

EFFICIENT RECOVERY OF BLOCK SPARSE SIGNALS VIA ZERO-POINT ATTRACTING PROJECTION

Jingbo Liu, Jian Jin, Yuantao Gu*

State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, CHINA

ABSTRACT

In this paper, we consider compressed sensing (CS) of block-sparse signals, i.e., sparse signals that have nonzero coefficients occurring in clusters. An efficient algorithm, called zero-point attracting projection (ZAP) algorithm, is extended to the scenario of block CS. The block version of ZAP algorithm employs an approximate $l_{2,0}$ norm as the cost function, and finds its minimum in the solution space via iterations. For block sparse signals, an analysis of the stability of the local minimums of this cost function under the perturbation of noise reveals an advantage of the proposed algorithm over its original non-block version in terms of reconstruction error. Finally, numerical experiments show that the proposed algorithm outperforms other state of the art methods for the block sparse problem in various respects, especially the stability under noise.

Index Terms— Compressed sensing, sparse recovery, block sparse, zero-point attracting projection.

1. INTRODUCTION

Compressed sensing (CS) [1], [2] addresses the problem of retrieving sparse signals from under-determined linear measurements. It enjoys the advantage of reducing computational complexity in the measurement stage, and therefore has shown a great potential in applications such as MRI imaging [3], wireless communication [4], pattern recognition [5], and source coding [6]. On the part of signal reconstruction in CS, one of the key problems is to retrieve the sparsest solution, i.e., the minimum l_0 norm solution to the equations of linear constraints:

$$\min \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ is the unknown sparse signal, $\mathbf{y} \in \mathbb{R}^m$ is the measurement, and typically $m < n$. Unfortunately, l_0 norm minimization problem is generally an NP hard problem. Previous work including [7] and [1] have shown that under some conditions, the sparsest solution can be obtained via convex relaxation, such as Basis Pursuit (BP). Another popular method for CS recovery problem is based on greedy pursuits, and its representative is orthogonal matching pursuit (OMP) [8].

The block sparse problem for compressed sensing was first introduced by Eldar et.al in [9]. The authors have shown that sampling problems over unions of subspaces can be converted into block-sparse recovery problems. Examples in applications can be found

in [10], [11] and [12]. Mathematically, a block-sparse signal can be represented as follows:

$$\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_N^T]^T, \quad (2)$$

where $\mathbf{x}_i = [x(iD - D + 1), \dots, x(iD)]$ is the i th block of \mathbf{x} with length D , and $n = N \times D$. A signal is K block-sparse if at most K out of the N blocks of the signal are non-zero. Similar to (1), the $l_{2,0}$ norm minimization for the block-sparse problem can be formulated as:

$$\min \|\mathbf{x}\|_{2,0}, \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (3)$$

where the $l_{p,q}$ norm of a block vector \mathbf{x} is defines as:

$$\|\mathbf{x}\|_{p,q} = \sum_{k=1}^N (\|\mathbf{x}_k\|_p)^q. \quad (4)$$

From (4) it is clear that the $l_{2,0}$ norm can be interpreted as the number of non-zero blocks of the signal. Like the l_0 minimization problem (1), solving (3) is also NP-hard. Although all the conventional recovery algorithms in CS is also applicable to the block CS problem, these algorithms cannot take advantage of the essential block-sparse characteristic of signals. To remedy this, Eldar et.al introduced two algorithms in [13] and [9]: the $l_{2,1}$ -opt and the Block orthogonal matching pursuits (BOMP). However, like their ancestors, these algorithms have their inherit drawbacks: $l_{2,1}$ -opt is quite slow and becomes worse as dimension increases; BOMP is faster, but its estimation accuracy is poorer in the presence of noise perturbation.

In contrast, a recently proposed algorithm called zero-point attracting projection (ZAP) [14] is an efficient sparse reconstruction method based on an idea different from the aforementioned convex relaxations and greedy pursuits: The authors choose a smooth function to approximate the l_0 norm and then finds its minimum in the solution space via iterations. Their simulations show that ZAP requires fewer measurements for exact reconstruction than the referenced algorithms in the experiment settings, while having tractable computational complexity. It is then interesting to explore the block-sparse reconstruction methods based on the idea of ZAP.

