# **ITERATIVE SOFT-THRESHOLDING FOR TIME-VARYING SIGNAL RECOVERY**

Aurèle Balavoine, Christopher J. Rozell, and Justin Romberg

School of Electrical and Computer Engineering, Georgia Institute of Technology

# ABSTRACT

Recovering static signals from compressed measurements is an important problem that has been extensively studied in modern signal processing. However, only recently have methods been proposed to tackle the problem of recovering a time-varying sequence from streaming online compressed measurements. In this paper, we study the capacity of the standard iterative soft-thresholding algorithm (ISTA) to perform this task in real-time. In previous work, ISTA has been shown to recover static sparse signals. The present paper demonstrates its ability to perform this recovery online in the dynamical setting where measurements are constantly streaming. Our analysis shows that the  $\ell_2$ -distance between the output and the target signal decays according to a linear rate, and is supported by simulations on synthetic and real data.

*Index Terms*— Iterative Soft-Thresholding, sparse recovery, Compressed Sensing, time-varying signal

## 1. INTRODUCTION

Recent work in signal processing, in particular in the field of Compressed Sensing (CS), has developed powerful mathematical tools to acquire large signals at much lower rates than the traditional Nyquist rate [1]. In the static setting, noisy linear measurements  $y \in \mathbb{R}^M$  of a sparse signal  $a^{\dagger} \in \mathbb{R}^N$  are obtained via a measurement matrix  $\Phi$ , i.e.  $y = \Phi a^{\dagger} + \epsilon$  with  $M \ll N$ . If the signal is *S*-sparse, meaning that only *S* coefficients in  $a^{\dagger}$  are non-zero, a well-studied approach is to solve the optimization program

$$\hat{a}^{\dagger} = \arg\min_{a} \frac{1}{2} \|y - \Phi a\|_{2}^{2} + \lambda \|a\|_{1}.$$
 (1)

If the target is sufficiently sparse and the measurement matrix satisfies some properties, the minimum of this optimization provably recovers the target signal accurately and robustly with respect to noise [2]. Unfortunately, the above optimization is non-smooth and computing its optimal solution is computationally expensive. A large amount of work has been put to develop efficient algorithms that solve or approximately solve this problem. One of the well-known approaches to sparse signal recovery is the iterative soft-tresholding algorithm (ISTA) [3]. Its simple update rule makes it intuitive and easy to implement. In addition, ISTA provably minimizes the  $\ell_1$ -minimization (or LASSO) objective in (1), which is known to yield strong recovery guarantees. Although ISTA tends to converge slowly, many state-of-the-art solvers are slight variations of the standard ISTA [4–7].

In this paper, we show the potential of ISTA to perform online recovery of time-varying signals. In our setting, the ISTA iterations are performed as the measurements are received without waiting for convergence, overcoming the issue of slow convergence. Our analysis proves that, despite not letting the algorithm converge, the output is able to track the target signal over time. We prove that the  $\ell_2$ -distance decreases at a linear rate and tends to a minimum that is intuitive and essentially optimal. Our analysis is supported by simulation results on both synthetic and real data.

# 2. BACKGROUND

#### 2.1. Iterative Soft-Thresholding Algorithm

The approach we are interested in is the well-known iterative soft-thresholding algorithm (ISTA). Its step rule can be viewed as a generalized gradient step and its  $k^{\text{th}}$  iterate  $a(k) \in \mathbb{R}^N$  is defined by<sup>1</sup>

$$a(k+1) = T_{\lambda} \left( a(k) + \tau \left( \Phi^T \left( y - \Phi a(k) \right) \right) \right), \quad (2)$$

where  $T_{\lambda}(\cdot)$  is the soft-thresholding function operating entrywise and defined by

$$T_{\lambda}(x_n) = \max\left\{ |x_n| - \lambda, 0 \right\} \operatorname{sign}(x_n).$$
(3)

The constant  $\tau$  represents the step size of the gradient step, which is usually required to be in the interval  $(0, ||\Phi^T \Phi||^{-1})$  to ensure convergence. Several papers have shown that this algorithm converges to the solution to (1) as k goes to infinity from any initial point a(0) with linear rate [8,9].

