MONOTONE OPERATOR SPLITTING FOR OPTIMIZATION PROBLEMS IN SPARSE RECOVERY

M.J. Fadili

GREYC CNRS-ENSICAEN-Université de Caen
14050 Caen France

J.-L. Starck

DAPNIA/SEDI-SAP CEA-Saclay
91191 Gif-sur-Yvette France

ABSTRACT

This work focuses on several optimization problems involved in recovery of sparse solutions of linear inverse problems. Such problems appear in many fields including image and signal processing, and have attracted even more interest since the emergence of the compressed sensing (CS) theory. In this paper, we formalize many of these optimization problems within a unified framework of convex optimization theory, and invoke tools from convex analysis and maximal monotone operator splitting. We characterize all these optimization problems, and to solve them, we propose fast iterative convergent algorithms using forward-backward and/or Peaceman/Douglas-Rachford splitting iterations. With non-differentiable sparsity-promoting penalties, the proposed algorithms are essentially based on iterative shrinkage. This makes them very competitive for large-scale problems. We also report some experiments on image reconstruction in CS to demonstrate the applicability of the proposed framework.

Index Terms— Convex analysis, Non-smooth optimization, Monotone operator splitting, Sparse recovery.

1. INTRODUCTION

The complex structures of natural signals and images require tools in order to make use of their intricate redundancies. To capture this complexity, we have witnessed a flurry of research activities where researchers spanning a wide range of viewpoints have advocated the use of sparsity and overcomplete signal/image representations. It has turned out that exploiting sparsity and overcompleteness offers striking benefits in a wide range of signal and image processing areas. The development of sparse representations in inverse problems is facing major challenges both on theoretical and practical sides. These problems include the design of fast optimization algorithms to handle real world data and application to many ill-posed problems. This work focuses on several optimization problems involved in linear inverse problems where the solution $s = \Phi x, \Phi : \mathcal{H} \to \mathcal{K}$, is assumed to be sparsely represented in an overcomplete dictionary Φ of waveforms $(\varphi_i)_{i \in \mathcal{I}}$. Φ is a generally a frame of $\mathcal{H}$. Our goal is the generic minimization of functions of the form $f = f_1 + f_2$

$$(\text{P}) \quad \min_{x \in \mathcal{H}} f_1(x) + f_2(x) \,, \qquad (1)$$

where f_1, f_2 are closed convex functions which are not necessarily differentiable, are not infinite everywhere, and their domains have non-empty intersection. f_1 is typically a convex sparsity-promoting penalty and f_2 is a function measuring the consistency to data. Problem (P) covers the case of the ℓ_1-norm decoder known also as Basis Pursuit [1, 2, 3] when $f_1(x) = \|x\|_1$, and $f_2(x)$ is the indicator function of the affine subspace $\{x \in \mathcal{H} | s = \mathrm{F}x\}$, $\mathrm{F} = \mathrm{H}\Phi$ is a linear operator which is typically a composition of a linear measurement or degradation operator H and the dictionary Φ. When the observation s is contaminated by noise, the equality constraint must be relaxed to a noise-aware variant. Problem (P) becomes typically Basis Pursuit DeNoising (BPDN) in its augmented Lagrangian form when $f_1(x) = \lambda \|x\|_1$ and $f_2(x) = \frac{1}{2} \|s - \mathrm{F}x\|^2$ [1], and is also known as Lasso [4]. This augmented Lagrangian form has received a considerable attention, see e.g. [5, 6, 7]. Only a few algorithms have been proposed to solve the more difficult to solve constrained form where f_2 is the indicator function of the closed convex set $\{x \in \mathcal{H} | \|s - \mathrm{F}x\| \leqslant \sigma\}$; e.g. [8, 9, 10]. The Dantzig selector [11] is also a special instance of (P) when f_2 is the indicator function of $\{x \in \mathcal{H} | \|\mathrm{F}^*(s - \mathrm{F}x)\|_\infty \leqslant \tau\}$.

In this paper, we formalize all these optimization problems within a unified framework of convex optimization theory, and invoke tools from convex analysis (e.g. duality, proximity operators) and maximal monotone operator splitting. We characterize all these optimization problems, and to solve them, we propose fast iterative and convergent algorithms combining forward-backward and/or Peaceman/Douglas-Rachford splitting iterations. With non-differentiable sparsity-promoting penalties, the proposed algorithms amount to iterative shrinkage. Furthermore, the computational burden of these algorithms is essentially invested in one application at each iteration of fast implicit operators associated to the involved dictionary Φ and the linear measurement operator H and their respective adjoints. This makes them very competitive for large-scale problems such as those rising in image processing.