In this paper, the ZAP algorithm is extended to the block sparse model. The block ZAP algorithm (BZAP) employs a smoother cost function to approximate the $l_{2,0}$ norm of the block sparse input, and then minimize this function via iterations. An analysis of the stability of the local minimums of the cost function gives a lower bound on the reconstruction error for BZAP than the original ZAP, approximately by a factor of $1/\sqrt{D}$. Simulations show that BZAP outperforms other state-of-the-art methods for the block sparse problem (BOMP, $l_{2,1}$ -opt) both in terms of incidence of exact recovery in

*This work was partially supported by National Natural Science Foundation of China (NSFC 60872087 and NSFC U0835003). The corresponding author of this paper is Yuantao Gu (gyt@tsinghua.edu.cn).

noiseless case, and the mean square deviation in the case of noise-contaminated measurements.

The remainder of the paper is organized as follows: section 2 presents the formulation of the BZAP algorithm. In section 3, an analysis of the l_2 stability of the local minimum of the cost function for BZAP is offered. Section 4 presents simulation results comparing BZAP with BOMP, $l_{2,1}$ -opt and the original ZAP. Finally, the whole paper is concluded in section 5.

Notation. Throughout this paper, we denote vectors by boldface lowercase letters, and matrices by boldface upper case letters. Given a matrix \mathbf{A} , \mathbf{A}^* is its Hermitian conjugate. \mathbf{A}^\dagger denotes the pseudo inverse of \mathbf{A} , that is, if \mathbf{A} has full row rank or full column rank, then

$$\mathbf{A}^\dagger = \begin{cases} \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}, & \mathbf{A} \text{ has full row rank;} \\ (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*, & \mathbf{A} \text{ has full column rank.} \end{cases} \quad (5)$$

Block support T is a subset of $\{1 \dots N\}$ indicating the non-zero blocks of \mathbf{x} , and T^c is its complement. We use \mathbf{x}_T to denote the vector formed by the blocks in \mathbf{x} indexed by T , and \mathbf{A}_T the sub-matrix that lies in the column blocks indexed by T . Notation $\|\cdot\|$ takes either the Euclid norm of a vector or the l_2 operator norm of a matrix.

2. BLOCK ZAP ALGORITHM

This section aims to extend the ZAP algorithm to the block sparse problem. One chief idea of BZAP is to employ a ‘smoother’ function:

$$J(\mathbf{x}) = \sum_{k=1}^N F(\|\mathbf{x}_k\|) \quad (6)$$

to approximate the $l_{2,0}$ norm of \mathbf{x} . Of course, there is a great liberty in the choice of the function F in (6). But to reduce computation complexity, we select

$$F_\alpha(w) = \begin{cases} 2\alpha|w| - \alpha^2 w^2 & |w| \leq \frac{1}{\alpha}; \\ 1 & \text{elsewhere,} \end{cases} \quad (7)$$

in the implementations, since its derivative is linear. Now from (4) we see that the $l_{2,0}$ norm of \mathbf{x} can be approximated as:

$$\|\mathbf{x}\|_{2,0} \approx \sum_{k=1}^N F_\alpha(\|\mathbf{x}_k\|), \quad (8)$$

So the problem (3) is transferred to

$$\min \sum_{k=1}^N F_\alpha(\|\mathbf{x}_k\|), \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (9)$$

Traditional methods of steepest descent together with a ‘projection’ step can be used to solve (9). That is, in the t th iteration, the solution is updated along the negative gradient direction of the sparse penalty, which in effect attracts the solution to the zero point:

$$\tilde{\mathbf{x}}(t+1) = \mathbf{x}(t) - \kappa \cdot \nabla J(\mathbf{x}(t)). \quad (10)$$

Since $\tilde{\mathbf{x}}(t+1)$ is generally not in the solution space, the next step is to project it back to the hyperplane of $\mathbf{A}\mathbf{x} = \mathbf{y}$:

$$\mathbf{x}(t+1) = \mathbf{P}\tilde{\mathbf{x}}(t+1) + \mathbf{Q}, \quad (11)$$

where $\mathbf{P} = \mathbf{I} - \mathbf{A}^\dagger \mathbf{A}$ is named as projection matrix and $\mathbf{Q} = \mathbf{A}^\dagger \mathbf{y}$. The attraction step (10) and projection step (11) are used alternately

