An Iterative Convex Optimization Solver with Side Information for Joint-Sparse Signal Recovery

Shih-Wei Hu, Sung-Hsien Hsieh, and Chun-Shien Lu*
Institute of Information Science, Academia Sinica,
*Email: lcs@iis.sinica.edu.tw

Gang-Xuan Lin
Department of Mathematics,
National Cheng-Kung University, Tainan, Taiwan

***Abstract*—We propose an iterative joint-sparse signal recovery algorithm for compressive sensing with multiple measurement vectors. Our algorithm is an iterative procedure to solve $\ell_{2,1}$-norm minimization with side information that is self-produced from the reconstructed signals at previous iterations instead of being given in advance. Our side information is designed with a theoretical foundation to further reduce the required number of measurements for successful recovery.**

I. Introduction

In the context of compressive sensing (CS) [1]–[3], multiple measurement vectors (MMVs) [4]–[7] has become a popular issue, where more than one signals are sensed by the same sensing matrix and these signals have joint-sparsity. Let $S \in \mathbb{R}^{n\times l}$ be the unknown matrix of l (> 1) original signals to be sensed by a sensing matrix $\Phi \in \mathbb{R}^{m\times n}$ ($m < n$), and let the matrix of measurement vectors be $Y \in \mathbb{R}^{m\times l}$, where $Y = \Phi S$. Suppose there exists an orthonormal basis Ψ such that $S = \Psi\bar{X}$, where $\bar{X} \in \mathbb{R}^{n\times l}$ is k joint-sparse and the ground truth. Given a dictionary $A = \Phi\Psi$ and observation Y, the joint-sparse signal recovery problem can be solved efficiently via convex minimization as:

$$\text{(ML1)} \quad \min_X \|X\|_{2,1} \quad \text{s.t.} \quad Y = AX.$$

We say the problem (ML1) succeeds if the unique optimal solution of (ML1) is the ground truth, $\bar{X}$.

A. Related Work

Wang and Yin [8] proposed iterative support detection via ℓ_1-minimization under the SMV model. At each iteration, it is first to estimate the support set via thresholding the recovered signal at the previous iteration. Then, the estimated support set treated as side information is fed into weighted ℓ_1-minimization for signal recovery. Chen and Huo [4] discovered that the required number of measurements in the MMVs problem is related to not only the sparsity but also the rank of observation Y. Motivated by [4], Davies and Eldar [6] proposed another greedy algorithm, called RA-ORMP. In [9], we studied when problem (ML1) with prior information succeeds and derive the phase transition of success rate inspired by [10], [11].

B. Contributions

We propose an iterative convex solver that explores the side information, which is self-produced from the output of solving (ML1). Compared with the traditional solver [8] based on weighted ℓ_1-minimization, the side information in our solver can include not only the support set but also the signal types. Especially, the side information originated from signal types is potential to let the number m of measurements be smaller than sparsity k (*i.e.*, $m < k$). This improvement transcends the performance limit of greedy algorithms with $m = k$.

II. Proposed Algorithm: MIC

We present an **M**MVs **I**terative **C**onvex solver (MIC). Let $\tilde{X}_i$ and W_i be the results at i^{th} iteration. Our algorithm is depicted as follows:

1) **Input**: $\tilde{X}_0 = \mathbf{0}$, Y, A, k, c_1, ϵ, λ, and $i = 0$. **Output**: $\tilde{X}$.
2) **Side information prediction**:
$$W_i = \begin{cases} \mathbf{0}, & \text{if } i = 0, \\ c_1 \times \text{sign}(\hat{X}_i), & \text{otherwise}, \end{cases} \quad (1)$$
where c_1 is used to enhance the strength of side information and $\hat{X}_i = \arg\min_X \|X - \tilde{X}_i\|_F$ *s.t.* $|\text{supp}(X)| = k$.
3) **Signal reconstruction**: Let $\tilde{X}_{i+1} = \arg\min_X \|X\|_{2,1} + \lambda\|X - W_i\|_{2,1}$ *s.t.* $Y = AX$.
4) **Stopping criterion**: Output $\tilde{X} = \tilde{X}_{i+1}$ if $\|\tilde{X}_{i+1} - \tilde{X}_i\|_F \le \epsilon$; otherwise, $i = i + 1$ and go back to Step 2.

The algorithm involves a major component that is **side information prediction**. By Theorem 3.4 in [9], W_i and c_1 in Eq. (1) are designed to decrease the required number of measurements for recovery. However, at the first iteration, since we don't have any side information about the original signal in advance, the initial matrix W_0 is set to a zero matrix. In other words, the first side information is predicted by the problem (ML1).

If (ML1) model fails to reconstruct the ground truth, we may assume $\tilde{X}_i = \bar{X} + Z$, where $Z \in \text{null}(A, l) = \{Z | AZ = \mathbf{0}\}$, because $\bar{X} + Z$ is an optimal point of the problem (ML1). On the other hand, given a Gaussian random matrix A, spark$(A) = m + 1$[1] almost surely implies $|\text{supp}(Z)| \ge m$. Moreover, since the energy of Z will distribute uniformly on at least m rows, it may not be sufficient to influence the values of $\bar{X}$ when $\bar{X}$ is sparse enough or m is not so small. Therefore, the best k-term approximation of $\tilde{X}_i$ still can identify the support set of $\bar{X}$.