2. NOTATION AND TERMINOLOGY

Let $\mathcal{H}$ a Hilbert space (here a real vector space) equipped with the inner product $\langle ., . \rangle$ and associated norm $\|.\|$. Let I be the identity operator on $\mathcal{H}$. The operator (induced) norm of A is denoted $\|\|A\|\| = \sup_{x \in \mathcal{H}} \frac{\|Ax\|}{\|x\|}$. Let $\|.\|_p, p \geq 1$ be the ℓ_p-norm with the usual adaptation for the case $p = +\infty$. Denote B_p^σ the closed ℓ_p-ball of radius $\sigma > 0$. Throughout the paper, we denote by $\rightharpoonup$ weak convergence and $\rightarrow$ strong convergence in $\mathcal{H}$.

A real-valued function $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ is coercive, if $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$. The domain of f is defined by $\text{dom} f = \{x \in \mathcal{H} : f(x) < +\infty\}$ and f is proper if $\text{dom} f \neq \emptyset$. We say that a real-valued function f is lower semi-continuous (lsc) if $\liminf_{x \rightarrow x_0} f(x) \geq f(x_0)$. $\Gamma_0(\mathcal{H})$ is the class of all proper lsc convex functions from $\mathcal{H}$ to $(-\infty, +\infty]$.

We note $\iota_\mathcal{C}$ the indicator of the convex set $\mathcal{C}$: $\iota_\mathcal{C}(x) = 0$ if $x \in \mathcal{C}$, and $+\infty$ otherwise. The interior of $\mathcal{C}$ will be denoted $\text{int} \mathcal{C}$ and its relative interior $\text{ri} \mathcal{C}$; see e.g. [12].

We recall that F is the frame synthesis operator associated to a frame of $\mathcal{H}$ if its adjoint F^* (analysis operator) satisfies the generalized Parseval relation with bounds c_1 and c_2

$$c_1 \|x\|^2 \leqslant \|\text{F}^* x\|^2 \leqslant c_2 \|x\|^2, \quad 0 < c_1 \leqslant c_2 < +\infty .$$

The frame is tight when $c_1 = c_2 = c$ and $\text{FF}^* = c\text{I}$.

3. PROXIMAL CALCULUS

We first define the notion of a proximity operator, which was introduced in [13] as a generalization of convex projection operator.

Definition 1 (Proximity operator [13]) *Let $f \in \Gamma_0(\mathcal{H})$. Then, for every $x \in \mathcal{H}$, the function $z \mapsto \frac{1}{2}\|x - z\|^2 + f(z)$ achieves its infimum at a unique point denoted by $\text{prox}_f x$. The uniquely-valued operator $\text{prox}_f = (\text{I} + \partial f)^{-1} : \mathcal{H} \rightarrow \mathcal{H}$ thus defined is the proximity operator of f. It will be convenient to introduce the reflection operator $\text{rprox}_f = 2\text{prox}_f -\text{I}$.*

The following lemma provides an important result on the proximity operator of the pre-composition of $f \in \Gamma_0(\mathcal{H})$ with a bounded affine operator $\text{A} := \text{F} \cdot -y$. It will be at the heart of many algorithms in the rest of the paper.

Lemma 2 *Let F be a bounded linear operator such that $\text{int}(\text{dom}(f) \cap \text{Im}(\text{A})) \neq \emptyset$. Then $f \circ \text{A} \in \Gamma_0(\mathcal{H})$ and*

(i) F is a tight frame. Then

$$\text{prox}_{f \circ \text{A}}(x) = x + c^{-1}\text{F}^* \left(\text{prox}_{cf} -\text{I}\right)(\text{F}x - y) . \quad (2)$$

(ii) F is a general frame. Let $\mu_t \in (0, 2/c_2)$. Define

$$u^{(t+1)} = \mu_t \left(\text{I} - \text{prox}_{\mu_t^{-1} f}\right) \circ \left(\mu_t^{-1} u^{(t)} + \text{A}p^{(t)}\right) , \quad (3)$$

$$p^{(t+1)} = x - \text{F}^* u^{(t+1)} . \quad (4)$$

Then $u^{(t)} \rightarrow \bar{u}$ and $p^{(t)} \rightarrow \text{prox}_{f \circ \text{A}} = x - \text{F}^* \bar{u}$. More precisely, both $u^{(t)}$ and $p^{(t)}$ converge linearly and the best convergence rate $O\left(\left(\frac{c_2 - c_1}{c_2 + c_1}\right)^t\right)$ is attained for $\mu_t \equiv 2/(c_1 + c_2)$.