Table 1. Procedure Outline of BZAP

Input: $\alpha, \kappa, \mathbf{A}, \mathbf{y}$;
Initialize BZAP: $\mathbf{x}_0(0) = \mathbf{A}^\dagger \mathbf{y}, t = 0$.
Repeat: (for time instant t);
Update $\tilde{\mathbf{x}}(t+1)$ with the zero attraction by (10) and (8);
Project $\mathbf{x}(t+1)$ back to the solution space by (11);
Update the index: $t = t + 1$.
Until: Block ZAP stop criterion is satisfied.

in the iterations, hence the name of zero-point attracting projection. The procedure of BZAP is summarized in TABLE 1.

Finally, we remark on the choice of parameters for the BZAP algorithm:

The choice of α : According to (7), parameter α determines the range of effect of the cost function. There is a tradeoff in the choice of α since a small α leads to a bad approximation of the $l_{2,0}$ norm, and produces many local minimums, while an overly large α limits the effective range. Empirically we have found that BZAP performs the best when $1/\alpha$ is around the square root of the variation of the non-zero entries in \mathbf{x} .

The choice of κ : The step length κ determines the speed of convergence and the accuracy of the estimation. A large κ will result in a fast convergence but a poor estimation. In our simulations, κ is decreased as the iterations approaches convergence, in order to ensure both speed of convergence and accuracy. More specifically, we let κ decrease by a factor of η ($\eta < 1$) whenever the cost function (6) starts to increase.

Stop conditions: The iteration (10) and (11) is terminated when any of the two following conditions is satisfied: (a) The total number of reductions of step length κ reaches a predefined number C_1 ; or (b) The total number of iterations exceeds a predefined number C_2 .

3. STABILITY OF THE LOCAL MINIMUM POINT

In this section, we consider the problem of estimating $\bar{\mathbf{x}}$ from the following noisy measurements:

$$\mathbf{y} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{v}. \quad (12)$$

While in [15], the authors have discussed the convergence of the ZAP iterations, in this work we mainly consider the stability of local minimums of the cost function of BZAP under noise perturbation.

Suppose α satisfies

$$1/\alpha < \|\bar{\mathbf{x}}_k\|, \quad k \in T, \quad (13)$$

define the closed ball $B(\bar{\mathbf{x}}, d)$ as a neighborhood of $\bar{\mathbf{x}}$, where

$$d = \min_{k \in T} (1/\alpha, \|\bar{\mathbf{x}}_k\| - 1/\alpha).$$

Let L be the solution space:

$$L := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{y} = \mathbf{A}\mathbf{x}\},$$

where \mathbf{y} is the measurement given in (12). Then, regarding the stability of the local minimizer of (9) in the noise-contaminated measurements, we have the following theorem:

Theorem 1. Suppose $\bar{\mathbf{x}}$ is a block sparse signal and \mathbf{A}_T has full column rank, then the minimizer $\hat{\mathbf{x}}$ of function (6) in the region $L \cap B(\bar{\mathbf{x}}, d)$ satisfies

$$\|\hat{\mathbf{x}} - \bar{\mathbf{x}}\| \leq 2\sqrt{N}(1 + \|\mathbf{A}_T^\dagger \mathbf{A}_{T^c}\|) \|\mathbf{A}^\dagger \mathbf{v}\| + \|\mathbf{A}_T^\dagger \mathbf{v}\|. \quad (14)$$

Proof. Let

$$\delta \mathbf{x} = \hat{\mathbf{x}} - \bar{\mathbf{x}} \quad (15)$$

be the difference between the local minimum of the cost function and the real signal. The aim is to bound $\|\delta \mathbf{x}\|$ with \mathbf{v} . Obviously,

$$\|\delta \mathbf{x}\| \leq \|\delta \mathbf{x}_T\| + \|\delta \mathbf{x}_{T^c}\|. \quad (16)$$

Then we will derive bounds on $\|\delta \mathbf{x}_T\|$ and $\|\delta \mathbf{x}_{T^c}\|$ respectively:

First, consider the bound on $\|\delta \mathbf{x}_{T^c}\|$. If $\mathbf{x} \in B(\bar{\mathbf{x}}, d)$, then

$$\|\mathbf{x}_k - \bar{\mathbf{x}}_k\| < d, \quad k = 1 \dots N,$$

and it follows from the definition of d that

$$\begin{aligned} \|\mathbf{x}_k\| &> 1/\alpha, & k \in T; \\ \|\mathbf{x}_k\| &< 1/\alpha, & k \in T^c. \end{aligned} \quad (17)$$

Therefore, with the cost functions defined in (6), (7), we have

$$\alpha \|\mathbf{x}_{T^c}\|_{2,1} \leq \sum_{k \in T^c} F_\alpha(\|\mathbf{x}_k\|) \leq 2\alpha \|\mathbf{x}_{T^c}\|_{2,1}. \quad (18)$$

Next, we will prove

$$\|\delta \mathbf{x}_{T^c}\| \leq 2\sqrt{N} \|\mathbf{A}^\dagger \mathbf{v}\| \quad (19)$$

by differentiating between the following two situations:

- 1) If $\|\mathbf{A}^\dagger \mathbf{v}\| \geq d$, then (19) automatically holds.
- 2) If $\|\mathbf{A}^\dagger \mathbf{v}\| < d$, we have

$$\|\delta \mathbf{x}_{T^c}\| \leq \|\delta \mathbf{x}_{T^c}\|_{2,1} \quad (20)$$

$$\leq \frac{1}{\alpha} \sum_{k \in T^c} F_\alpha(\|\mathbf{x}_k\|) \quad (21)$$

$$= \min_{\mathbf{x}} \frac{1}{\alpha} \sum_{k \in T^c} F_\alpha(\|\mathbf{x}_k\|), \text{ s.t. } \|\mathbf{x}\| < d, \mathbf{A}\mathbf{x} = \mathbf{v} \quad (22)$$

$$\leq \min_{\mathbf{x}} 2\|\mathbf{x}\|_{2,1}, \text{ s.t. } \|\mathbf{x}\| < d, \mathbf{A}\mathbf{x} = \mathbf{v} \quad (23)$$

$$\leq 2\|\mathbf{A}^\dagger \mathbf{v}\|_{2,1} \quad (24)$$

$$\leq 2\sqrt{N} \|\mathbf{A}^\dagger \mathbf{v}\|, \quad (25)$$

where the definition of $\delta \mathbf{x}$ is used in the derivation of (22), relation (18) in the derivation of (23), and the fact that $\mathbf{A}^\dagger \mathbf{v}$ is a feasible point for the constraint of (23) in the derivation of (24). To conclude, (19) holds in both situations.

Finally, we derive a bound on $\|\delta \mathbf{x}_T\|$. Since

$$\mathbf{A}_T \delta \mathbf{x}_T = \mathbf{v} - \mathbf{A}_{T^c} \delta \mathbf{x}_{T^c},$$

we have

$$\delta \mathbf{x}_T = \mathbf{A}_T^\dagger (\mathbf{v} - \mathbf{A}_{T^c} \delta \mathbf{x}_{T^c}),$$

therefore by triangular inequality and (19),

$$\begin{aligned} \|\delta \mathbf{x}_T\| &\leq \|\mathbf{A}_T^\dagger \mathbf{v}\| + \|\mathbf{A}_T^\dagger \mathbf{A}_{T^c} \delta \mathbf{x}_{T^c}\| \\ &\leq \|\mathbf{A}_T^\dagger \mathbf{v}\| + 2\sqrt{N} \|\mathbf{A}_T^\dagger \mathbf{A}_{T^c}\| \|\mathbf{A}^\dagger \mathbf{v}\|. \end{aligned} \quad (26)$$

Then, combining (25) and (26) yields the final result (14). \square

Now, we remark on the improvement of BZAP over ZAP: Since ZAP can be seen as the $D = 1$ special case of BZAP, when theorem 1 is applied to ZAP, the bound becomes

$$\|\hat{\mathbf{x}} - \bar{\mathbf{x}}\| \leq 2\sqrt{n}(1 + \|\mathbf{A}_T^\dagger \mathbf{A}_{T^c}\|) \|\mathbf{A}^\dagger \mathbf{v}\| + \|\mathbf{A}_T^\dagger \mathbf{v}\|. \quad (27)$$

