GENERAL TOTAL VARIATION REGULARIZED SPARSE BAYESIAN LEARNING FOR ROBUST BLOCK-SPARSE SIGNAL RECOVERY

Aditya Sant, Markus Leinonen† and Bhaskar D. Rao**

*Department of Electrical and Computer Engineering, University of California San Diego
†Centre for Wireless Communications – Radio Technologies, University of Oulu, Finland

ABSTRACT

Block-sparse signal recovery without knowledge of block sizes and boundaries, such as those encountered in multi-antenna mmWave channel models, is a hard problem for compressed sensing (CS) algorithms. We propose a novel Sparse Bayesian Learning (SBL) method for block-sparse recovery based on popular CS based regularizers with the function input variable related to total variation (TV). Contrary to conventional approaches that impose the regularization on the signal components, we regularize the SBL hyperparameters. This iterative TV-regularized SBL algorithm employs a majorization-minimization approach and reduces each iteration to a convex optimization problem, enabling a flexible choice of numerical solvers. The numerical results illustrate that the TV-regularized SBL algorithm is robust to the nature of the block structure and able to recover signals with both block-patterned and isolated components, proving useful for various signal recovery systems.

***Index Terms*—** Compressed Sensing, Block-sparsity, Sparse Bayesian Learning, Total Variation, Majorization-minimization.

1. INTRODUCTION

Block-sparse signal recovery has various applications in wireless communication, audio, and image processing. We are primarily interested in such signal recovery for mmWave channel estimation where the received signal is composed of angular multipath components that impinge on the receive antenna as clustered rays [1, 2]. One of the main challenges with block-sparse recovery is to model inter-element dependency, in addition to the sparsity constraint. Since block sizes can be unequal and block boundaries are unknown, the number of possible signal "blocks" involved in the search grows exponentially with the grid size. Hence, there is a need to impose structure on signal recovery algorithms. This also presents an inherent trade-off between computational complexity and block-sparse modeling for arbitrarily sized blocks with unknown boundaries.

Considering known block partitions, compressed sensing (CS) [3,4] algorithms have been modified for block-sparse signals, which include Group-Lasso [5], Group Basis Pursuit [6], Model-based CoSaMP [7], and Block-OMP [8]. Early attempts for block-sparse recovery under unknown block partitions include Struct-OMP in [9] and the method based on graphical models in [10]. The works [11,12] were the first SBL approaches and developed the Block SBL (BSBL) algorithm. Using Bayesian CS, [13] incorporated a spike-and-slab prior to model both block and individual sparsity.

The Pattern-coupled SBL (PC-SBL) [14] enforced block-sparse structures by coupling the underlying SBL parameters. The Non-uniform Burst Sparsity algorithm proposed in [15] improved on [14] through Variational Bayesian Inference. Coupled priors were also used in the Extended-BSBL (EBSBL) method in [16]. The works in [17, 18] enforce block structures using a cluster-structured prior.

We propose a novel total variation (TV) based regularizer for SBL to promote block-sparse signal recovery. Specifically, we enforce block sparsity in the hyperparameter space of SBL to promote uninterrupted zero regions of the estimated signal. This is achieved using common regularizers from CS acting on the hyperparameter TV input variable, instead of signal components. The framework is quite general and allows for an exploration of a wide range of regularizers utilizing the experience from CS. To the best of our knowledge, this is the first work to apply a TV type penalty in the hyperparameter space of SBL to encourage block-sparsity. Majorization-minimization is used to convexify the proposed SBL formulation and develop an iterative algorithm. Numerical results show that by inducing a soft TV prior on the parameters, the TV-regularized SBL method is robust to sparsity structure; the algorithm attains definitive recovery from strict block-sparsity to fully random sparsity.

