SPARSE RECOVERY FOR DISCRETE TOMOGRAPHY

Yen-ting Lin, Antonio Ortega and Alexandros G. Dimakis

Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, California, 90089-2564, U.S.A

ABSTRACT

Discrete tomography (DT) focuses on the reconstruction of a discrete valued image from few projection angles. Prior knowledge about the image can greatly increase the quality of the reconstructed image, especially when a small number of projections are available. In this paper, we show that DT can be formulated as a sparse signal recovery problem. By using a well designed dictionary, it is possible to represent a binary image with very few coefficients. Starting from this concept, we modify the reweighed l_1 algorithm to achieve a sparse solution and preserve the binary property of image. Preliminary simulation results show that our algorithm can outperform conventional continuous reconstruction methods in cases when very limited data is available.

Index Terms— Discrete tomography, sparse signal recovery, compressive sensing

1. INTRODUCTION

Discrete Tomography (DT) requires reconstructing an image from very few projection angles, with each pixel in the image consisting of only a small number of discrete values [1]. While in many tomographic reconstruction problems the target image does not necessarily have discrete intensity values, using prior knowledge to reconstruct a discrete-valued approximation can lead to better quality, especially if the number of measurements is very limited [2]. When the number of projection angles is small, the number of unknown variables required for reconstruction far exceeds the amount of measured data. Thus, a DT reconstruction involves finding a discrete valued solution to a highly under determined linear system, for which a unique solution is unlikely to exist.

Many algorithms have been proposed to solve the DT problem. In combinatorics, researchers are interested in reconstructing a special class of images, *hv-convex* objects, with only horizontal, vertical and diagonal projection angles. Herman and Kuba gave comprehensive reviews in their books [1, 3]. Another approach is to start from continuous reconstruction methods then force the solution to converge into discrete values. Batenburg proposed a discrete algebraic reconstruction algorithm to iteratively update the object boundary [4]. Fishburn [5] and Weber [6] used linear programming (LP) relaxation to solve this NP-hard problem.

In recent years a significant amount of work has been devoted to sparse signal recovery, e.g., compressed sensing, but only limited effort has addressed scenarios where the data to be reconstructed is discrete in nature [7, 8]. In particular, no work has considered how

to take advantage of the potential sparseness of $2D$ binary images under certain transformations.

In practice, while in most cases images to be reconstructed are not binary (or even discrete valued), DT can be applied to scenarios where the target image can be segmented into high contrast regions. For example, in computed tomography (CT), researchers are interested in reconstructing images from fewer X-ray projection angles in order to reduce the radiation dose for the patients. In an angiography only two types of regions need to be considered: blood vessels enhanced by a radio-opaque contrast agent and the "background" which represents the rest of body [2]. Another example is geophysical computed tomography. The physical characteristics between different underground layers can change dramatically. In contrast to medical tomography where the measurements are taken from all viewing angles, the region of interest is scanned from very limited angles [9, 10]. Other applications can be found in different fields: chromosome analysis, transmission electron microscope [3], and experimental fluid dynamics [7].

In this paper we propose to solve the DT in the transform domain by exploiting sparseness characteristics. We expect this transform domain sparseness to be a better model for typical images of interest than spatial domain sparseness (which would involve assuming that typical images contain very few non-zero pixels). Specifically, we observe that binary images with few discontinuities along x and y axis have a sparse representation as linear combinations of unit step functions. Every binary image can then be represented in terms of basis in this dictionary multiplied by coefficient vector $\mathbf{u}, u_i \in \{0, -1, 1\}$. Note that step functions can also be applied to approximate general discrete valued images, for which a sparse approximation will lead to piecewise constant reconstruction.

For image reconstruction, we use LP to search for a sparse solution by minimizing $\|\mathbf{u}\|_1$ rather than $\|\mathbf{u}\|_0$. With LP we can also relax the integer constraint, $u_i \in \{0, -1, 1\}$, and instead find solutions such that $u_i \in [-1, 1]$. We modify the reweighed l_1 minimization algorithm [11] to encourage the solution to reach integer values and be sparse. From the property of reweighing the l_1 norm, the solution is given by the intersection of reweighed l_1 ball and the feasible region. If the solution stays in a non-integer value, we randomly change the shape of reweighed l_1 ball to make it converge to integers. Preliminary simulation results show that our method can converge to a sparse integer solution.