It will be shown later that the term $\|\mathbf{A}_T^\dagger \mathbf{v}\|$ in (14) is equivalent with the error of a so-called ‘oracle estimator’, which gives a lower bound on the mean square error for the recovery problem. The other term, $2\sqrt{n}(1 + \|\mathbf{A}_T^\dagger \mathbf{A}_{T^c}\|) \|\mathbf{A}^\dagger \mathbf{v}\|$ in (27) is reduced by BZAP by a factor of \sqrt{D} in (14). This reveals that the reconstruction via BZAP is more stable than via ZAP in the case of noisy measurements.

4. SIMULATION RESULTS

In this section, the proposed BZAP algorithm is compared with the conventional ZAP, BOMP, and $l_{2,1}$ -opt algorithm. In all examples, the measurement matrix \mathbf{A} has 40 rows and 100 columns, with independent entries following the distribution of $\mathcal{N}(0, 1)$. The block is of the size $D = 4$. The locations of nonzero blocks in the unknown sparse signal $\bar{\mathbf{x}}$ are randomly chosen, and the values of nonzero elements are independently drawn from the Rademacher distribution.

For the BZAP algorithm, we set $\kappa = 1$, $\alpha = 1$, $\eta = 0.1$, $C_1 = 4$, and $C_2 = 1200$; Therefore it's easily checked that condition (13) in theorem 1 is always satisfied. For ZAP and $l_{2,1}$ -opt, we adopt the same stop conditions and control of step size as in the implementation of BZAP.

4.1. Recovery rate for different block sparsity

In this first experiment, the exact recovery rate for different algorithms in the noiseless case is compared. We define exact recovery when the squared deviation $\|\bar{\mathbf{x}} - \hat{\mathbf{x}}\|^2 / \|\bar{\mathbf{x}}\|^2$ is smaller than 10^{-6} . One thousand independent simulations are conducted to calculate the empirical exact reconstruction rate.

As is shown in Fig.1, the proposed BZAP algorithm outperforms all the other referenced algorithms in the experiment condition. That is, BZAP can achieve exact reconstruction of sparse signals when there are more non-zero elements: while other algorithms exactly reconstruct the signal when K is no more than 3, BZAP can achieve this when K is up to 4. ZAP gives a poor estimation because it is the only one of the algorithms that does not employ the block sparse nature of the signal.

4.2. Mean square deviation (MSD) in the presence of noise

In this experiment, the noise-contaminated measurements is formulated as in (12). The observational signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = 10 \log \left(\frac{\|\mathbf{A}\bar{\mathbf{x}}\|^2}{\|\mathbf{v}\|^2} \right). \quad (28)$$

In the simulation the SNR ranges from 10dB to 50dB. The noise vector \mathbf{v} is first generated with independent entries following the normal distribution and then re-scaled to the fit the designed SNR.

To compare the reconstruction error, the mean-square deviation associated with different algorithms is calculated, which is defined as follows:

$$\text{MSD} = \frac{\mathbf{E} \|\hat{\mathbf{x}} - \bar{\mathbf{x}}\|^2}{\mathbf{E} \|\bar{\mathbf{x}}\|^2}. \quad (29)$$

To calculate the empirical expectation in (29), we take the average of the squared norms over 10^5 independent simulations.

Regarding the MSD lower bound, consider the following oracle estimator: suppose the support T is known, then the minimum variance unbiased estimate of \mathbf{x} is the least square estimate:

$$\hat{\mathbf{x}} = \mathbf{A}_T^\dagger \mathbf{y}. \quad (30)$$

The reconstruction error is $\|\mathbf{A}_T^\dagger \mathbf{v}\|$, therefore the MSE is given by

$$\mathbf{E}(\|\hat{\mathbf{x}} - \bar{\mathbf{x}}\|^2) = \sigma^2 \text{tr}[(\mathbf{A}_T^* \mathbf{A}_T)^{-1}], \quad (31)$$

which should be lower than the achievable MSE for any practical estimators.

