Weighted Block Sparse Bayesian Learning for Basis Selection

Ahmed Al Hilli, and Athina Petropulu

Department of Electrical and Computer Engineering, Rutgers University
The work was supported by NSF under Grant ECCS-1408437

Introduction

- In block sparse signals groups of entries are active simultaneously.
- Block sparse signal recovery can be formulated as finding the signal with the minimum number of active groups that describes the observation.
- In general, this is a complex problem. A relaxation is to find the signal with the smallest sum of group energies.
- Bayesian approaches have also been proposed for group sparse problems by generalizing Sparse Bayesian Learning (SBL).
- In this paper, a Weighted Block Sparse Bayesian Learning is proposed for sparse vector recovery.

Block Sparse Bayesian Learning (BSBL)

- Consider the system $\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}$, $\mathbf{n} \sim N(0, \sigma^2\mathbf{I})$, $\mathbf{G} \in R^{m\times n}$, and $\mathbf{x} \in R^n$.
- Assume each block $\mathbf{x}_i \in R^{d_i\times 1}$ in $\mathbf{x}$ follows a parametrized multivariate Gaussian distribution, i.e.,

$$p(\mathbf{x}_i; g_i, \mathbf{B}_i) \sim \mathcal{N}(0, g_i\mathbf{B}_i), \quad (1)$$

where $g_i \geq 0$ controls the block sparsity of $\mathbf{x}$.
- Assuming independence between the blocks, $p(\mathbf{x})$ can be written as $p(\mathbf{x}) \sim \mathcal{N}(0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\Sigma}_0 = \text{diag}\{g_1\mathbf{B}_1, ..., g_m\mathbf{B}_m\}$.
- The posterior of $\mathbf{x}$ is [1]

$$p(\mathbf{x}; \mathbf{y}, \sigma^2, \{g_i, \mathbf{B}_i\}_{i=1}^m) = \mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x), \quad (2)$$

where

$$\boldsymbol{\mu}_x = \boldsymbol{\Sigma}_0\mathbf{G}^T(\sigma^2\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma}_0\mathbf{G}^T)^{-1}\mathbf{y}, \quad (3)$$

and

$$\boldsymbol{\Sigma}_x = (\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2}\mathbf{G}^T\mathbf{G})^{-1}. \quad (4)$$

- Given the parameters σ^2 and $\{g_i, \mathbf{B}_i\}_{i=1}^m$, the Maximum a Posteriori (MAP) estimate of $\mathbf{x}$ is $\hat{\mathbf{x}} = \boldsymbol{\mu}_x$.
- The parameters can be estimated by minimizing

$$L(\sigma^2, \{g_i, \mathbf{B}_i\}_{i=1}^m) = \log|\sigma^2\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma}_0\mathbf{G}^T| + \mathbf{y}^T(\sigma^2\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma}_0\mathbf{G}^T)^{-1}\mathbf{y}. \quad (5)$$

- Differentiating L w.r.t. g_i, σ^2, and $\mathbf{B}_i$, and equating to zero we get

$$g_i = \frac{1}{d_i}\text{Tr}[\mathbf{B}_i^{-1}(\boldsymbol{\Sigma}_x^i + \boldsymbol{\mu}_x^i(\boldsymbol{\mu}_x^i)^T)], \; i = 1, 2, ..., m, \quad (6)$$

$$\sigma^2 = \frac{\|\mathbf{y} - \mathbf{G}\boldsymbol{\mu}_x\|_2 + \text{Tr}[\boldsymbol{\Sigma}_x\mathbf{G}^T\mathbf{G}]}{M}, \quad (7)$$

Block Sparse Bayesian Learning (BSBL)

- Differentiating L w.r.t. $\mathbf{B}_i$, and equating to zero we get

$$\mathbf{B}_i = \frac{1}{m}\sum_{i=1}^{m}\frac{\boldsymbol{\Sigma}_x^i + \boldsymbol{\mu}_x^i(\boldsymbol{\mu}_x^i)^T}{g_i}, \quad (8)$$

where $\boldsymbol{\mu}_x^i$ is the i^{th} block in $\boldsymbol{\mu}_x$, $\boldsymbol{\Sigma}_x^i$ is the corresponding i^{th} principal diagonal block in $\boldsymbol{\Sigma}_x$, and d_i is the length of the i^{th} block.