The paper is organized as follows. In Sections 2 and 3, we present the formulation of the discrete tomography problem and the proposed sparse reconstruction algorithm. In Section 4, we show the reconstruction result from noiseless and noisy measurements and provide a performance comparison. In Section 5 we conclude this work and discuss some further directions.

2. PROBLEM FORMULATION

In this section, we introduce the formulation of our problem. For simplicity, we focus on binary signals. Let $f(\alpha, \beta)$ be a binary 2D image and let the measurements be parallel projections along different angles. This method can be easily extended to higher dimensions and to non-binary discrete valued signals. The relationship between the projection measurements and signal f can be modeled using the Radon transform

$$P_{\theta,f}(t) = \int\limits_{-\infty}^{\infty} \int f(\alpha, \beta)\delta(\alpha \cdot \sin(\theta) + \beta \cdot \cos(\theta) - t)d\alpha d\beta \quad (1)$$

The reconstruction problem is how to recover the signal f based on a set of different viewing angles $\{\theta_i, i = 1, \ldots, m\}$. Note the Radon
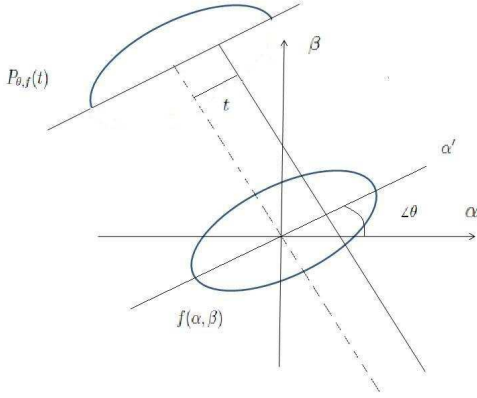

Fig. 1. Coordinate system for the Radon transform

transform is a linear transform. If the image has $p \times q$ pixels, we can reshape the 2D image f into a 1D vector $\mathbf{x}$ with dimension $n = p \cdot q$. For each projection angle, equally spaced measurements are made along the α' axis (see Figure 1). If we take γ points along the α' axis, the measured data for one specific angle will be a dimension γ vector $\mathbf{p}$. Then we can write the projection operator in matrix form:

$$\mathbf{W}_\theta \mathbf{x} = \mathbf{p}_\theta, \quad (2)$$

where $\mathbf{W}_\theta$ is a $\gamma \times n$ line projection matrix with angle θ. If we have d different viewing angles, we can represent the projection matrices and measurements as:

$$\mathbf{A}\mathbf{x} = \mathbf{y}, \quad (3)$$

with

$$\mathbf{A} = \begin{bmatrix} \mathbf{W}_{\theta_1} \\ \vdots \\ \mathbf{W}_{\theta_d} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{p}_{\theta_1} \\ \vdots \\ \mathbf{p}_{\theta_d} \end{bmatrix} \quad (4)$$
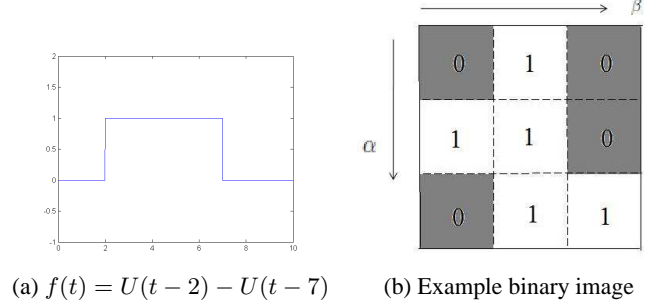
$\mathbf{A}$ will be a $m \times n$ line projection matrix that maps $\mathbf{x}$ onto measurement data $\mathbf{y}$, where $m = \gamma \cdot d$. The reconstruction problem is to, given the projection matrix $\mathbf{A}$ and data $\mathbf{y}$, with $m << n$, finding a binary solution $\mathbf{x}$, $x_i \in \{0, 1\}$. This is an integer, linear inverse problem.

