ON THE BENEFITS OF THE BLOCK-SPARSITY STRUCTURE IN SPARSE SIGNAL RECOVERY

Hwanjoon (Eddy) Kwon and Bhaskar D. Rao

Department of Electrical and Computer Engineering,
University of California at San Diego, La Jolla, CA 92093-0407, USA

ABSTRACT

We study the problem of support recovery of block-sparse signals, where nonzero entries occur in clusters, via random noisy measurements. By drawing analogy between the problem of block-sparse signal recovery and the problem of communication over Gaussian multi-input and single-output multiple access channel, we derive the sufficient and necessary condition under which exact support recovery is possible. Based on the results, we show that block-sparse signals can reduce the number of measurements required for exact support recovery, by at least '$1/(\text{block size})$', compared to conventional or scalar-sparse signals. The minimum gain is guaranteed by increased signal to noise power ratio (SNR) and reduced effective number of entries (i.e., not individual elements but blocks) that are dominant at low SNR and at high SNR, respectively. When the correlation between the elements of each nonzero block is low, a larger gain than '$1/(\text{block size})$' is expected due to, so called, diversity effect, especially in the moderate and low SNR regime.

Index Terms— Support recovery, Block-sparse signals, MISO-MAC channel capacity

1. INTRODUCTION

The probem of sparse signal recovery has recently received much attention and involves the estimation of a sparse signal $\mathbf{X} \in \mathbb{R}^m$ in high dimension with a small number of nonzero entries, via linear measurements $\mathbf{Y} = A\mathbf{X} + \mathbf{Z}$, where $A \in \mathbb{R}^{n\times m}$ is referred to as the measurement matrix and $\mathbf{Z}$ is the measurement noise. The goal is to reconstruct the signal $\mathbf{X}$ from as few number of menasurements as possible.

In many applications, it is important to find the exact support of the sparse signal [1], [2]. In the noiseless environment (i.e., $\mathbf{Z} = 0$), sufficient conditions to exactly recover the support of the sparse signal have been derived in [3]-[5]. In the presence of measurement noise, information theoretic tools have proven useful in understanding the performance tradeoff for support recovery of sparse signals [6]-[10]. In particular, Jin et al. [10] identified the connection between the problem of sparse signal support recovery and the problem of communication over Gaussian multiple access channel (MAC), based on which they derived sharper asymptotic tradeoffs between the signal dimension, the number of nonzero entries, and the number of measurements for exact support recovery in the noisy setting.

In some applications, the nonzero entries of sparse signals often take place in clusters [11]-[12]. Such signals are referred to as block-sparse [13]-[15]. To elaborate, this model is a good approximation to the signal model in applications like EEG/MEG where the brain activities are in localized regions rather than at a single point. This model is also consistent with communication channel modeling where an ideal sparse channel consisting of a few specular multi-path components has a discrete time, bandlimited, baseband representation which exhibits a block sparse structure with the block centers determined by the arbitrary arrival times of the multi-path components. Consequently, understanding the limits of support recovery for block structured sparse signals has important practical ramifications. Under the noiseless assumption, Stojnic et al. [13] presented sufficient conditions under which the solution to minimize the sum of l_2-norms of each block finds the sparsest solution. Eldar et al. [14] studied conditions on the measurement matrix ensuring block-sparse signals can be recovered through various optimization techniques in the noiseless environment. On the other hand, Ben-Haim and Eldar [15] examined the ability of greedy algorithms to estimate a block sparse parameter vector from noisy measurements.