The Proposed Weighted Block Sparse Bayesian (WBSBL) Approach

- Consider $\alpha_i = \frac{1}{g_i} \sim$ **Gamma**(a_i, b_i)
- Using a Type II maximum likelihood procedure as in BSBL, the cost function to be minimized is

$$L(\sigma^2, \{g_i, \mathbf{B}_i\}_{i=1}^m) = \log|\sigma^2\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma}_0\mathbf{G}^T| + \mathbf{y}^T(\sigma^2\mathbf{I} + \mathbf{G}\boldsymbol{\Sigma}_0\mathbf{G}^T)^{-1} + 2\sum_{i=1}^{m}\frac{b_i}{g_i} + 2\sum_{i=1}^{m}a_i\log(g_i). \quad (9)$$

- Differentiating w.r.t. g_i, σ^2, and $\mathbf{B}_i$, we get

$$g_i = \frac{\text{Tr}[\mathbf{B}_i^{-1}(\boldsymbol{\Sigma}_x^i + \boldsymbol{\mu}_x^i(\boldsymbol{\mu}_x^i)^T)] + 2b_i}{d_i + 2a_i} \quad (10)$$

and σ^2 and $\mathbf{B}_i$ are as described in (7) and (8), respectively.
- Suppose we have access to a weight vector $\mathbf{w}$, which contains large values corresponding to active $\mathbf{x}_i$ blocks, and low values corresponding to non active blocks in $\mathbf{x}$.
- Set $a_i = \frac{1}{w_i}$ and $b_i = w_i$. Assuming that $w_i \neq 0$, update the rule for g_i as

$$g_i = \frac{\text{Tr}[\mathbf{B}_i^{-1}(\boldsymbol{\Sigma}_x^i + \boldsymbol{\mu}_x^i(\boldsymbol{\mu}_x^i)^T)^T] + 2w_i}{d_i + 2/w_i}. \quad (11)$$

Simulation Results

- 1000 Monte Carlo simulations are performed. In each trial, a matrix $\mathbf{G}$ with entries following $N(0, 1)$ is constructed.
- k indices are randomly selected as the locations of the non-zero active blocks.
- The values of the non-zero entries in all the blocks are taken from $N(5, 0.25)$.
- Additive white Gaussian noise is added to $\mathbf{Gx}$ with SNR equal to 5 dB, 2 dB, and 0 dB.

Simulation Results

- MUSIC based on 100 snapshots is used to construct the weighting vector $\mathbf{w}$.