The simulation results are shown in Fig.2. In this experiment, the proposed BZAP algorithms again outperforms other estimators in terms of MSD, and in fact closely follows the oracle bound. The

BOMP algorithm, although guarantees higher exact recovery rate in the noiseless case, is very unstable under the perturbation of noise.

5. CONCLUSION

In this paper, we have extended the ZAP algorithm to the block-sparse problem, by introducing a cost function to approximate the $l_{2,0}$ norm of the signal. The stability of the local minimum of the cost function in BZAP is studied, which reveals an advantage of BZAP over the original ZAP by employing block sparsity of block sparse signals. Finally, simulation results show that BZAP out-performs BOMP, $l_{2,1}$ -opt and the original ZAP both in terms of the incidence of exact recovery in the noiseless case, and the mean square deviation in the noisy measurements.

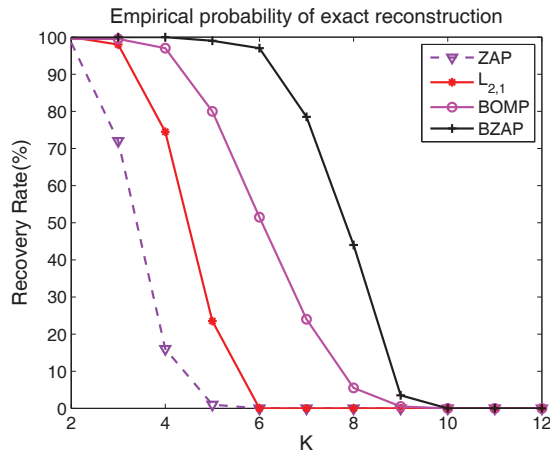


Fig. 1. Recovery of an input signal from $\mathbf{y} = \mathbf{Ax}$, where \mathbf{x} is a block sparse signal with a block sparsity level of K .

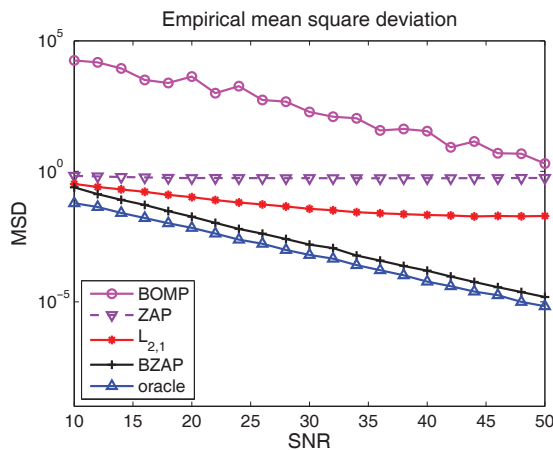


Fig. 2. Recovery of an input signal from $\mathbf{y} = \mathbf{Ax} + \mathbf{v}$, where \mathbf{x} is a block sparse signal with a block sparsity level of $K = 4$.

6. REFERENCES

- [1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, February 2006.
- [3] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRL: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182-1195, December 2007.
- [4] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressive wireless sensing," *Proc. 5th Intl. Conf. on Information Processing in Sensor Networks (IPSN' 06)*, pp.134-142, April 2006.
- [5] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031-1044, June 2010.
- [6] G. Valenzise, G. Prandi, M. Tagliasacchi, and A. Sarti, "Identification of sparse audio tampering using distributed source coding and compressive sensing techniques," *Journal on Image and Video Processing*, vol. 2009, January 2009.
- [7] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp.4203-4215, December 2005.
- [8] J. A. Tropp, and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655 - 4666, December 2007.
- [9] Y. C. Eldar, M. Mishali, "Robust Recovery of Signals From a Structured Union of Subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp.5302-5316, November 2009.
- [10] Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 2986 - 2997, August 2009.
- [11] M. Mishali and Y. C. Eldar, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993-1009, March 2009.
- [12] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, no. 1, pp. 3752, 1967.
- [13] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Compressed sensing of block-sparse signals: Uncertainty relations and efficient recovery," *CoRR*, 2009.
- [14] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409-420, April 2010.
- [15] X. Wang, Y. Gu, "Proof of convergence and performance analysis for zero-point attracting projection algorithm," submitted to *IEEE Trans. Signal Processing*.