In this paper, we provide the asymptotic sufficient and necessary condition under which the exact support recovery of the block-sparse signal is possible in a noisy environment. Our focus is on the case where the block sizes are the same for all the blocks of a sparse signal. We find that $n = \log(m)/c(\mathbf{X})$ is sufficient and necessary for the exact support recovery. We provide a complete characterization of $c(\mathbf{X})$ that depends on the number of nonzero blocks and nonzero values of each block, based on which we discuss how much gain a known block structure can provide in terms of the minimum number of measurements for accurate support recovery. We first show that the block-sparsity can reduce the minimum number of measurements, by at least $\frac{1}{\text{block size}}$, compared to conventional or scalar-sparse signals with the same nonzero values. There is a twofold reason why $\frac{1}{\text{block size}}$ is guaranteed: increased SNR and reduced number of effective entries (i.e., not individual elements but blocks). It is shown that the former mainly plays at low SNR while the latter at high SNR. In addition, so called, diversity effect is discussed through which the block-sparsity can further reduce the minimum number of measurements, especially in the regime of moderate and low SNR. Our work can be viewed as a generalization of the work in [10] in the sense that we interpret our problem as the problem of communications over Gaussian multi-input and single-output (MISO) MAC.

The rest of this paper is organized as follows. In Section 2, we formulate the block-sparse signal model and the problem of support recovery. In Section 3, we introduce an interpretation of the problem from the perspective of information theory. Our main results and their implications are presented in Section 4. Finally, we conclude our work in Section 5.

2. SIGNAL MODEL AND PROBLEM FORMULATION

The overall model is the same as the standard sparse signal recovery problem except for the signal $\mathbf{X} = [X_1, \cdots, X_m]^{\mathrm{T}} \in \mathbb{R}^m$. In-

This research was supported by NSF grant CCF-1144258.

formally, the signal is divided into m_b blocks, each of size b, i.e., $\mathbf{X} = [\mathbf{X}_1^{\mathrm{T}}, \cdots, \mathbf{X}_{m_b}^{\mathrm{T}}]^{\mathrm{T}}$, where $\mathbf{X}_i$ is a vector of b entries. Block sparsity refers to the fact that most of the blocks are zero except for a few nonzero blocks. A more formal definition is required to derive the results and is provided next.

We are given the total number of nonzero elements k, the block size b, the total number of blocks m_b(=m/b), the number of nonzero blocks k_b(=k/b), and a vector of the k nonzero values $\mathbf{w} = [w_1, \cdots, w_k]^{\mathrm{T}} \in \mathbb{R}^k$. We assume m_b and k_b are integers. Define i-th block of $\mathbf{w}$ as $\mathbf{w}_i = [w_{(i-1)b+1}, \cdots, w_{ib}]^{\mathrm{T}}$. Generate $\mathbf{B} = [B_1, \cdots, B_{k_b}]^{\mathrm{T}}$ such that $B_1, \cdots, B_{k_b}$ are chosen uniformly at random from $\{1, 2, \cdots m_b\}$ without replacement. (For notational simplicity, let $[m_b]$ denote $\{1, 2, \cdots m_b\}$.) Then, the signal of interest $\mathbf{X} = \mathbf{X}(\mathbf{w}, \mathbf{B}, b)$ generated on block basis such as

$$\mathbf{X}_i = \begin{cases} \mathbf{w}_j & \text{if } i = B_j \\ \mathbf{0}^{\mathrm{T}} & \text{if } i \notin \{B_1, \cdots, B_{k_b}\} \end{cases} \tag{1}$$

where $\mathbf{0}$ is a zero vector.

We measure $\mathbf{X}$ through the linear operation

$$\mathbf{Y} = A\mathbf{X} + \mathbf{Z} \tag{2}$$

where $A \in \mathbb{R}^{n\times m}$ is the measurement matrix, $\mathbf{Z} \in \mathbb{R}^n$ is the measurement noise, and $\mathbf{Y} \in \mathbb{R}^n$ is the noisy measurement. We further assume that the elements of A are independently generated according to $\mathcal{N}(0,1)$ and the noise Z_i are independently and identically distributed (i.i.d.) according to $\mathcal{N}(0, \sigma_z^2)$. SNR of i-th element and SNR of j-th block are defined as X_i^2/σ_z^2 and $\|\mathbf{X}_j\|^2/\sigma_z^2$, respectively.