(iii) If $c_1 = 0$, apply (3) with $\mu_t \in (0, 2/c_2)$. Then $u^{(t)} \rightharpoonup \bar{u}$ and $p^{(t)} \rightarrow \text{prox}_{f \circ \text{A}} = x - \text{F}^* \bar{u}$ at the rate $O(1/t)$.

The proof is based on Fenchel-Rockafellar duality [12] and is detailed in a forthcoming long version of the paper. In the special case of tight frames, we recover the same result as in [14]. Note that $f \circ \text{A} \in \Gamma_0(\mathcal{H})$ is always verified when F corresponds to a frame. The recursion (3) is basically a forward-backward splitting [7] applied to the dual problem. Other solvers can be used instead such as Nesterov's scheme [15].

4. SPARSE RECOVERY INVERSE PROBLEMS

4.1. Splitting schemes

Recall that our goal is the minimization of $f = f_1 + f_2$ in (1) for $f_1, f_2 \in \Gamma_0(\mathcal{H})$ such that their domain has a non-empty intersection. Let T the set of solutions of problem (P), $T = \{x \in \mathcal{H} | x \in (\partial f)^{-1}(0)\}$. When $T \neq \emptyset$, the proximal-type recursion constructed as $x^{(t+1)} = \text{prox}_{\beta f}(x^{(t)}), \beta > 0$ is a fundamental algorithm for solving (1), and $x^{(t)} \rightharpoonup \bar{x} \in T$. The main difficulty with this iteration is that $\text{prox}_{\beta f}$ may be hard to compute in general, depending on the function f. This is for instance the case in most inverse problems arising in image and signal processing. Splitting methods for problem (P) are algorithms that do not attempt to evaluate the proximity mapping $\text{prox}_{\beta f}$ of the combined function f, but instead perform a sequence of calculations involving separately the individual proximity (resolvent) operators $\text{prox}_{\beta f_1}$ and $\text{prox}_{\beta f_2}$.

The Douglas/Peaceman-Rachford (DR) family is the most general pre-existing class of monotone operator splitting methods. Given a fixed scalar $\beta > 0$ and a sequence $\alpha_t \in (0, 2)$, this class of methods may be expressed via the recursion

$$x^{(t+1)} = \left[\left(1 - \frac{\alpha_t}{2}\right)\text{I} + \frac{\alpha_t}{2}\text{rprox}_{\beta f_1} \circ \text{rprox}_{\beta f_2}\right] x^{(t)} . \quad (5)$$

The DR splitting has been brought to light in the image and signal processing literature very recently by [14]. Its most general form is due to [16] where its convergence is established.

Theorem 3 ([16]) *Suppose that $T \neq \emptyset$. Let $\beta \in (0, +\infty)$, $(\alpha_t)_{t \in \mathbb{N}}$ be a sequence in $(0, 2)$ such that $\sum_{t \in \mathbb{N}} \alpha_t(2 - \alpha_t) = +\infty$, and let $(a_t)_{t \in \mathbb{N}}$ and $(b_t)_{t \in \mathbb{N}}$ be sequences in $\mathcal{H}$ such that $\sum_{t \in \mathbb{N}} \alpha_t (\|a_t\| + \|b_t\|) < +\infty$. Fix $x^{(0)} \in \mathcal{H}$ and define the*

recursion,

$$x^{(t+1/2)} = \text{prox}_{\beta f_2}\left(x^{(t)}\right) + b_t \ , \tag{6}$$

$$x^{(t+1)} = x^{(t)} + \alpha_t \left(\text{prox}_{\beta f_1}\left(2x^{(t+1/2)} - x^{(t)}\right) + a_t - x^{(t+1/2)}\right) \ . \tag{7}$$

Then $x^{(t)} \rightharpoonup \bar{x}$ and $\text{prox}_{\beta\partial f_2}(\bar{x}) \in T$.

