Maximin Analysis of Message Passing Algorithms for Recovering Block Sparse Signals

Armeen Taeb  
University of Colorado at Boulder  
armeen.taeb@colorado.edu

Arian Maleki  
Columbia University  
arian@stat.columbia.edu

Christoph Studer  
Rice University  
studer@rice.edu

Richard G. Baraniuk  
Rice University  
richb@rice.edu

Abstract—We consider the problem of recovering a block (or group) sparse signal from an underdetermined set of random linear measurements, which appear in compressed sensing applications such as radar and imaging. Recent results of Donoho, Johnstone, and Montanari have shown that approximate message passing (AMP) in combination with Stein’s shrinkage outperforms group LASSO for large block sizes. In this paper, we prove that for a fixed block size and in the strong undersampling regime (i.e., having very few measurements compared to the ambient dimension), AMP cannot improve upon group LASSO, thereby complementing the results of Donoho et al.

I. PROBLEM STATEMENT

We analyze the recovery of a block (or group) sparse signal \( \mathbf{x} \in \mathbb{R}^N \) with at most \( k \) nonzero entries from an underdetermined set of linear measurements \( \mathbf{y} = \mathbf{A}\mathbf{x} \), where \( \mathbf{A} \in \mathbb{R}^{n \times N} \) is i.i.d zero-mean Gaussian with unit variance. We consider the asymptotic setting where \( \delta = n/N \) and \( \rho = k/n \) represent the undersampling and sparsity parameters, respectively, and \( N, n, k \to \infty \); we furthermore assume that all blocks are of the same size \( B \). The signal \( \mathbf{x} \) is partitioned into \( M \) blocks with \( N = MB \). In what follows, we will denote a particular block by \( \mathbf{x}_B \). Suppose that the elements of \( \mathbf{x}_B \) are drawn from a distribution \( F(\mathbf{x}_B) = (1 - \epsilon)\delta_0(\|\mathbf{x}_B\|_2) + \epsilon G(\mathbf{x}_B) \), where \( \epsilon = \rho \delta \), and \( \delta_0 \) is the Dirac delta function; \( G \) is a probability distribution that is typically unknown in practice. We define the block soft-thresholding function as follows [1]:

\[
\eta^B(\mathbf{y}; \tau) \equiv \mathbf{y}_B/\|\mathbf{y}_B\|_2 \max\left\{\|\mathbf{y}_B\|_2 - \tau, 0\right\}.
\]

Two popular algorithms for recovering block sparse signals in compressed sensing are group LASSO [2] and approximate message passing (AMP) [3]. Group LASSO searches for a vector \( \mathbf{x} \) that solves \( \mathbf{x}^* = \arg\min_{\mathbf{x}} \left( \sum_{B = 1}^{M} \|\mathbf{x}_B\|_2 : \mathbf{y} = \mathbf{A}\mathbf{x} \right) \). AMP is an iterative algorithm for computing \( \mathbf{x} \). Concretely, AMP is initialized by \( \mathbf{x}_t^0 = 0 \) and \( \mathbf{c}_t^0 = 0 \), and iteratively performs the following computations [4]:

\[
\mathbf{x}^{t+1} = \eta^B(\mathbf{x}^t + A^*\mathbf{c}^t) \quad \text{and} \quad \mathbf{c}^t = \mathbf{y} - A\mathbf{x}^t + \varepsilon^t.
\]

Here, \( \mathbf{c} \) is a correction term that depends on the previous iterations to significantly improve the convergence of AMP; \( \mathbf{x}^t \) is the sparse estimate at iteration \( t \), and \( \eta \) is a nonlinear function that imposes (block) sparsity.

The performance of compressed sensing recovery algorithms can be characterized accurately by their phase-transition (PT) behavior. Specifically, we define a two-dimensional phase space \( (\delta, \rho) \in [0, 1] \) that is partitioned into two regions: “success” and “failure”, with these regions separated by the PT curve \( (\delta, \rho(\delta)) \). For the same value of \( \delta \), algorithms with higher PT outperform algorithms with lower PT, i.e., guarantee the exact recovery for more nonzero entries \( k \).

II. MAIN RESULTS

The thresholding function \( \eta^B \) determines the performance of AMP. Indeed, different choices of \( \eta^B \) may lead to fundamentally different performance behaviors. If \( \eta^\text{std} \) from (1) is used, then the performance of AMP is equivalent to that of group LASSO [1], [3]. However, \( \eta^\text{std} \) is not necessarily optimal for block (or group) sparse signals, and finding the optimal thresholding function is of significant practical interest. In this case, we are interested in optimal thresholding functions that satisfy the following maximin criterion:

\[
\rho^*(\delta) \equiv \sup_{\eta^B} \rho^B(\delta, G).
\]

Such an \( \{\eta^B\}_{i = 1}^\infty \) provides the best PT performance for the least favorable distribution—a reasonable assumption, since the distribution \( G \) is typically unknown in most practical applications.

Our first contribution characterizes the behavior of the optimal PT \( \rho^*(\delta) \) of AMP in the strong undersampling regime.

Theorem 1: The optimal PT of AMP follows the behavior

\[
\rho^*(\delta) \sim \frac{B}{2\log(1/\delta)} \quad \text{for} \quad \delta \to 0.
\]

Our second contribution confirms that the optimal PT \( \rho^*(\delta) \) of AMP coincides with that of group LASSO. Consequently, block soft-thresholding as in (1) is optimal for strong undersampling.

Theorem 2: The PT of group LASSO follows the behavior

\[
\rho^B(\delta) \sim \frac{B}{2\log(1/\delta)} \quad \text{for} \quad \delta \to 0.
\]

The combination of Theorems 1 and 2 reveals that, for a fixed block size \( B \) and \( \delta \to 0 \), the optimal PT \( \rho^*(\delta) \) of AMP coincides with the PT \( \rho^B(\delta) \) of group LASSO. This result has two implications in the strong undersampling regime and for fixed block sizes: (i) Block soft-thresholding is optimal and (ii) AMP in combination with Stein’s shrinkage as in [1] cannot outperform group LASSO.

We emphasize that our observations do not contradict the results of [1], which show that AMP outperforms group LASSO for large block sizes. By contrast, Theorems 1 and 2 combined with [1, Section 3.2] show that as \( \delta \) decreases, AMP with Stein’s shrinkage requires larger block sizes to outperform group LASSO. In fact, we conclude that under strong undersampling and for fixed block sizes, no thresholding function can enable AMP to outperform group LASSO with respect to the phase transition behavior.

REFERENCES