Now we define the decomposition of 2D binary image f in terms of unit step functions:

$$f(\alpha, \beta) = \sum_{i=1}^{p} v(i, \beta) \cdot U(t - i), \quad \forall \beta \quad (5)$$
$$= \sum_{j=1}^{q} h(\alpha, j) \cdot U(t - j), \quad \forall \alpha \quad (6)$$

For every β, $v(i, \beta)$ decomposes $f(\alpha, \beta)$ along the α axis with unit step function $U(t)$, while $h(\alpha, j)$ decomposes $f(\alpha, \beta)$ along the β axis for each α. The unit step functions are the basis in this transform, and this transform is invertible. Because $f(\alpha, \beta)$ only has binary values $\{0, 1\}$, we can see that $v(i, \beta)$ and $h(\alpha, j)$ will also have integer values in the set $\{0, -1, 1\}$.

(a) $f(t) = U(t - 2) - U(t - 7)$ (b) Example binary image

$$\left\{ \begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{array} \right\} \qquad \left\{ \begin{array}{ccc} 0 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{array} \right\}$$

(c) Decomposition along α axis   (d) Decomposition along β axis

Fig. 2. Decomposition of binary image

Since vector $\mathbf{x}$ is obtained by reshaping image f, we can also reshape $v(i, \beta)$ and $h(\alpha, j)$ into 1D vectors $\mathbf{u}_\alpha$ and $\mathbf{u}_\beta$. We define $\mathbf{x} = \mathbf{T}_\alpha \mathbf{u}_\alpha$, a unit-step transform of the original image along the α axis, and similarly $\mathbf{x} = \mathbf{T}_\beta \mathbf{u}_\beta$ as the transform along the β axis. Using this notation, $\mathbf{T}_\alpha$ and $\mathbf{T}_\beta$ are well defined linear operators on $\mathbf{x}$ and are invertible. We can now define $\mathbf{D}$ as $\mathbf{D} = [\mathbf{T}_\alpha \mid \mathbf{T}_\beta]$, which will be an over-complete dictionary. In this dictionary we can always find $\mathbf{u}$ such that $\mathbf{x} = \mathbf{D}\mathbf{u}$ with $\mathbf{u} \in \{-1, 0, 1\}$. For example, if the image only has few "stripes" in α direction, we can choose the all the coefficients from $\mathbf{T}_\alpha$ and set the coefficients from $\mathbf{T}_\beta$ to zero. Instead of solving all possible $\mathbf{x}$, we seek to find reconstructed binary images that have a sparse representation under this dictionary $\mathbf{D}$. In order to solve the integer inverse problem with $x_i \in \{0, 1\}$, we use linear programing relaxation to search the solution $x_i \in [0, 1]$. Because the number of measurements m is much smaller than the number of variables n, this is still an under determined linear inverse problem. Now, we restate this problem as a sparse signal recovery formulation with a new measurement matrix $\widetilde{\mathbf{A}} = \mathbf{A} \cdot \mathbf{D}$:

$$\begin{aligned} \min \quad & \|\mathbf{u}\|_1 \quad (7) \\ \text{subject to} \quad & \widetilde{\mathbf{A}}\mathbf{u} = \mathbf{y}, \quad \widetilde{\mathbf{A}} = \mathbf{A} \cdot \mathbf{D} \\ & \mathbf{D}\mathbf{u} = \mathbf{x} \geq 0 \\ & \mathbf{D}\mathbf{u} = \mathbf{x} \leq 1 \end{aligned}$$

Candès and Tao [12] have shown that if the measurement matrix $\widetilde{\mathbf{A}}$ has the restricted isometry property (RIP) then it can be guaranteed that a perfect reconstruction signal $\mathbf{u}$ can obtained. However, the RIP property of the measurement matrix $\widetilde{\mathbf{A}}$ depends on the number of viewing angles and the chosen transform, and it's only a sufficient condition for perfect reconstruction. Generally speaking, we would like to choose a dictionary $\mathbf{D}$ such that the signal $\mathbf{x}$ has sparse representation and $\widetilde{\mathbf{A}}$ also has low coherence.