Email: {aurele.balavoine,crozell,jrom}@gatech.edu.

<sup>&</sup>lt;sup>1</sup>We indicate the iterate number k in parenthesis, similar to a time index, and the  $n^{\text{th}}$  entry as a subscript:  $a_n(k)$ .

### 2.2. Related work

Encouraged by the positive results for static signals recovery, recent papers have proposed approaches that apply CS techniques to streaming signals. In [10-12], extensions of the classic Kalman filter that take sparsity into account are developed. In [13, 14], a Bayesian inference approach is used, where a probabilistic model of the target's dynamics needs to be adjusted. In [15, 16], an optimization program is set-up to take into account the time-correlation between samples and the recovery is performed blockwise. The methods presented in [17–21] are iterative schemes combined with techniques to reduce complexity. While they show good performance in practice, these approaches lack theoretical guarantees for the convergence and accuracy, or only present theoretical results in the static setting. In [22], a proof of convergence is given, but there is no accuracy result and it is not clear if the conditions for convergence apply when the target is time-varying.

In our analysis, no a priori information on the signal's dynamics is required. The standard ISTA is studied when the input contains measurements of a time-varying signal, and no convergence criterion is imposed before moving to the next time instance. As a consequence, the number of iterations performed by ISTA for each measurement is entirely dictated by the sampling rate. Our results show that even in this iteration-limited setting, ISTA is able to track the underlying time-varying signal, and gets closer in the  $\ell_2$ -distance as the number of iterations increases. While other papers, such as [12] had pointed out that limiting the number of iterations still yields good convergence in practice, to the best of our knowledge no theoretical analysis has been performed.

### 3. TIME-VARYING SIGNAL RECOVERY

## 3.1. Signal and Measurement Model

We consider the problem where the target  $a^{\dagger}$  varies with time. A new measurement is received every  $P^{\text{th}}$  ISTA iterate:

$$y(kP) = \Phi a^{\dagger}(kP) + \epsilon(kP), \qquad \forall k \ge 0 \tag{4}$$

where the noise vector  $\epsilon(kP)$  may also vary with time. As a consequence, the ISTA  $l^{\text{th}}$  iterate is:

$$u(l+1) = a(l) + \tau \left( \Phi^T \left( y(l) - \Phi a(l) \right) \right)$$
  
$$a(l+l) = T_\lambda(u(l+l)) \qquad , \forall l \ge 0 \quad (5)$$

For iterates l of the form l = kP + i, with i = 0, ..., P - 1, the target signal  $a^{\dagger}(kP + i)$  and the measurements y(kP + i)are treated as constant signals, while the differences between consecutive samples are assumed to be bounded:

$$a^{\dagger}(kP+i) = a^{\dagger}(kP), \ \forall k \ge 0, \forall i = 0, \dots, P-1,$$
 (6)

$$\left\|a^{\dagger}(kP) - a^{\dagger}(kP-1)\right\|_{2} \le \mu, \qquad \forall k \ge 0.$$
<sup>(7)</sup>

We also assume that the energy in the target is bounded:

$$\left\|a^{\dagger}(l)\right\|_{2} \le \eta, \qquad \forall l \ge 0, \tag{8}$$

We introduce the extra variable u(l) above to simplify notations in the proof of the main result. Note that because the measurement vector y(l) changes every P iterations, if P is small, ISTA never converges to the minimum of (1). This approach is of great interest for scenarios where the measurements are streaming at very high rates or computational resources are limited.

# 3.2. Main Result

The following theorem shows that the number of non-zero elements in the output remains bounded. It also provides an expression for the evolution of the  $\ell_2$ -distance between the output and the target signal.