III. Simulation

We compare MIC with RA-ORMP [6] and (ML1) with two types of signals: (1) non-zero entries drawn from ± 1 with equal probability and (2) Gaussian distribution. The verification procedure below was repeated 100 times for each set of parameters, composed of m and k, under $l = 5$, $n = 100$, and stopping criterion $\epsilon = 10^{-3}$.

1) Construct $\bar{X}$ according to one of two signal types.
2) Draw a standard normal matrix $A \in \mathbb{R}^{m\times n}$ to sample signals.
3) Run the proposed MIC algorithm to output $\tilde{X}$.
4) Declare success if $\|\bar{X} - \tilde{X}\|_F \le \epsilon$.

In Figs. 1 and 2, we show the results with $l = 5$ for random signal and binary signal, respectively. It is found that MIC outperforms RA-ORMP because MIC can achieve the performance limit $m = k$ under small l while RA-ORMP requires larger l to utilize the information of rank. MIC also improves (ML1) remarkably.

[1] Spark$(A) = \min |I|$, where $I \in \mathcal{I} = \{I | A_I v = \mathbf{0}$ and $v \ne \mathbf{0}, I \subseteq 2^{[n]}/\{\emptyset\}.\}$, A_I is a submatrix of A, and $[n] = \{1, 2, ..., n\}$.

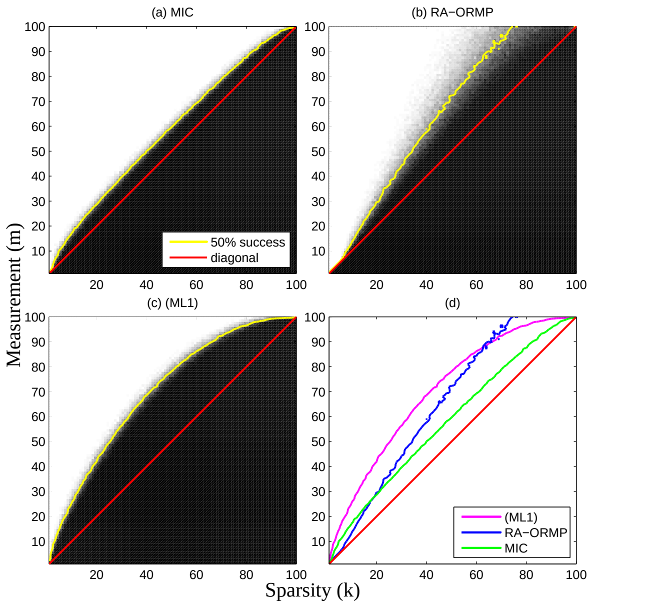

Fig. 1. Phase transition diagram of successful probability for random signal. (a) Proposed method MIC. (b) RA-ORMP. (c) Problem (ML1). (d) Comparison of phase transition curves achieving 50% successful probability. (Best viewed on a color display.)

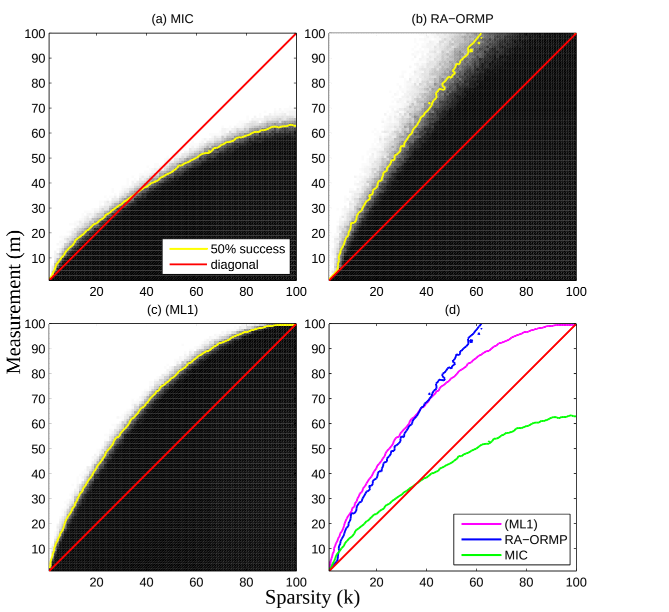

Fig. 2. Same setting in Fig. 1 but for binary signal.

REFERENCES

[1] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

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

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

[4] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4634–4643, Dec 2006.

[5] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation: Part i: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.

[6] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1135–1146, Feb 2012.

[7] S.-W. Hu, G.-X. Lin, S.-H. Hsieh, and C.-S. Lu, "Phase transition of joint-sparse recovery from multiple measurements via convex optimization," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, April 2015, pp. 3576–3580.

[8] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 462–491, 2010.

[9] S. W. Hu, G. X. Lin, S. H. Hsieh, W. J. Liang, and C. S. Lu, "Performance analysis of joint-sparse recovery from multiple measurement vectors with prior information via convex optimization," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 4368–4372.

[10] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: phase transitions in convex programs with random data," *Information and Inference*, vol. 3, no. 3, pp. 224–294, 2014.

[11] J. F. C. Mota, N. Deligiannis, and M. R. D. Rodrigues, "Compressed sensing with prior information: Optimal strategies, geometry, and bounds," *CoRR*, vol. abs/1408.5250, 2014. [Online]. Available: http://arxiv.org/abs/1408.5250