With the presence of noise, the measurements will be $\mathbf{y} = \widetilde{\mathbf{A}}\mathbf{u} + \mathbf{z}$, $\mathbf{z}$ is the noise with $\|\mathbf{z}\|_2 \leq \epsilon$. In order to operate with noisy measurements we change the constraints to:

$$\begin{aligned} \min \quad & \|\mathbf{u}\|_1 \\ \text{subject to} \quad & \|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\|_2 \leq \delta \\ & \mathbf{Du} \geq \mathbf{0} \\ & \mathbf{Du} \leq \mathbf{1} \end{aligned} \tag{8}$$

where δ is the search allowance for noisy measurements. This approach is closely related to LASSO [13] and sparse Bayesian learning [14], where the goal is to achieve a good trade-off between data fitting and model complexity. Without knowing the noise power ϵ, selecting an appropriate δ is non-trivial. If we choose δ to be too large, then the hyper ball $\left\|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\right\|_2$ will contain the origin $\mathbf{u} = \mathbf{0}$. In that case, $\|\mathbf{n}\|_1 = 0$ is the global minimum and the solution will be a zero vector. On the other hand, if δ is too small, the intersection of hyper ball $\left\|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\right\|_2$ and $\mathbf{0} \leq \mathbf{Du} \leq \mathbf{1}$ may be empty. Boufounos et al [15] have proposed a cross validation method to estimate noise power. With a good choice of δ, we can assume that the hyper ball $\left\|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\right\|_2$ contain the actual solution $\mathbf{u}_0$ we want to recover. But it is also possible that some other point $\widetilde{\mathbf{u}_0}$ inside the hyper ball will have lower $\|\widetilde{\mathbf{u}_0}\|_1$. Donoho et al [16] proved that under certain conditions for $\widetilde{\mathbf{A}}$ the noisy reconstructed solution $\widetilde{\mathbf{u}_0}$ is at least as sparse as the ideal solution $\mathbf{u}_0$, and the atoms in $\widetilde{\mathbf{u}_0}$ belong to $\mathbf{u}_0$.

3. PROPOSED ALGORITHM

We modify the reweighed l_1 minimization algorithm proposed by Candès [11]. We choose $\mathbf{u}_0 = \widetilde{\mathbf{A}}'\mathbf{y}$ as our initial starting point for LP and iteratively search for a sparse solution. In each iteration, this method minimizes the weighted l_1 ball $\min_{\mathbf{u}} \sum_i w_i |u_i|$, which more closely approximates the l_0 norm. By reweighing we change the shape of l_1 ball and make it "sharper" in some dimensions. But in our case the linear region constraint will cut off the feasible region. It is possible that the solution will be trapped in these intersection points (see figure 3). In order to jump off these lo-
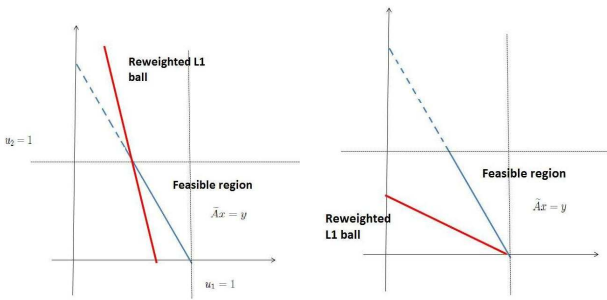
Fig. 3. Non-integer solution case: Because the feasible region is cut off by the boundary condition, the intersection of l_1 ball and feasible region falls in the middle. After changing the shape of l_1 ball, an integer valued solution can be reached.

cal non-integer value, every 5 iterations we check the solution to find out the non-integer dimensions u_i and reset their weights w_i with a

random value. This results in random changes to the shape of the l_1 ball, making it possible to find the correct vertices.

For the noisy case we are searching the solution in $\left\|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\right\|_2 \leq \delta$ region, therefore, even if we have the correct shape of reweighed l_1 ball the solution will not be exactly all integer. It will always touch the inner bound of the region (See Figure 4). In our simulations, we use $\sum_i x_i \cdot (1 - x_i)$ [6] as the metric for non-integer solution. This metric will be 0 if $x_i \in \{0, 1\}$. Our experimental results show that our algorithm tends to find solution with more integer values, while non-randomized method stop after a few iterations providing solutions with fewer integer values.
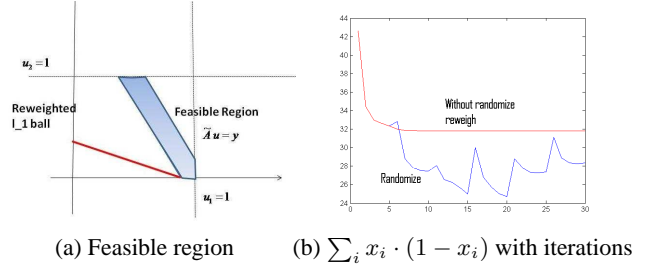
(a) Feasible region (b) $\sum_i x_i \cdot (1 - x_i)$ with iterations

Fig. 4. The solution for noisy case: The reweighed l_1 ball will always touch the inner bound.