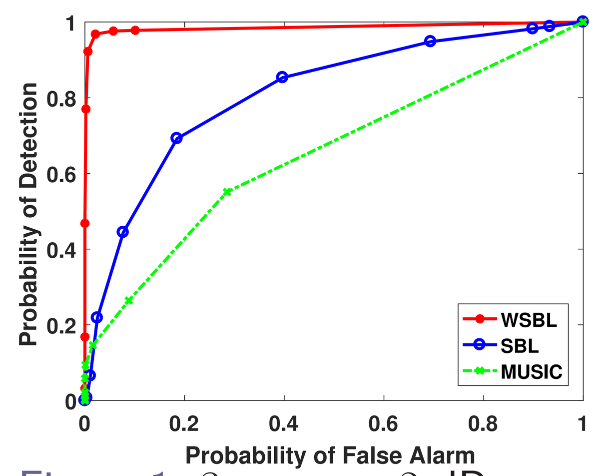

Figure 1: 3 sources, 2 dB SNR, block size of 2

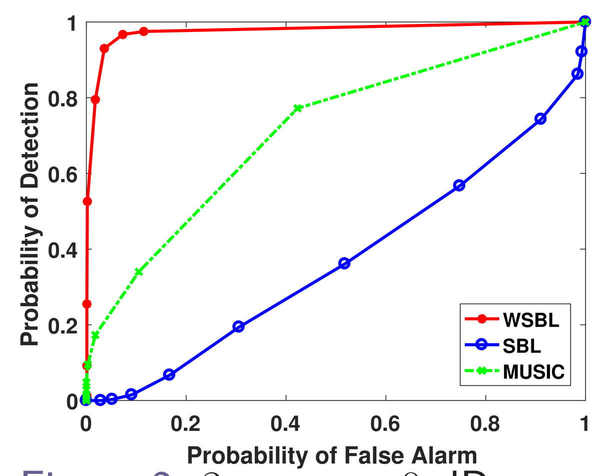

Figure 2: 3 sources, 0 dB SNR, block size of 2

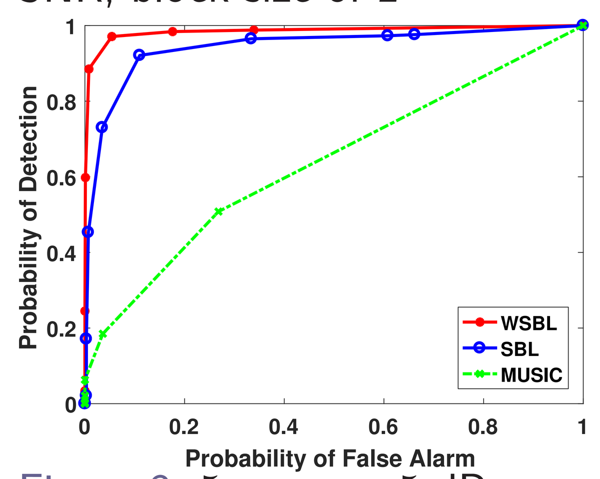

Figure 3: 5 sources, 5 dB SNR, block size of 2

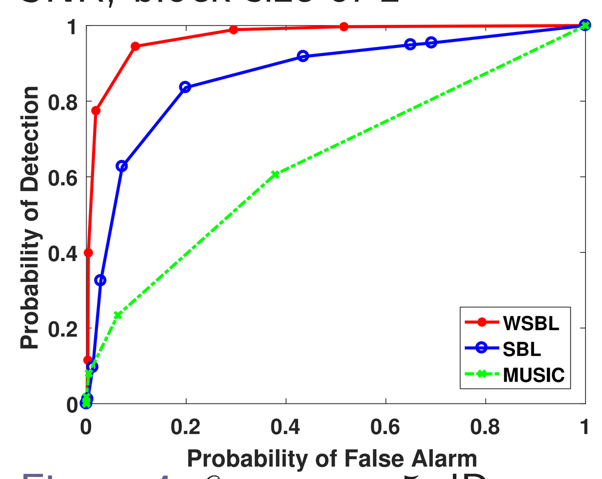

Figure 4: 6 sources, 5 dB SNR, block size of 2

Conclusion

- Weighted Block Sparse Bayesian Learning approach has been proposed, which assigns distinct variance priors to each block, giving some hyperparameters more importance over the others.
- The importance of a specific parameter is obtained based on a rough estimate of the underlying block sparse vector, obtained via a methods that does encourage sparsity.
- Simulations have shown significant improvement in terms of probability of detection and false alarm, especially at low SNR scenarios.
- WBSBL degrades slower as the number of active block increased, as compared to BSBL.

References

- [1] Zhang, Zhilin, and Bhaskar D. Rao. "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation." IEEE Transactions on Signal Processing 61.8 (2013): 2009-2015.