Upon observing the noisy measurement $\mathbf{Y}$, the goal is to recover the support of $\mathbf{X}$. Throughout this paper, we assume k and b are known but $\mathbf{w}$ are unknown. In effect, the problem is equivalent to the recovery of $\mathbf{B}$ since the support of $\mathbf{X}$ is determined by ($\mathbf{B}$, b). The performance metric is the average probability error in support recovery. Note that the probability here is taken over the random $\mathbf{B}$, A, and $\mathbf{Z}$.

3. AN INFORMATION THEORETIC PERSPECTIVE ON BLOCK-SPARSE SIGNAL RECOVERY

Jin et al. [10] revealed the similarities between the standard sparse recovery problem and the communication problem over MAC, where multiple senders simultaneously transmit information, so called, codewords, to a common receiver which tries to correctly detect the codewords of each sender. As an extension, this paper identifies the connection between the support recovery of the block-spare signal and the communication over MISO MAC, which gives a fresh insight into our problem and allows us to use various results developed in channel coding theorem relevant to the MISO MAC.

First, we briefly review the problem of communication over Gaussian MISO MAC. Suppose l senders wish to transmit information to a common receiver. Each sender i has N_t transmit antennas and the receiver has a single receive antenna. Each sender i has access to a codebook $\mathcal{C}^{(i)} = \{C_1^{(i)}, C_2^{(i)}, \cdots C_{m^{(i)}}^{(i)}\}$, where $C_j^{(i)} \in \mathbb{R}^{n\times N_t}$ is a MISO codeword and $m^{(i)}$ is the number of codewords in $\mathcal{C}^{(i)}$. The rate for the sender i, $R^{(i)} \triangleq \frac{\log m^{(i)}}{n}$. To transmit information, each sender chooses a codeword from its codebook. Let q_i be the codeword index chosen by sender i. Then, the received signal $\mathbf{Y} \in \mathbb{R}^n$ at the receiver is given by

$$\mathbf{Y} = C_{q_1}^{(1)}\mathbf{h}_1 + C_{q_2}^{(2)}\mathbf{h}_2 + \cdots + C_{q_l}^{(l)}\mathbf{h}_l + \mathbf{Z} \tag{3}$$

where $\mathbf{h}_i \in \mathbb{R}^{N_t}$ is the MISO channel gain associated with sender i and $\mathbf{Z} \in \mathbb{R}^n$ is the noise with Z_j i.i.d. according to $\mathcal{N}(0, \sigma_z^2)$.

Upon receiving $\mathbf{Y}$, the receiver determines the codewords transmitted by each sender. Since the senders interfere with each other, there is an inherent tradeoff among their operating rates. The notion of capacity region is introduced to capture this tradeoff by characterizing all possible rate tuples $\mathcal{R} \triangleq (R^{(1)}, R^{(2)}, \cdots, R^{(l)})$ at which reliable communication can be achieved with diminishing error probability. Under the assumption that MISO channel gain $\mathbf{h}_i$ is unknown to each sender and each sender obeys the power constraint $\|C_j^{(i)}\|_F^2/n < 1$, the capacity region of the Gaussian MISO MAC is given by

$$\left\{\mathcal{R} : \sum_{i\in\mathcal{T}} R^{(i)} \le \frac{1}{2}\log\left(1 + \frac{1}{N_t\sigma_z^2}\sum_{j\in\mathcal{T}}\|\mathbf{h}_i\|^2\right), \forall\mathcal{T} \subseteq [l]\right\}. \tag{4}$$

Indeed, $\mathcal{R}$ is achieved via the codebooks of which elements are independently generated according to $\mathcal{N}(0, 1/N_t)$. Recall that each element of our measurement matrix A is generated according to $\mathcal{N}(0,1)$. This implies that when we think of our problem as the MISO MAC communication problem as discussed below, the total power of each sender increases proportionally to the block size b, as opposed to the standard MISO MAC.

