \Phi(\Omega_2 \begin{bmatrix} y \\ \mathbf{h} \end{bmatrix}) + \sum_{m=1}^M \iota_{C_m}(\mathbf{h}). \quad (9)$$

3. PRIMAL-DUAL PROXIMAL ALGORITHM

To perform the minimization in (8), we propose to employ the Monotone+Lipschitz Forward-Backward-Forward (M+LFBF) algorithm recently proposed in [10]. This algorithm belongs to the class of primal-dual algorithms and avoid some matrix inversions necessary in many other approaches [11, 12]. It requires to compute the proximity operators of each of the terms in (8) (note that when one term has a Lipschitz continuous gradient, the gradient can be computed instead of the proximity operator). Some relevant information about proximity operators are recalled next.

3.1. Proximity operators

Let \mathcal{H} be a separable real Hilbert space with 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 [13] of $\phi \in \Gamma_0(\mathcal{H})$ is defined as

$$\text{prox}_\phi: \mathcal{H} \rightarrow \mathcal{H}: u \mapsto \arg \min_{v \in \mathcal{H}} \frac{1}{2} \|v - u\|^2 + \phi(v). \quad (10)$$

If C is a nonempty closed convex set of \mathcal{H} , then prox_{ι_C} reduces to the projection Π_C onto C . Note that this operator possesses numerous mathematical properties [14, 15].

3.2. M+LFBF algorithm

For many standard data fidelity terms, $\psi \in \Gamma_0(\mathbb{R}^N)$ is differentiable with a μ -Lipschitz continuous gradient for some $\mu \in]0, +\infty[$. In this case,

$$\nabla \Psi = -\Omega_1^\top \nabla \psi(z - \Omega_1 \cdot). \quad (11)$$

In addition, it can be assumed that Φ is a possibly nonsmooth function belonging to $\Gamma_0(\mathbb{R}^K \times \mathcal{G})$.

The iterative approach allowing us to solve (8) is described in Algorithm 1 where \mathbf{L}^* denotes the adjoint operator of \mathbf{L} . The use of this algorithm requires to be able to compute

the norm of each linear operator involved in the criterion or at least an upper bound of it. Here, we can derive:

$$\|\Omega_1\| = \|\mathbf{I} \ \mathbf{R}\| \leq \sqrt{1 + \|\mathbf{R}_0\|^2 + \dots + \|\mathbf{R}_{J-1}\|^2} \quad (12)$$

where $\|\mathbf{R}_j\| = \max_{n \in \{0, \dots, N-1\}} \|\mathbf{R}_j^{(n)}\|$ for every $j \in \{0, \dots, J-1\}$ and $\|\Omega_2\| = \max(\|\mathbf{F}\|, \|\mathbf{L}\|)$. At each iteration i , the stepsize $\gamma^{[i]}$ must be chosen so as to satisfy the following rule: let $\beta = \mu + \sqrt{\|\Omega_2\|^2 + M}$ and $\epsilon \in]0, \frac{1}{\beta+1}[$, then $\gamma^{[i]} \in [\epsilon, \frac{1-\epsilon}{\beta}]$. It must be emphasized that the choice of the stepsize is crucial for the convergence speed of the algorithm. If the norms of the matrices are too high, then the algorithm will converge slowly since the stepsize will be small. In this case, a preconditioned version of the algorithm can be employed [16].

Algorithm 1 M+LFBF

Set $\gamma^{[i]} \in [\epsilon, \frac{1-\epsilon}{\beta}]$. Set $\begin{bmatrix} y^{[0]} \\ \mathbf{h}^{[0]} \end{bmatrix} \in \mathbb{R}^{N+Q}$, $v_1^{[0]} \in \mathbb{R}^K$, $u_1^{[0]} \in$

\mathcal{G} , $(u_{2,m}^{[0]})_{1 \leq m \leq M} \in (\mathbb{R}^Q)^M$

for $i = 0, 1, \dots$ **do**

Gradient computation

$$\begin{bmatrix} s_1^{[i]} \\ t_1^{[i]} \end{bmatrix} = \begin{bmatrix} y^{[i]} \\ \mathbf{h}^{[i]} \end{bmatrix} - \gamma^{[i]} \left(\nabla \Psi \left(\begin{bmatrix} y^{[i]} \\ \mathbf{h}^{[i]} \end{bmatrix} \right) + \begin{bmatrix} F^\top v_1^{[i]} \\ \mathbf{L}^* u_1^{[i]} + \sum_{m=1}^M u_{2,m}^{[i]} \end{bmatrix} \right)$$