**Theorem 1.** Assume that the dictionary  $\Phi$  and the gradient step size  $\tau$  in (5) satisfy

$$1 \le \left\| \Phi^T \Phi \right\| < 1/\tau, \tag{9}$$

where  $\|\cdot\|$  is the operator norm. Define  $c = \|\tau \Phi^T \Phi - I\| < 1$ . If the target signal satisfies conditions (7) and (8), the initial point a(0) contains less than q active nodes and the following condition holds for some q > 0

$$\max\left\{\left\|u(0)\right\|_{2}, \ c\lambda\sqrt{q}\right\} + \eta + \sigma \le \lambda\sqrt{q}, \qquad (10)$$

then

- the output a(l) never contains more than q active nodes for all l ≥ 0; and
- 2. the  $\ell_2$ -distance between the output and the target signal satisfies  $\forall l \ge 0$ , letting  $i = (l \mod P)$  (i.e.,  $\exists k \ge 0$  such that l = kP + i, with  $0 \le i \le P 1$ )

$$\begin{aligned} \left| a(l+1) - a^{\dagger}(l) \right\|_{2} &\leq c^{l} \left( \left\| a(1) - a^{\dagger}(0) \right\|_{2} - W \right) \\ &+ \frac{c^{i+1}}{1 - c^{P}} \mu + D, \end{aligned}$$
(11)

where

$$D = (1 - c)^{-1} (\lambda \sqrt{q} + \sigma), \qquad (12)$$

$$W = \frac{c\mu}{1 - c^P} + D. \tag{13}$$

This theorem shows that every  $P^{\text{th}}$  iteration, the  $\ell_2$ -distance between the output a(kP) and the target signal  $a^{\dagger}(kP-1)$  remains bounded and converges as k goes to infinity towards

$$D + \frac{c^{P}}{1 - c^{P}}\mu = (1 - c)^{-1} \left(\lambda \sqrt{q} + \sigma\right), + \frac{c^{P}}{1 - c^{P}}\mu$$

with a linear rate of convergence. Note that this final value is essentially optimal. The first term  $(1 - c)^{-1} (\lambda \sqrt{q} + \sigma)$ corresponds to the error we expect from solving (1). Together with the bound (10), they resemble the terms of Corollary 5.1 in [23] for the static case. The additional term  $c^P(1 - c^P)^{-1}\mu$  behaves like  $\mu/P$  and corresponds to the error we expect from having a time-varying input. The larger the variations in the target, the larger  $\mu$  is, which corresponds to a more difficult signal to track and a larger error. Moreover, the slower the target varies, the larger P can be, and as expected, the smaller the final error is. When P goes to infinity, this additional term disappears.

## 3.3. Proof

*Proof.* First, we show by induction on l that  $||u(l)||_2 \le \lambda \sqrt{q}$ . Note that if  $||u(l)||_2 \le \lambda \sqrt{q}$ , then no more than q entries in u(l) have absolute value greater than  $\lambda$ , so no more than q entries are non-zero in  $a(l) = T_{\lambda}(u(l))$ . Thus, this induction proves 1. of the theorem.

We will use the fact that (9) implies  $c = \|\tau \Phi^T \Phi - I\| < 1$ and  $\tau \|\Phi\| \le 1/\sqrt{\|\Phi^T \Phi\|} \le 1$ . Also note that if  $\|u(l)\|_2 \le \lambda \sqrt{q}$ , then (3) implies that  $\|a(l)\|_2 \le \|u(l)\|_2 \le \lambda \sqrt{q}$  and  $\|u(l) - a(l)\|_2 \le \|u(l)\|_2 \le \lambda \sqrt{q}$ .