Next, consider how our problem can be mapped to the above communication problem. The measurement model (2) has an alternative form

$$\mathbf{Y} = A_{B_1}\mathbf{w}_1 + A_{B_2}\mathbf{w}_2 + \cdots + A_{B_{k_b}}\mathbf{w}_{k_b} + \mathbf{Z} \tag{5}$$

where $A_i \in \mathbb{R}^{n\times b}$ is i-th block of columns of A corresponding to $\mathbf{X}_i$. In contrast to (3), (5) can be viewed as the received signal at the common receiver equipped with a single receive antenna, signals from k_b senders, each equipped with b transmit antennas, where sender i chooses codeword index B_i and transmits codeword A_{B_i} over MISO channel $\mathbf{w}_i$. Based on the relations, we see that support recovery of the block-sparse signal is equivalent to detection of the MISO codewords so that the performance limit of the former can be derived from the capacity region of the equivalent MISO MAC. One important restriction in our problem is that the codebook A is common to all the senders, resulting in the same rate. Consequently, the capacity region of interest is defined in one dimensional space such as $R = \frac{\log m}{n} < c(\mathbf{w}, b)$ at which the exact codeword detection is possible, where $c(\mathbf{w}, b)$ can be derived as below, from (4) with $R^{(1)} = R^{(2)} = \cdots = R^{(k_b)}$, and in the consideration of the increased transmit power by a factor of b.

$$c(\mathbf{w}, b) = \min_{\mathcal{T}\subseteq[k_b]}\left[\frac{1}{2|\mathcal{T}|}\log\left(1 + \frac{1}{\sigma_z^2}\sum_{j\in\mathcal{T}}\|\mathbf{w}_j\|^2\right)\right]. \tag{6}$$

This quantity can be understood as a symmetric capacity of the MISO MAC with total transmit power of each sender boosted by a factor of b. The symmetric capacity [19] indicates the maximum rate at which all the senders can communicate reliably.

To make a long story short, we can infer $\frac{\log m}{n} < c(\mathbf{w}, b)$ is a sufficient condition for the accurate support recovery of the block sparse signal. This turns out to be the true sufficient condition in the next section.

4. MAIN RESULTS AND THEIR IMPLICATIONS

In this section, we first present the sufficient and necessary condition under which exact support recovery of block-sparse signals is possi-

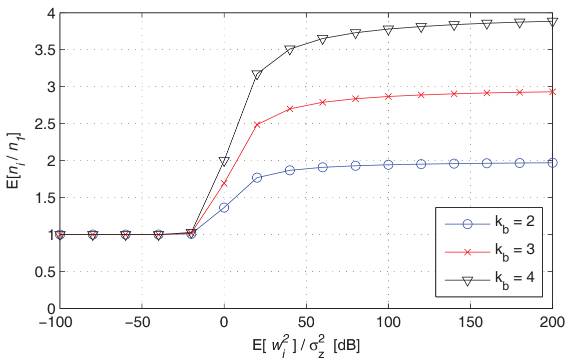

Fig. 1. Increase in the minimum number of measurements for the exact support recovery compared to the case when the number of nonzero block $k_b = 1$. (The block size is fixed to 2. Each nonzero element is uniformly chosen at random from $\{\pm 1\}$. n_i denotes the minimum number of measurements for exact support recovery when $k_b = i$.)

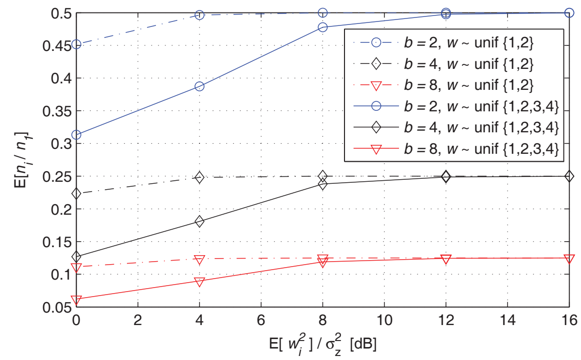

Fig. 2. Reduction in the minimum number of measurements for the exact support recovery compared to the scalar-sparse signal due to the block structure with block size b of 2, 4, and 8. (The total number of nonzero elements is fixed to 8. 'w∼unif$\{1, 2\}$' indicates w_i are uniformly chosen at random from $\{\pm 1, \pm 2\}$. n_i denotes the minimum number of measurements for exact support recovery when $b = i$.)

ble[1]. We then discuss how the minimum number of measurements for exact support recovery is affected by different parameters in various scenarios. In addition, we present a set of numerical results which help us understand our problem in more depth.