2. NEW TOTAL VARIATION REGULARIZERS FOR BLOCK-SPARSE SIGNAL RECOVERY VIA SBL

We consider a multiple measurement vector (MMV) problem which involves simultaneous estimation of L block-sparse source vectors $\mathbf{x}_l \in \mathbb{C}^N$ from a collection of noisy linear measurements.

$$\mathbf{y}_l = \mathbf{A}\mathbf{x}_l + \mathbf{n}_l,\ l = 1, \ldots, L, \tag{1}$$

where $\mathbf{y}_l \in \mathbb{C}^M$ is a measurement vector at time instant l, $\mathbf{A} \in \mathbb{C}^{M\times N}$ is a fixed known measurement matrix, and $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0}, \lambda\mathbf{I})$ is a noise vector, independent of $\mathbf{x}_l$. Source vectors and noise vectors are assumed to be independent and identically distributed (i.i.d.) across the time instants. The same sparsity pattern is shared among the collection of vectors $\{\mathbf{x}_l\}_{l=1}^L$. Thus, the signal ensemble $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_L]$ is *block-row-sparse*. We assume that both the block sizes and their locations are *unknown*.

The early works imposed a block structure on SBL inference through a specific deterministic or stochastic parameterization of the signal; for example, BSBL [12] relies on a pre-determined block partition. Whereas, to handle dynamic block sizes, algorithms like PC-SBL [14,15] impose *explicit coupling* on the variables. Our approach improves on one limitation of such coupling based approaches: reduced sensitivity to isolated spurious components. We first provide a brief overview of the SBL framework for sparse signal recovery.

The work of A. Sant and B. D. Rao has been financially supported by ONR Grant No. N00014-18-1-2038 and the UCSD Center for Wireless Communications. The work of M. Leinonen has been financially supported in part by Walter Ahlström Foundation through Tutkijat Maailmalle program, Infotech Oulu, the Academy of Finland (grant 323698) and (grant 319485), and Academy of Finland 6Genesis Flagship (grant 318927).

2.1. SBL Framework: Generalized Cost Function

There are various advantages of SBL for MMV sparse signal recovery, motivating our choice: *(i)* The M-SBL [19] parameter estimation abstracts each row of $\mathbf{X}$ by a single parameter (γ_i), reducing the number of parameters to be estimated from NL to N compared to CS approaches; *(ii)* It falls under the class of methods that are correlation-aware which have shown superior ability to find sparse solutions [20]; *(iii)* SBL shows great promise for sparse signal recovery under correlated sources and ill-conditioned dictionaries [21].

We now describe the SBL inference. With an additive Gaussian noise model (1), the SBL framework [19] assumes a parametric Gaussian distribution for each signal $\mathbf{x}_l \in \mathbb{C}^N$ ($l \in \{1, \ldots, L\}$) as $p(\mathbf{x}_l; \boldsymbol{\gamma}) = \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Gamma}|}} \exp\left(-\frac{1}{2}\mathbf{x}_l^{\mathrm{H}} \boldsymbol{\Gamma}^{-1} \mathbf{x}_l\right)$, where $\boldsymbol{\gamma} = [\gamma_1 \cdots \gamma_N]^{\mathrm{T}} \in \mathbb{R}_+^N$ is a vector of *hyperparameters*, adjusting the variance of each signal component $x_{l,i}$, $i = 1, \ldots, N$, and $\boldsymbol{\Gamma} \triangleq \mathrm{diag}(\boldsymbol{\gamma})$. The hyperparameter values $\boldsymbol{\gamma}$ reflect the sparsity profile of the block-row-sparse $\mathbf{X}$; a suitable prior on $\boldsymbol{\gamma}$ can lead $\mathbf{x}_l$ to model many interesting sparse priors, e.g., Gaussian scale mixtures.

The posterior density $p(\mathbf{x}_l|\mathbf{y}_l; \boldsymbol{\gamma})$ is also Gaussian as $\mathcal{CN}(\boldsymbol{\mu}_{\mathbf{x}_l|\mathbf{y}_l;\boldsymbol{\gamma}}, \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y};\boldsymbol{\gamma}})$, where

$$\boldsymbol{\mu}_{\mathbf{x}_l|\mathbf{y}_l;\boldsymbol{\gamma}} = \lambda^{-1} \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y};\boldsymbol{\gamma}} \mathbf{A}^{\mathrm{H}} \mathbf{y}_l, \; \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y};\boldsymbol{\gamma}} = \left(\lambda^{-1} \mathbf{A}^{\mathrm{H}} \mathbf{A} + \boldsymbol{\Gamma}^{-1}\right)^{-1}. \tag{2}$$