The sequences a_t and b_t in (6)-(7) play a prominent role as they formally establish the robustness of the DR algorithm to numerical errors when computing the proximity operators of f_1 and f_2 (e.g. via a nested inner iteration).

4.2. Sparse recovery with equality-constraint

From now, we assume that $\mathcal{H}$ is *finite-dimensional*. Let us first consider the general equality-constrained problem

$$(\mathsf{P}_{\text{eq}}): \quad \min_{x \in \mathcal{H}} \ \Psi(x) \quad \text{s.t.} \quad s = \mathrm{H}\Phi x \ , \tag{8}$$

$\Psi \in \Gamma_0(\mathcal{H})$ is a sparsity-promoting potential function, H is a bounded linear operator, and the equality constraint corresponds to an affine subspace that we denote $\mathcal{C}_{\text{eq}} = \{x \in \mathcal{H} | s = \mathrm{H}\Phi x\}$. We assume the following

Assumption 4 *(i) We suppose in the sequel that Ψ is additive; that is $\Psi(x) = \sum_{i\in\mathcal{I}} \psi_i(x_i)$. (ii) Ψ is coercive. (iii) $\mathrm{F} = \mathrm{H}\Phi$ is surjective. (iv) $s \in \mathrm{F}(\text{ri dom}\,\Psi)$.*

The following proposition gives the algorithm to solve (P_{eq}) and establishes its convergence.

Proposition 5 *Suppose that Assumptions 4(i)-(iv) are satisfied. Let β, $(\mu_t)_{t\in\mathbb{N}}$, $(a_t)_{t\in\mathbb{N}}$ and $(b_t)_{t\in\mathbb{N}}$ satisfy the conditions in Theorem 3. Fix $x^{(0)} \in \mathcal{H}$ and apply the DR recursion (6)-(7) with*

$$\text{prox}_{\beta f_2}(x) = \mathcal{P}_{\mathcal{C}_{\text{eq}}}(x) = x + \mathrm{F}^*\left(\mathrm{F}\mathrm{F}^*\right)^{-1}\left(s - \mathrm{F}x\right) \ , \tag{9}$$

$$\text{prox}_{\beta f_1}(x) = \text{prox}_{\beta\Psi}(x) = \left(\text{prox}_{\beta\psi_i}(x_i)\right)_{1\leqslant i\leqslant|\mathcal{I}|} \ . \tag{10}$$

Then $x^{(t)} \to \bar{x}$ and $\mathcal{P}_{\mathcal{C}_{\text{eq}}}(\bar{x})$ is a solution to (P_{eq}).

To implement this DR splitting algorithm, we need to express $\text{prox}_{\beta\psi_i}$. Its closed-form expression is given in [17] for a large class of penalty functions ψ_i. For instance, with $\psi_i(x_i) = \lambda|x_i|$, $\text{prox}_{\beta\psi_i} = \text{ST}_\lambda(x)$ the soft-thresholding operator.

4.3. Sparse recovery with inequality constraint

We then consider the inequality-constrained problem

$$(\mathsf{P}_\sigma): \quad \min_{x \in \mathcal{H}} \ \Psi(x) \quad \text{s.t.} \quad \|s - \mathrm{H}\Phi x\| \leqslant \sigma \ , \tag{11}$$

which can be viewed as a noise-aware variant of (P_{eq}). (P_σ) only asks the reconstruction be consistent with the data such

that the reconstruction error energy is within the noise level σ. Problem (P_σ) is challenging to solve for large-scale problems. The following proposition is a consequence of Lemma 2 and Theorem 3.

Proposition 6 *Under Assumptions 4(i)-(iii) and $0 \in \text{int}\left(s + B_2^\sigma\right) - \mathrm{F}(\text{ri dom}\,\Psi)$. Define β, $(\mu_t)_{t\in\mathbb{N}}$, $(a_t)_{t\in\mathbb{N}}$ and $(b_t)_{t\in\mathbb{N}}$ as before. Fix $x^{(0)} \in \mathcal{H}$ and apply the DR recursion (6)-(7) with $\text{prox}_{\beta f_1}$ as in (10) and*

$$\text{prox}_{\beta f_2}(x) = \mathcal{P}_{\mathcal{C}_\sigma}(x) \ , \tag{12}$$

where $\mathcal{P}_{\mathcal{C}_\sigma}(x)$ is given by Lemma 2 with $y = s$, $\text{prox}_f(u) = \mathcal{P}_{B_2^\sigma}(u) = (\mathrm{I} - \text{prox}_{\sigma\|\cdot\|})(u)$ and $\text{prox}_{\sigma\|\cdot\|}(u) = u\left(1 - \frac{\sigma}{\|u\|}\right)_+$. Then $x^{(t)} \to \bar{x}$ and $\mathcal{P}_{\mathcal{C}_\sigma}(\bar{x})$ is a solution to (P_σ).