Proximity operator computation

$$s_2^{[i]} = v_1^{[i]} + \gamma^{[i]} F y^{[i]}$$

$$w_1^{[i]} = s_2^{[i]} - \gamma^{[i]} \text{prox}_{(\gamma^{[i]})^{-1} \varphi}((\gamma^{[i]})^{-1} s_2^{[i]})$$

$$t_{2,1}^{[i]} = u_1^{[i]} + \gamma^{[i]} \mathbf{L} \mathbf{h}^{[i]}$$

$$(t_{2,2,m}^{[i]})_{1 \leq m \leq M} = (u_{2,m}^{[i]} + \gamma^{[i]} \mathbf{h}^{[i]})_{1 \leq m \leq M}$$

$$w_{2,1}^{[i]} = t_{2,1}^{[i]} - \gamma^{[i]} \text{prox}_{(\gamma^{[i]})^{-1} \tilde{\rho}}((\gamma^{[i]})^{-1} t_{2,1}^{[i]})$$

$$(w_{2,2,m}^{[i]})_{1 \leq m \leq M} = (t_{2,2,m}^{[i]} - \gamma^{[i]} \Pi_{C_m}(\frac{t_{2,2,m}^{[i]}}{\gamma^{[i]}}))_{1 \leq m \leq M}$$

Averaging

$$q_{1,1}^{[i]} = w_1^{[i]} + \gamma^{[i]} F s_1^{[i]}$$

$$v_1^{[i+1]} = v_1^{[i]} - s_2^{[i]} + q_{1,1}^{[i]}$$

$$q_{2,1}^{[i]} = w_{2,1}^{[i]} + \gamma^{[i]} \mathbf{L} t_1^{[i]}$$

$$(q_{2,2,m}^{[i]})_{1 \leq m \leq M} = (w_{2,2,m}^{[i]} + \gamma^{[i]} t_1^{[i]})_{1 \leq m \leq M}$$

$$u_1^{[i+1]} = u_1^{[i]} - t_{2,1}^{[i]} + q_{2,1}^{[i]}$$

$$(u_{2,m}^{[i+1]})_{1 \leq m \leq M} = (u_{2,m}^{[i]} - t_{2,2,m}^{[i]} + q_{2,2,m}^{[i]})_{1 \leq m \leq M}$$

Update

$$\begin{bmatrix} y^{[i+1]} \\ \mathbf{h}^{[i+1]} \end{bmatrix} = \begin{bmatrix} y^{[i]} \\ \mathbf{h}^{[i]} \end{bmatrix} - \gamma^{[i]} \left(\nabla \Psi \left(\begin{bmatrix} s_1^{[i]} \\ t_1^{[i]} \end{bmatrix} \right) + \begin{bmatrix} F^\top w_1^{[i]} \\ \mathbf{L}^* w_{2,1}^{[i]} + \sum_{m=1}^M w_{2,2,m}^{[i]} \end{bmatrix} \right)$$

end for

4. A DEEPER LOOK AT THE CRITERION AND PROXIMITY OPERATORS

The noise is usually assumed to be additive, zero-mean, white and Gaussian. In this case, as ψ accounts for the noise statis-

tics, it can be chosen equal to $\psi = \frac{1}{2\sigma^2} \|\cdot\|^2$. Then, the Lipschitz constant μ is equal to $\|\Omega_1\|^2/\sigma^2$.

Concerning φ , a separable form can be adopted, where we can set, for every $k \in \{1, \dots, K\}$, $\varphi_k = \kappa_{k,1}|\cdot| + \kappa_{k,2}|\cdot|^{p_k}$ with $p_k \in]1, +\infty[$ and $(\kappa_{k,1}, \kappa_{k,2}) \in [0, +\infty[^2$. Closed form expressions of the considered power functions are indeed available for typical values of the exponent [14].