At iteration l = 0, (10) implies that  $||u(0)||_2 \le \lambda \sqrt{q}$ . Now, assume that for some  $l \ge 0$ ,  $||u(l)||_2 \le \lambda \sqrt{q}$ .

$$\begin{aligned} \|u(l+1)\|_{2} &= \|a(l) + \tau \Phi^{T} (y(l) - \Phi a(l))\|_{2} \\ &= \|- \left(\tau \Phi^{T} \Phi - I\right) a(l) \\ &+ \tau \Phi^{T} \Phi a^{\dagger}(l) + \tau \Phi^{T} \epsilon(l)\|_{2} \\ &\leq \|\tau \Phi^{T} \Phi - I\| \|a(l)\|_{2} \\ &+ \tau \|\Phi^{T} \Phi\| \|a^{\dagger}(l)\|_{2} + \tau \|\Phi\| \sigma \\ &\leq c \|a(l)\|_{2} + \|a^{\dagger}(l)\|_{2} + \sigma \\ &\leq c \lambda \sqrt{q} + \eta + \sigma \qquad \leq \lambda \sqrt{q}. \end{aligned}$$

So the property holds at l + 1 and the induction is proven.

We now show by induction on l that (11) holds. At l = 0, the property obviously holds. Assume that it holds for some  $l \ge 0$ . There exist a unique  $k \ge 0$  and a unique  $0 \le i \le P-1$ such that l = kP + i. Applying 1., we compute

$$\begin{split} \left\| a(l+2) - a^{\dagger}(l+1) \right\|_{2} \\ &\leq \|a(l+2) - u(l+2)\|_{2} + \|u(l+2) - a^{\dagger}(l+1)\|_{2} \\ &\leq \|u(l+2)\|_{2} + \|u(l+2) - a^{\dagger}(l+1)\|_{2} \\ &= \lambda \sqrt{q} + \|\tau \Phi^{T} \epsilon(l) \\ &\qquad + \left(\tau \Phi^{T} \Phi - I\right) \left(a^{\dagger}(l+1) - a(l+1)\right)\|_{2} \\ &\leq \lambda \sqrt{q} + \sigma \\ &\qquad + c \left\|a^{\dagger}(l+1) - a^{\dagger}(l)\right\|_{2} + c \left\|a(l+1) - a^{\dagger}(l)\right\|_{2}. \end{split}$$

There are two scenarios: i = P - 1 or  $0 \le i \le P - 2$ .

First case: When i = P - 1, l = (k + 1)P - 1 and (7) yields  $\|a^{\dagger}(l+1) - a^{\dagger}(l)\|_{2} \le \mu$ . So, using the induction hypothesis (11) at l gives

$$\begin{split} \left\| a(l+2) - a^{\dagger}(l+1) \right\|_{2} \\ &\leq c \left\| a(l+1) - a^{\dagger}(l) \right\|_{2} + c\mu + \lambda \sqrt{q} + \sigma \\ &\leq c \left( c^{l} \left( \left\| a(1) - a^{\dagger}(0) \right\|_{2} - W \right) + \frac{c^{P}}{1 - c^{P}} \mu + D \right) \\ &\quad + \lambda \sqrt{q} + \sigma + c\mu \\ &\leq c^{l+1} \left[ \left\| a(1) - a^{\dagger}(0) \right\|_{2} - W \right] \\ &\quad + \frac{c^{P+1}}{1 - c^{P}} \mu + cD + \lambda \sqrt{q} + \sigma + c\mu \\ &\leq c^{l+1} \left[ \left\| a(1) - a^{\dagger}(0) \right\|_{2} - W \right] + \frac{c}{1 - c^{P}} \mu + D. \end{split}$$

So the induction hypothesis holds for l + 1 = (k + 1)P. Second case: When  $0 \le i \le P - 2$ , (6) yields

Since l + 1 = kP + (i + 1), with  $1 \le i + 1 \le P - 1$ , this proves the induction hypothesis in the second case and finishes the proof.