If the data fidelity constraint in (P_σ) is replaced by any ℓ_p-norm for $p \geq 1$, the algorithm described above can be extended easily by appropriate substitution of $\mathcal{P}_{B_2^\sigma}(u)$ by the projector onto the ℓ_p-ball B_p^σ in Proposition 6. For example, in the case of uniform quantization noise, a good choice would be $p = \infty$, implying $\mathcal{P}_{B_\infty^\sigma}(u) = \left(\frac{u_i}{\max(|u_i|/\sigma, 1)}\right)_i$. Another interesting variant is when the noise is impulsive where the value $p = 1$ is advocated, and the projector onto $\mathcal{P}_{B_1^\sigma}$ can be computed through soft-thresholding.

4.4. The Dantzig Selector

Let's now turn to the problem

$$(\mathsf{P}_{\text{DS}}): \quad \min_{x \in \mathcal{H}} \ \Psi(x) \quad \text{s.t.} \quad \|\mathrm{F}^*(s - \mathrm{F}x)\|_\infty \leqslant \tau \ . \tag{13}$$

To avoid the unique trivial solution $x = 0$, we assume that $\tau < \|\mathrm{F}^*s\|_\infty$. The compact convex constraint set in (13) is $\mathcal{C}_{\text{DS}}$. The Dantzig Selector [11] is when Ψ is the ℓ_1-norm. In this case, (P_{DS}) can be recast as a linear program [11].

Let $\mathrm{G} = \mathrm{F}^*\mathrm{F}$ be the (self-adjoint) Gram operator and $z = \mathrm{F}^*s$. Beside Assumption 4(i) and (iii), suppose that $0 \in \text{int}\left(z + B_\infty^\tau\right) - \mathrm{G}(\text{ri dom}\,\Psi)$. Again, by straightforward application of Lemma 2 and Theorem 3 we arrive at

Proposition 7 *Fix $x^{(0)}$ and apply the DR recursion (6)-(7) with $\text{prox}_{\beta f_1}$ as in (10) and*

$$\text{prox}_{\beta f_2}(x) = \mathcal{P}_{\mathcal{C}_{\text{DS}}}(x) \ , \tag{14}$$

where $\mathcal{P}_{\mathcal{C}_{\text{DS}}}(x)$ is given by Lemma 2(iii) with G in lieu of F, $y = z$ and $(\mathrm{I} - \text{prox}_f)(u) = (\mathrm{I} - \mathcal{P}_{B_\infty^\tau})(u) = \text{ST}_\tau(u)$. Then $x^{(t)} \to \bar{x}$ and $\mathcal{P}_{\mathcal{C}_{\text{DS}}}(\bar{x})$ is a solution to (P_{DS}).

4.5. Analysis prior

In the problems considered above, we seek a sparse set of coefficients x and the solution image is synthesized from these representation coefficients $\bar{s} = \Phi x$. Such a prior is called

m	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}
DR	0.58	1.17	2.52	3.80	6.43	15.18	36.20
LARS	0.28	1.00	3.65	6.60	10.54	24.77	59.38
LP-IP	2.07	9.85	45.35	163.62	631.57	3191.38	13175.63
StOMP	0.34	0.84	2.36	3.61	6.38	16.65	43.13

Table 1. Computation times in seconds for exact CS reconstruction from Hadamard measurements.

a synthesis prior. A less common approach seeks an image s whose coefficients $x = \Phi^* s$ are sparse. Such a prior is called an analysis prior. This is typically the case for the Total Variation regularization. (P_{eq}), (P_σ), its Lagrangian form or (P_{DS}) can be modified to incorporate such a prior. For example, (P_σ) becomes

$$(Q_\sigma): \quad \min_{s \in \mathcal{K}} \Psi(\Phi^* s) \quad \text{s.t.} \quad \|s - Hx\|_2 \leqslant \sigma, \quad (15)$$

and similarly for the other problems. Despite its apparent difficulty, it turns out that the analysis-prior problems such as (Q_σ) can still be solved using (6)-(7). The only modification is the computation of the proximity operator of $f_1 = \Psi \circ \Phi^*$. The expression of the latter is easily accessible owing to Lemma 2(iii). Indeed, we have the following corollary.