4.1. Sufficient and necessary condition

Theorem 1 (*Sufficient and necessary condition*): For given k, $\mathbf{w}$, and b that are bounded above, as $m \to \infty$ the sufficient and necessary condition under which reliable support recovery of the block-sparse signal is possible with diminishing error probability is

$$n = \frac{\log m}{c(\mathbf{w}, b)} \tag{7}$$

where $c(\mathbf{w}, b)$ is given by (6).

Note that the constant c($\mathbf{w}$, b) is explicitly characterized by the number of nonzero blocks k/b and their SNR $\|\mathbf{w}_j\|^2/\sigma_z^2$.

4.2. Implications of the theorem

Theorem 1 indicates that the minimum number of measurements required for the exact support recovery is inversely proportional to the symmetric capacity, $c(\mathbf{w}, b)$. How can a block structured signal increase the symmetric capacity in the equivalent MISO MAC, or, decrease the minimum number of measurements for the exact support recovery, compared to the scalar-sparse signal with the same k? As discussed in the previous section, the block size $b > 1$ leads to a boost in transmit power for each sender by a factor of b in the equivalent MISO MAC problem and therefore causes SNR to be raised by the same amount, which certainly augments the symmetric capacity. Another factor is the reduced effective number of entries by a factor of $1/b$ due to the fact that the support of signal is recovered not on individual element basis but on block basis. Corollary 1 shows how the two factors play dissimilarly in different SNR regime, where a scenario different from our problem setup is considered to individually assess the impact of the two factors on reducing the minimum number of measurements. This evaluation is interesting in itself, but the results also shed light on understanding why the minimum required number of measurements is decreased by at least $\frac{1}{\text{block size}}$, which will be subsequently discussed.

Corollary 1 (*Dominant parameter*): Suppose sparse signals $\mathbf{X}^{(1)} \in \mathbb{R}^m$ and $\mathbf{X}^{(2)} \in \mathbb{R}^m$ are generated according to our signal model, under the assumption that $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ have k_1 and k_2 nonzero blocks, respectively, with the same block size b, and with the constant l_2-norm w for each nonzero block. Note that the total number of nonzero elements are different between the two signals. Let n_1 and n_2 be the minimum required number of measurements for exact support recovery of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, respectively. Let $\gamma = w^2/\sigma_z^2$ (SNR of each block). Then,

$$n_2/n_1 \approx k_2/k_1 \text{ for high } \gamma \text{ and} \tag{8}$$

$$n_1 \approx n_2 \propto 1/\gamma \text{ for low } \gamma. \tag{9}$$

This can be shown using the property of the symmetric capacity (6) that $\mathcal{T}_1 = [k_1]$ and $\mathcal{T}_2 = [k_2]$ minimize (6) when $\|\mathbf{w}_j\|$ is constant for all j, based on which we have $n_i = k_i \log m / \log(1 + k_i\gamma)$, which, in turn, leads to the results. Corollary 1 indicates that when block SNR γ is high, the minimum number of measurements is mainly determined by the number of nonzero blocks, while it is mostly affected by γ, not by the number of nonzero blocks, when γ is low. Note that this observation is irrelevant to the block size and therefore is still valid for scalar-sparse signals. Figure 1 shows a set of relevant numerical results, where b is fixed to 2 and each nonzero element is uniformly selected at random from $\{\pm 1\}$. The curves represents the average ratio of the minimum number of measurements when $k_b = i$, i =2, 3, and 4, denoted by n_i to the minimum number of measurements when $k_b = 1$, denoted by n_1. To elaborate, take a close look at the curve with $k_b = 2$. At low SNR, E$[n_2/n_1]$ is close to one, which indicates that even though the number of nonzero blocks are increased from 1 to 2, the number of measurements to detect the two position indices in $\mathbf{X}$ is still similar to that in the absence of the second. On the other hand, in the regime of intermediate SNR, E$[n_2/n_1]$ increases as SNR grows, converging to '2' at high SNR that is the number of total nonzero blocks. The above observations agree with (9) and (10).