## 4. SIMULATIONS

### 4.1. Synthetic data

We start by testing the approach on synthetic data. A sparse vector  $a^{\dagger}(0)$  of length N = 512 with S = 40 non-zero locations chosen at random by drawing 40 amplitudes from a normal Gaussian distribution and normalizing them to have norm 1. Then, we generate 99 consecutive time samples of  $a^{\dagger}(k)$  as follows:

$$a^{\dagger}(k+1) = \sqrt{\frac{\eta^2 - \mu^2}{\eta^2}} a^{\dagger}(k) + \frac{\mu}{\sqrt{S}} v(k)$$

where v(k) is a vector with same sparsity pattern as  $a^{\dagger}$  and amplitudes drawn from a standard Gaussian distribution. We pick  $\eta = 1$  and  $\mu = 0.5$ . Each resulting sample  $a^{\dagger}(k)$ has energy equal to  $\eta$  and innovation with energy proportional to  $\mu$  in expectation. The measurement matrix  $\Phi$ , of



**Fig. 1**.  $\ell_2$ -distance between the target and the output at the *P*<sup>th</sup> ISTA iterations for 5 values of *P* averaged over 1000 trials.



Fig. 2. Ratio of the maximum number of non-zero elements q over the sparsity level S for several values of  $\lambda$  and S averaged over 100 trials.

size  $256 \times 512$ , is filled with random normal Gaussian variables and its columns are normalized to 1. Gaussian white noise with standard deviation  $0.3 \|\Phi a^{\dagger}(0)\|_2 / \sqrt{M}$  is added to the measurements, which corresponds to a moderate level of noise. We run ISTA in (3.1) for 5 different values of P, and average over 1000 such trials. The step size  $\tau$  is chosen equal to  $1/\|\Phi^T\Phi\|$ , and the threshold is  $\lambda = 0.4$ . In Fig.1, we plot the  $\ell_2$ -error  $\|a(kP) - a^{\dagger}(kP - 1)\|_2$ . We observe that the curves tend to a final value that matches the behavior of  $c^P (1 - c^P)^{-1}$  predicted by the theorem as k goes to infinity.

Next, we vary the threshold  $\lambda$  and the sparsity level S, and generate for each pair 10 time samples of  $a^{\dagger}(k)$  and associated measurements y(k) in the same fashion as above. We run ISTA for P = 10 iterations per measurement. In Fig.2, we plot the average over 100 such trials of the ratio of the maximum number of non-zero elements q in a(l) over the sparsity level S. The plot shows that the maximum number of non-zero elements remains small (q is mostly contained between 1S and 10S), which matches the theorem's prediction.

### 4.2. Real data

We follow the work in [24], where an efficient approach to image acquisition and recovery, called "compressive imaging", is developed. Since natural images are sparse in wavelets, the matrix  $\Phi$  used is  $\Phi = AB$ , where A con-



**Fig. 3**. rMSE and CPU time for each random 100 consecutive frames of the recovered foreman video sequence wavelet coefficients averaged over 50 trials.

sists of M = 0.25N random rows of a noiselet matrix and B is a dual-tree discrete wavelet transform (DT-DWT) [25]. We compare the performance of ISTA against TFOCS, a state-of-the-art LASSO solver applied independently on each frame [26], BPDN-DF that adds additional time-dependent regularization between frames, as well as RWL1-DF which additionally performs reweighting at each iteration [11]. On Fig.3, the regularized mean-squared error, defined by  $rMSE(k) = \frac{||a(k) - a^{\dagger}(k)||_2}{||a^{\dagger}(k)||_2}$ , and the CPU time per frame, averaged over 50 random sequences of 100 frames

of the foreman video <sup>2</sup>, are plotted. The results show that for both P = 3 and P = 10, the gain in time of ISTA is significant. For P = 10, the rMSE is similar to the one for BPDN and BPDN-DF. The rMSE for RWL1-DF is much lower, due to the additional reweighting steps, however, the CPU time for this method is much larger than the other four, making it impractical for applications where the measurements are streaming at very high rates.

# 5. CONCLUSION

In this paper, we showed that the standard ISTA can be used as an online solver to recover time-varying signals from streaming compressed measurements. Our analysis provides an analytic expression for the  $\ell_2$ -distance between the ISTA iterate and the target signal that depends on the number of time samples as well as the number of iterations performed per measurement. This analysis is particularly useful in the scenario where the measurements are streaming at very high rates and only a few iterations can be performed. Simulations showed that the theoretical behavior is followed in practice, and that the method is applicable to real data. Our analysis could potentially be extended to other solvers for sparse recovery and help analyze their behavior when the iteration is stopped before convergence.