Corollary 8 *Let* $\mu_t \in (0, 2/\left\|\left|\Phi\right|\right\|^2)$. *Define*

$$p^{(t+1)} = s - \mu_t \Phi \left(I - \text{prox}_{\mu_t^{-1} \Psi} \right) \circ \left(\mu_t^{-1} u^{(t)} + \Phi^* (s - \Phi u^{(t)}) \right).$$

Then $p^{(t)} \to \text{prox}_{\Psi \circ \Phi^*}$ *at the rate* $O(1/t)$.

5. NUMERICAL EXAMPLES

The proposed framework has been applied successfully to several inverse problems such as denoising, component separation, inpainting and CS. For obvious space limitation, we only report experiments on CS reconstruction.

In the first one, we solve (P_{eq}) with $\Psi(x) = \|x\|_1$, H is $m \times 4m$ random sensing matrix generated from the Hadamard ensemble, $\Phi = I$ (canonical basis), the support of x was selected uniformly at random with sparsity level $0.05m$ and normally distributed non-zero entries. Four algorithms were compared: our DR scheme (9)-(10), the LARS/LASSO [4], linear-programming with interior point solver [1, 3] and StOMP [18]. The results are shown in Table 1 (similar results were observed for Fourier measurements). As claimed, our DR splitting solver is among the fastest, and becomes faster than the other algorithms as m gets large. Fig 1 depicts the image recovered using (10)-(12) from 17% noisy (SNR=30dB) Fourier measurements of the 256^2 Mondrian image. The dictionary Φ was the wavelet transform.

6. CONCLUSION

In this paper, we proposed a general framework for solving several optimization problems arising in sparse solutions of
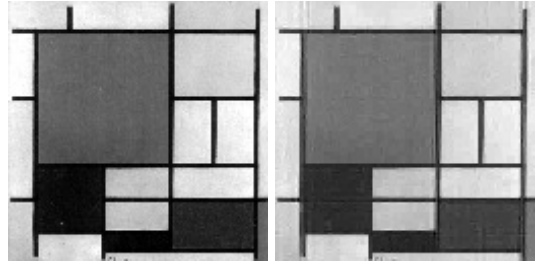

Fig. 1. Left: original 256^2 Mondrian image. Right: recovery from 17% noisy (SNR=30dB) Fourier measurements with the wavelet sparsity basis, SNR=22dB.

linear inverse problems. The algorithms were characterized and proven convergent. Sparse recovery from CS measurements illustrated the applicability of this framework.

7. REFERENCES

[1] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[4] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B,*, vol. 58, no. 1, pp. 267–288, 1996.

[5] M. Figueiredo and R. Nowak, "An em algorithm for wavelet-based image restoration," *IEEE Trans. on Image Processing*, vol. 12, pp. 906–916, 2003.

[6] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraints," *Comm. Pure Appl. Math.*, vol. 112, pp. 1413–1541, 2004.

[7] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, 2005.

[8] E. Candès, J. Romberg, , and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[9] E. Van Den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," Tech. Rep. TR-2008-01, Department of Computer Science, University of British Columbia, 2008.

[10] D.L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[11] E. Candès and T.Tao, "The dantzig selector: Statistical estimation when p is much larger than n," *Ann. Stat.*, vol. 35, pp. 2392–2404, 2007.

[12] R.T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.

[13] J.-J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *CRAS Sér. A Math.*, vol. 255, pp. 2897–2899, 1962.

[14] P. L. Combettes and J.-. Pesquet, "A Douglas-Rachford splittting approach to nonsmooth convex variational signal recovery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 564–574, 2007.

[15] Y. Nesterov, "Gradient methods for minimizing composite objective function," CORE Discussion Papers 2007076, UCL, Sept. 2007.

[16] P. L. Combettes, "Solving monotone inclusions via compositions of nonexpansive averaged operators," *Optimization*, vol. 53, no. 5-6, pp. 475–504, December 2004.

[17] M. J. Fadili, J.-L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *The Computer Journal*, vol. 52, pp. 64–79, 2007, in press.

[18] D.L. Donoho, Y. Tsaig, I. Drori, and J-L. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. on Information Theory*, 2006, submitted.