[1]The proof of the theorem is omitted due to space limitations. It can be derived by expanding on the proofs provided in [10].

Now, let us focus on the discussion on how much a block structure in sparse signals can reduce the number of measurements needed for precise support recovery. For the succeeding results, we assume sparse signals $\mathbf{X}^{(1)} \in \mathbb{R}^m$ and $\mathbf{X}^{(2)} \in \mathbb{R}^m$ are generated according to our signal model with the same number of nonzero elements k, the same nonzero values $\mathbf{w}$, the block size b_1 for $\mathbf{X}^{(1)}$, and the block size b_2 for $\mathbf{X}^{(2)}$. Let $k_1(= k/b_1)$ and $k_2(= k/b_2)$ indicate the number of nonzero blocks of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, respectively, where they are assumed to be integers. Let n_1 and n_2 be the minimum required number of measurements for exact support recovery of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, respectively.

Corollary 2 (*Minimum gain from block-sparsity*): For any k_1, k_2, and $\mathbf{w}$ that are bounded above, if k_2 is a multiple of k_1, then

$$\frac{n_2}{n_1} \leq \frac{b_1}{b_2}. \tag{10}$$

This can be proved using the property of (6) that for any $\mathcal{T}_2 \subseteq [k_2]$, the term $1/(2|\mathcal{T}_2|)\log\big(1 + \frac{1}{\sigma_z^2}\sum_{j\in\mathcal{T}_2}\|\mathbf{w}_j\|^2\big)$ in (6) for $\mathbf{X}^{(2)}$ appears in (6) for $\mathbf{X}^{(1)}$, multiplied by b_2/b_1. We also use the inequality, $\min[a_1, a_2] \leq 1/\alpha \min[\alpha a_1, \alpha a_2, a_3, a_4], \alpha, \forall a_i \in \mathbb{R}$. Note that if k_2 is not a multiple of k_1, (10) still holds for most $\mathbf{w}$, but, there exists some extraordinary $\mathbf{w}$ with which (10) is not true, e.g., $[100000,\ 0.00001,...]^{\mathrm{T}}$. When $\mathbf{X}^{(1)}$ is scalar-sparse ($b_1 = 1$), we have $n_2 \leq (1/b_2)n_1$, i.e., the block-sparse signal can reduce the minimum number of measurements for exact support recovery, by at least $\frac{1}{\text{block size}}$, compared to the scalar-sparse signals. This is thanks to the composite effect of the increased SNR and the reduced effective number of entries. According to Corollary 1, the former mainly plays at low SNR while the latter at high SNR.

Corollary 3 (*Constant magnitude induces the minimum gain*): If $|w_j|$ is constant for all j, then $n_2/n_1 = b_1/b_2$.

This can be shown using the property of (6) that $\mathcal{T}_1 = [k_1]$ and $\mathcal{T}_2 = [k_2]$ minimize (5) when $|w_j|$ is constant. The results imply that the block-sparsity gain is small when all the nonzero elements are similar in magnitude. On the other hand, the block-sparsity gain can be much larger than $\frac{1}{\text{block size}}$ if the correlation in magnitude between the elements of each nonzero block is low. For example, in the case when $\mathbf{w} = [0.25, -5, -0.15, 5]^{\mathrm{T}}$ and $\sigma_z^2 = 0.2$, the scalar-sparse signals require 26 and 52 times more measurements for exact support recovery than the block-sparse signals of block size 2 and the block-sparse signals of block size 4, respectively. This large gap is due to, so called, 'diversity' - for a scalar sparse signal, when an element has a much smaller magnitude than the others, it makes the symmetric capacity (6) small and consequently requires a large number of measurements. On the other hand, when the support is detected on cluster basis, the above problem can be significantly alleviated unless all the elements in each cluster have small magnitudes. Then, one may have a relevant question, "Once the correlation in magnitude is small, is the diversity gain substantial?" The following corollary answers this question.