For simplicity, a similar parametric form can be adopted for the regularization function $\tilde{\rho}$. However, a more general framework can be followed by appropriately defining the linear operator \mathbf{L} . For our purposes, the regularization applied on \mathbf{h} can be chosen separable with respect to multiplicity indices j and thus, we get:

$$\tilde{\rho}(\mathbf{L}\mathbf{h}) = \sum_{j=0}^{J-1} \tilde{\rho}_j(L_j \text{Diag}(h_j)) \quad (13)$$

where, for every $j \in \{0, \dots, J-1\}$, $L_j \in \mathbb{R}^{N' \times NP_j}$ and $\tilde{\rho}_j \in \Gamma_0(\mathbb{R}^{N' \times NP_j})$. Of particular interest is the case when $\tilde{\rho}_j$ is equal to the nuclear norm $\|\cdot\|_*$, up to a multiplicative positive constant. If $N' = NP_j$ and $L_j = \mathbf{I}$, we retrieve the classical ℓ_1 norm penalty on h_j . However, following the work in [17] on sparse regression, a better choice may be $N' = N$ and $L_j = R_j$. The proximity operator of the nuclear norm can be computed easily by thresholding the singular values of the matrix in its argument (see [18]).

Concerning the last term in (7), we propose to set $M = 2$ and to employ

$$C_1 = \left\{ \mathbf{h} \in \mathbb{R}^Q \mid \forall(j, p), \forall n \in \left\{ 0, \dots, \left\lfloor \frac{N}{2} \right\rfloor - 1 \right\}, \right. \\ \left. |h_j^{(2n+1)}(p) - h_j^{(2n)}(p)| \leq \varepsilon_{j,p} \right\} \quad (14)$$

$$C_2 = \left\{ \mathbf{h} \in \mathbb{R}^Q \mid \forall(j, p), \forall n \in \left\{ 1, \dots, \left\lfloor \frac{N-1}{2} \right\rfloor \right\}, \right. \\ \left. |h_j^{(2n)}(p) - h_j^{(2n-1)}(p)| \leq \varepsilon_{j,p} \right\}. \quad (15)$$

These constraints prevent strong variations of corresponding coefficients of the impulse response, estimated at two consecutive time samples. The bounds $\varepsilon_{j,p} \in [0, +\infty[$ 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. Note that in each subset, the involved variables are decoupled. The proximity operators of the associated indicator functions are given by the projections onto these sets, which reduces to projections onto a set of hyper-slabs of \mathbb{R}^2 .

5. SIMULATIONS

To demonstrate the performance of the proposed approach, we process seismic data represented in Fig. 2. We use a real-

istic seismic signal y and templates r_1 and r_2 ($J = 2$), and we synthetically generated the observations according to model (1). For these data, $P_1 = 10$, $P_2 = 14$ and $\sigma = 0.08$.

The criterion to be minimized is defined by (7) where ψ and $(\varphi_k)_{1 \leq k \leq K}$ have been chosen as explained in Section 4. F is a wavelet frame and, for every $k \in \{1, \dots, K\}$, $p_k \equiv 2$ and subband-adaptive values of $\kappa_{k,1}$ and $\kappa_{k,2}$ are set. The constraints C_1 and C_2 are chosen according to (14) and (15), where $\varepsilon_{1,p} = 0.1$ and $\varepsilon_{2,p} = 0.07$ for every p .

A first restoration result is displayed in Fig. 2 when $\tilde{\rho}$ is the nuclear norm. It can be observed that the multiples s are very well estimated. The primary y is faithfully retrieved in the first half part where its amplitude takes significant values. It is only partly unveiled on parts where the multiple energy is higher. Note that actual geophysical signal processing proceeds in an iterative manner, allowing estimated multiple subtraction from the noisy data, instead of direct primary estimation as performed here, or refinements on newly, even partly detected primary zones. We thus consider the method of sufficient performance and potential at the present stage.