Algorithm 1 Reweighed l_1 norm minimization for noisy measurement

1: Choose the allowance error δ in observed data and stop criteria ρ.
2: Define the maximal iteration number l_{max} and set the iteration number $l = 0$. Initialize the weight $\mathbf{w}_i^{(0)} = 1$, $\forall i = 1, \ldots, n$ and decide Δ. Use $\mathbf{u}^0 = \widetilde{\mathbf{A}}'\mathbf{y}$ as the starting point for LP.
3: Solve the reweighed l_1 norm linear programming problem with starting point $\mathbf{u}^{(l-1)}$.

$$\begin{aligned} \mathbf{u}^{(l)} = \quad & \arg\min \left\|\mathbf{W}^{(l)}\mathbf{u}\right\|_1 \\ \text{subject to} \quad & \left\|\mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u}\right\|_2 \leq \delta \\ & \mathbf{Du} = \mathbf{x} \geq \mathbf{0} \\ & \mathbf{Du} = \mathbf{x} \leq \mathbf{1} \end{aligned} \tag{9}$$

4: Update the weights: for each $i = 1, \ldots, n$

$$\mathbf{w}_i^{l+1} = \frac{1}{|\mathbf{u}_i^l| + \Delta} \tag{10}$$

5: Reset the weights for non-integer solutions: Every 5 iterations, we reset the w_i corresponding to non-integer u_i by a random number that is uniformly distributed between $[0, 2/\Delta]$

$$\mathbf{w}_i^{l+1} \leftarrow U(0, 2/\Delta) \tag{11}$$

6: Terminate the iteration if x converges $\left\|\mathbf{u}^{l+1} - \mathbf{u}^l\right\|_2 \leq \rho$ or reach the maximum number of iteration. Otherwise, increase l and go to step 3.

4. EXPERIMENTAL RESULT

We use CVX [17] in order to solve the LP. The test binary image is a simplified Sheep-Logan phantom and has size 32×32 pixels. The measurements are taken with 49 equally spaced projections for each angle, and 5 different projection angles uniformly between $[0, 180]$ are selected. The measuring matrix $\widetilde{A}$ will be a 245×1024 matrix. In order to test the noisy measurement, we added to the measurement data white Gaussian noise with zero mean and variance $\epsilon = \{0, 0.1, 0.25, 0.5, 1\}$. We choose the search allowance δ to be equal to the noise variance. For comparison, we list the mean square error achieved by Weber [6] and by a filtered back-projection reconstruction.

Our algorithm is very sensitive to the noise level. As we showed in Figure 4, the solution is always given by the inner bound of the region $\left\| \mathbf{y} - \widetilde{\mathbf{A}}\mathbf{u} \right\|_2 \leq \delta$. Choose a larger δ will always push the solution $\mathbf{u}$ closer to $\mathbf{0}$. From the simulation results, we notice that the reweighed l_1 minimization does lead to solutions taking boundary values $\mathbf{x}_i \in \{0, 1\}$. For the reconstruction, it shows our method performs better in low noise cases. A possible explanation is that, for high noise levels, the feasible region $\left\| \mathbf{y} - \tilde{\mathbf{A}}\mathbf{u} \right\| \leq \delta$ has been increased, and solution converges closer to $\mathbf{0}$. Therefore for higher noise level, we find fewer supports that satisfy the conditions of [16].

5. CONCLUSIONS

We have presented a new sparse reconstruction formulation for discrete tomography, which focuses on reconstruction of binary images that have a sparse representation using a well designed dictionary. Our algorithm uses LP to relax the integer solution condition and also search the the sparse representation. We introduce a randomized reweighed l_1 minimization to enhance convergence to a binary solution.