<sup>&</sup>lt;sup>2</sup>http://www.hlevkin.com/TestVideo/foreman.yuv

## 6. REFERENCES

- E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [2] E. J. Candès and Y. Plan, "Near-ideal model selection by 11 minimization," Ann. Stat., vol. 37, pp. 2145–2177, Oct. 2009.
- [3] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Aug. 2004.
- [4] A. Beck and M. Teboulle, "A fast iterative shrinkagethresholding algorithm with application to wavelet-based image deblurring," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2009, pp. 693–696.
- [5] J. M Bioucas-Dias and M. A.T. Figueiredo, "A new TwIST: two-step iterative Shrinkage/Thresholding algorithms for image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2992–3004, Dec. 2007.
- [6] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
- [7] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, 2009.
- [8] K. Bredies and D. A. Lorenz, "Linear convergence of iterative soft-thresholding," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 813–837, Dec. 2008.
- [9] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for 11-minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, Jan. 2008.
- [10] N. Vaswani, "Kalman filtered compressed sensing," in *IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 893–896.
- [11] A. S. Charles and C. J. Rozell, "Dynamic filtering of sparse signals using reweighted 11," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2013, pp. 6451–6455.
- [12] W. Dai, D. Sejdinovic, and O. Milenkovic, "Gaussian dynamic compressive sensing," in *Int. Conf. Sampling Theory and Appl.*, Singapore, May 2011.
- [13] D. Sejdinovic, C. Andrieu, and R. Piechocki, "Bayesian sequential compressed sensing in sparse dynamical systems," in *48th Annu. Allerton Conf. Commun., Control, Comput.*, Sept. 2010, pp. 1730–1736.
- [14] J. Ziniel, L.C. Potter, and P. Schniter, "Tracking and smoothing of time-varying sparse signals via approximate belief propagation," in *Asilomar Conf. Signals, Syst., Comput.*, Nov. 2010, pp. 808–812.
- [15] D. Angelosante, G.B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *16th Int. Conf. Digit. Signal Process.*, July 2009, pp. 1–8.
- [16] D. Angelosante, J.A. Bazerque, and G.B. Giannakis, "Online adaptive estimation of sparse signals: Where RLS meets the 11norm," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3436– 3447, July 2010.

- [17] M. S. Asif and J. Romberg, "Dynamic updating for 11 minimization," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 421–434, Apr. 2010.
- [18] Y.V. Zakharov and V.H. Nascimento, "DCD-RLS adaptive filters with penalties for sparse identification," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3198–3213, June 2013.
- [19] K. Slavakis, Y. Kopsinis, and S. Theodoridis, "Adaptive algorithm for sparse system identification using projections onto weighted 11 balls," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 3742–3745.
- [20] B. Babadi, N. Kalouptsidis, and V. Tarokh, "SPARLS: the sparse RLS algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4013–4025, Aug. 2010.
- [21] Y. Chen, Y. Gu, and A.O. Hero, "Sparse LMS for system identification," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2009, pp. 3125–3128.
- [22] K. Slavakis, Y. Kopsinis, S. Theodoridis, and S. McLaughlin, "Generalized thresholding and online sparsity-aware learning in a union of subspaces," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3760–3773, Aug. 2013.
- [23] Tong Zhang, "Some sharp performance bounds for least squares regression with 11 regularization," *Ann. Statist.*, vol. 37, no. 5A, pp. 2109–2144, Oct. 2009.
- [24] J. Romberg, "Imaging via compressive sampling," *IEEE Sig. Process. Mag.*, vol. 25, no. 2, pp. 14–20, Mar. 2008.
- [25] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.
- [26] S. R. Becker, E. J. Candès, and M. C. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Math. Prog. Comput.*, vol. 3, no. 3, pp. 165–218, Sept. 2011.