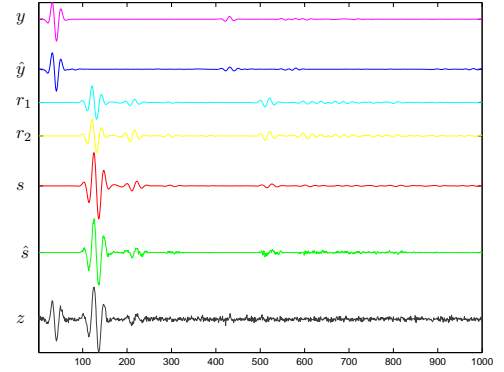


Fig. 2. Considered seismic signals, from top to bottom: original (unknown) signal y , estimated \hat{y} , first template r_1 , second template r_2 , multiple (unknown) s , estimated \hat{s} , and observed signal z ($\sigma = 0.08$).

Table 1 shows the signal-to-noise ratios (SNR, averaged over 100 noise realisations) obtained in the estimation of y and s . Simulations have been run for three a priori $\tilde{\rho}$ and for 4 noise levels (σ).

$\sigma \setminus \tilde{\rho}$	SNR _{y}			SNR _{s}		
	ℓ_1	ℓ_2	$\ \cdot\ _*$	ℓ_1	ℓ_2	$\ \cdot\ _*$
0.01	20.90	21.23	23.57	24.36	24.68	26.74
0.02	20.89	21.16	23.51	22.53	23.02	23.76
0.04	19.00	19.90	20.67	20.15	20.14	19.84
0.08	17.55	16.81	17.34	16.96	16.56	15.96

Table 1. SNR for the estimations of y and s considering different a priori functions $\tilde{\rho}$ and 4 noise levels. Symlet wavelets (undecimated) of length 8 over 4 resolution levels are used.

6. REFERENCES

- [1] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*, Prentice Hall, Englewood Cliffs, NJ., USA, 1980.
- [2] A. K. Takahata, E. Z. Nadalin, R. Ferrari, L. T. Duarte, R. Suyama, R. R. Lopes, J. M. T. Romano, and M. Tygel, “Unsupervised processing of geophysical signals: A review of some key aspects of blind deconvolution and blind source separation,” *IEEE Signal Process. Mag.*, vol. 29, no. 4, pp. 27–35, Jul. 2012.
- [3] J. F. Claerbout and F. Muir, “Robust modeling with erratic data,” *Geophysics*, vol. 38, no. 5, pp. 826–844, Oct. 1973.
- [4] J. Morlet, “Seismic tomorrow, interferometry and quantum mechanics,” in *Soc. Expl. Geophys. Annual International Meeting*, Denver, CO, USA, Oct. 1975.
- [5] D. J. Verschuur, *Seismic multiple removal techniques: past, present and future*, EAGE Publications, 2006.
- [6] M. O. Frijlink, R. Shahidi, and F. J. Herrmann, and R. G. van Borselen, “A Comparison of Standard Adaptive Subtraction and Primary-multiple Separation in the Curvelet Domain,” in *Proc. EAGE*, Barcelona, Spain, Jun. 14-17, 2010.
- [7] 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*, vol. 77, no. 6, pp. V183–V192, Nov.-Dec. 2012.
- [8] D. Gragnaniello, C. Chaux, J.-C. Pesquet, and L. Duval, “A convex variational approach for multiple removal in seismic data,” in *Proc. Eur. Sig. Image Proc. Conf.*, Bucharest, Romania, Aug. 27-31, 2012, pp. 215–219.
- [9] L. Jacques, L. Duval, C. Chaux, and G. Peyré, “A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity,” *Signal Process.*, vol. 91, no. 12, pp. 2699–2730, Dec. 2011.
- [10] P. L. Combettes and J.-C. Pesquet, “Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators,” *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, Jun. 2012.
- [11] 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.
- [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] J.-J. Moreau, “Proximité et dualité dans un espace hilbertien,” *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.
- [14] 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.
- [15] 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.
- [16] A. Repetti, É. Chouzenoux, and J.-C. Pesquet, “A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise,” in *Proc. Eur. Sig. Image Proc. Conf.*, Bucharest, Romania, Aug. 27-31, 2012, pp. 1553–1557.
- [17] É. Grave, G. Obozinski, and F. Bach, “Trace Lasso: a trace norm regularization for correlated designs,” in *Proc. Ann. Conf. Neur. Inform. Proc. Syst.*, Dec. 12-14, 2011, pp. 2187–2195.
- [18] J.-F. Cai, E. J. Candès, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.