Future work will be focus on the relationship between the designed dictionary $\mathbf{D}$ and the corresponding measuring matrix $\widetilde{\mathbf{A}} = \mathbf{A} \cdot \mathbf{D}$. For example, we can always increase the size of dictionary to obtain a sparser representation, but this will also increase the coherence for the measuring matrix. Another possible direction is to use the discrete value property to enhance signal recovery from noisy measurements.

Fig. 5. The testing phantom

Fig. 6. Reconstruction with noise $\rho = 0.1$

Fig. 7. S. Weber method with noise $\rho = 0.1$

Fig. 8. Filtered back-projection result with noise $\rho = 0.1$

Fig. 9. Mean square error with different noise level

6. REFERENCES

[1] G.T. Herman and A. Kuba, *Discrete Tomography: Foundations, Algorithms, and Applications*, Birkhauser, 1999.

[2] GT Herman and A. Kuba, "Discrete tomography in medical imaging," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1612–1626, 2003.

[3] G.T. Herman and A. Kuba, *Advances in discrete tomography and its applications*, Citeseer, 2007.

[4] KJ Batenburg and J. Sijbers, "DART: A fast heuristic algebraic reconstruction algorithm for discrete tomography," in *Proceedings of the IEEE International Conference on Image Processing, San Antonio, Texas*, 2007.

[5] P. Fishburn, P. Schwander, L. Shepp, and RJ Vanderbei, "The discrete Radon transform and its approximate inversion via linear programming," *Discrete Applied Mathematics*, vol. 75, no. 1, pp. 39–61, 1997.

[6] S. Weber, C. Schnorr, T. Schule, and J. Hornegger, "Binary tomography by iterating linear programs," *COMPUTATIONAL IMAGING AND VISION*, vol. 31, pp. 183, 2006.

[7] S. Petra, A. Schroder, B. Wieneke, and C. Schnorr, "On sparsity maximization in tomographic particle image reconstruction," *Lecture Notes in Computer Science*, vol. 5096, pp. 294–303, 2008.

[8] V. Stankovic, L. Stankovic, and S. Cheng, "Compressive sampling of binary images," in *CISP08 Congress on Image and Signal Processing*.

[9] KA Dines and RJ Lytle, "Computerized geophysical tomography," *Proceedings of the IEEE*, vol. 67, no. 7, pp. 1065–1073, 1979.
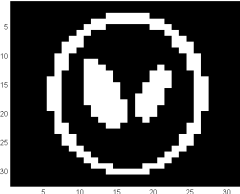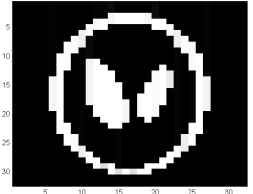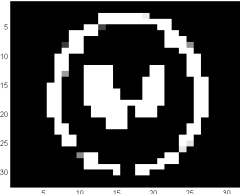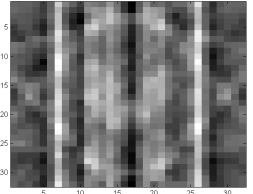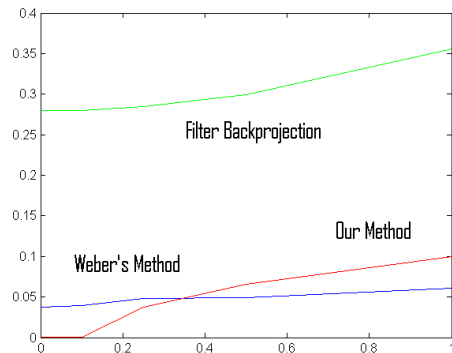
[10] YT Lin and Antonio Ortega, "Reconstruction algorithm for high contrast velocity travel time tomography," in *accepted by the IEEE International Conference on Acoustics, Speech, and Signal Processing , Dallas, Texas*, 2010.

[11] E.J. Candes, M.B. Wakin, and S.P. Boyd, "Enhancing sparsity by reweighted l 1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[12] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[13] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[14] D.P. Wipf and B.D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[15] P. Boufounos, M. Duarte, and R. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *IEEE Workshop on Statistical Signal Processing*. Citeseer, 2007, pp. 299–303.

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

[17] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," *avialable at http://www. stanford. edu/boyd/cvx*, vol. 1.