Corollary 4 (*Gain in the low noise power regime*): For any $\mathbf{w}$ bounded above, when σ_z^2 is low, $n_2/n_1 \approx b_1/b_2$.

As $\sigma_z^2 \to 0$, for any $\mathbf{w}$, $\mathcal{T}_1 = [k_1]$ and $\mathcal{T}_2 = [k_2]$ minimize (6), based on which we can show n_2/n_1 tends to b_1/b_2. The results indicate that the block-sparsity can reduce the minimum number of measurements by no more than $\frac{1}{\text{block size}}$ when the noise power is small, regardless of the correlation in magnitude between the elements of each nonzero block.

The numerical results in Figure 2 illustrate the reduction in the minimum number of measurements for exact support recovery due to the block-sparsity, compared to the scalar-sparse signal, where three block sizes $b = 2$, 4, 8 and two different probability distributions for random w_i are examined with 8 nonzero elements fixed for all the cases. It is observed that at high SNR, the gain is equal to the minimum gain $\frac{1}{\text{block size}}$ regardless of distribution of w_i, while at low and moderate SNR, the gain is larger than the minimum gain due to the diversity effect, proportional to the variance of w_i.

5. CONCLUSION

In this paper, we showed that a known block structure in sparse signals can provide a significant gain in the minimum number of measurements for accurate support recovery. To derive the results, we used the analytical framework that relies on the connection between the support recovery of block-sparse signals and communications over the Gaussian MISO multiple access channel. For future work, it would be interesting to consider the block-sparse signal recovery with multiple measurement vectors as well as the support recovery of blocks-sparse signals with irregular block sizes.

6. REFERENCES

[1] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, pp. 14-30, 2001.

[2] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," *IEEE ICASSP*, pp. 1357-1360, 2007.

[3] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution," *Comm. Pure Appl. Math*, vol. 59, pp. 797-829, 2004.

[4] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203-4215, 2005.

[5] J. A. Tropp, "Greedy is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, 2004.

[6] D. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presense of noise," *IEEE Trans. Inf. Theory*, 2006.

[7] M. Wainwright, "Sharp thresholds for high-dimensional and noisy sparsity recovery using l_1-constrained quadratic programming (lasso)," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2183-2202, 2009.

[8] M. Wainwright, "Information-theoretic limits on sparsity recovery in the highdimensional and noisy setting," *IEEE Trans. Inf. Theo.*, Dec 2009.

[9] A. K. Fletcher, S. Rangan, and V. K. Goyal, "Necessary and sufficient conditions for sparsity pattern recovery," *IEEE Trans. Inf. Theory*, Dec 2009.

[10] Y. Jin, Y. Kim, and B. D. Rao, "Performance tradeoffs for exact support recovery of sparse signals," *Proc. IEEE ISIT*, 2010.

[11] I. f. gorodnitsky, J. S. George, and B. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *j. Electroencephalogr. Clin. Neurophysiol.*, 1995.

[12] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi, "Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays," *IEEE J. Sel. Topics Signal Process.*, Jun. 2008.

[13] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the Reconstruction of Block-Sparse Signals with an Optimal Number of Measurements," *IEEE Trans. Sig. Process.*, Aug. 2009.

[14] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-Sparse Signals: Uncertainty Relations and Efficient Recovery," *IEEE Trans. Sig. Process.*, June 2010.

[15] Z. Ben-Haim and Y. C. Eldar, "Near-Oracle Performance of Greedy Block-Sparse Estimation Techniques From Noisy Measurements," *IEEE J. Sel. Topics Signal Process.*, Sep. 2011.

[16] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[17] A. Lapidoth, "Nearest neighbor decoding for additive non-gaussian noise channels," *IEEE Trans. Inf. Theory*, 1996.

[18] T. cover and J. Thomas, *Elements of Information Theory*, Wiley, 2006.

[19] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.