For a given $\boldsymbol{\gamma}$, the estimate of each signal $\{\mathbf{x}_l\}_{l=1}^L$ is formed as $\hat{\mathbf{x}}_{l,\mathrm{SBL}} = \boldsymbol{\mu}_{\mathbf{x}_l|\mathbf{y}_l;\boldsymbol{\gamma}}$ according to (2). Following [19], the hyperparameter estimation is done through Type-II maximum a posteriori (MAP) estimation of the posterior $p(\boldsymbol{\gamma}|\mathbf{y}_1, \ldots, \mathbf{y}_L)$ over $\boldsymbol{\gamma}$, i.e.,

$$\boldsymbol{\gamma}^* = \underset{\boldsymbol{\gamma} \succeq \mathbf{0}}{\mathrm{argmin}} \; L \log |\boldsymbol{\Sigma}_{\mathbf{y}}| + \textstyle\sum_{l=1}^L \mathbf{y}_l^{\mathrm{H}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}_l - \log p(\boldsymbol{\gamma}), \tag{3}$$

where $\boldsymbol{\Sigma}_{\mathbf{y}} = \lambda \mathbf{I} + \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^{\mathrm{H}}$ is the measurement covariance matrix and $\log p(\boldsymbol{\gamma})$ is the *hyperprior* on $\boldsymbol{\gamma}$. The expression in (3) is the generalized MMV SBL cost function. This optimization is *non-convex* due to the concave term $\log |\boldsymbol{\Sigma}_{\mathbf{y}}|$; convexity of $\log p(\boldsymbol{\gamma})$ depends on the prior. We elaborate on the optimization strategies in Sec. 3.

2.2. SBL with Novel TV-based Regularizers

Most existing approaches enforce structure by working directly on $\mathbf{x}_l$, a more challenging and less efficient approach for the complex block-sparsity problem. *Informative* priors/regularizers $\log p(\boldsymbol{\gamma})$ in (3) can help improve inference [22–24]. Although block-sparse methods have been developed using SBL, the priors are often strong, thereby biasing the methods and making them brittle. Surprisingly, some simple regularizers seem to have been overlooked and we show them to be quite effective and, more importantly, robust.

To support block-sparse solutions, we opt for a regularizer that combines various sparse regularizers developed in CS, with Total Variation (TV) [25–27]. To this end, we denote the hyperprior as $\beta\, T(\boldsymbol{\gamma}) \triangleq -\log p(\boldsymbol{\gamma})$, where β is a non-negative weighting parameter and $T(\cdot)$ is a general TV-type penalty of vector $\boldsymbol{\gamma}$. Regardless of $T(\cdot)$, we refer to our developed method collectively as **TV-SBL**.

We now describe the motivation of our regularizer $T(\boldsymbol{\gamma})$. For maximally (random) sparse solutions, an appropriate choice is $T(\boldsymbol{\gamma}) = \sum_i I(\gamma_i)$, where the indicator function $I(\cdot)$ is an exact measure of sparsity as $I(\gamma_i) = 1$ for $\gamma_i > 0$ and zero otherwise. Since the function $T(\boldsymbol{\gamma})$ is intractable, many surrogate measures have been used, the most common one being the ℓ_1-norm in CS.

Using the indicator function to help block-sparsity, the main driver of our work is $T(\boldsymbol{\gamma}) = \sum_i I(|\gamma_i - \gamma_{i-1}|)$, i.e., TV on $\boldsymbol{\gamma}$. This assumes equal variances of the entries within a block and thus optimally counts the number of edges in the underlying signal. Armed with this ideal measure, we can use tractable measures developed in CS on the TV inspired input variable $|\gamma_i - \gamma_{i-1}|$ to identify appropriate block structures. It is noteworthy that imposing this regularizer on the hyperparameters rather than the source vectors $\mathbf{x}_l$ is an important distinction and also key to the success of our approach. CS theory has developed many regularizers which are monotonically increasing and concave on the positive orthant to promote sparsity. We discuss two options to illustrate the potential of the TV framework.

1) Linear TV: Conventional Smoother

The Linear TV regularizer is equivalent to the ℓ_1 penalty in CS and is given by the form

$$T(\boldsymbol{\gamma}) = \textstyle\sum_{i=2}^N |\gamma_i - \gamma_{i-1}|. \tag{4}$$

It has also been used in different signal processing applications to preserve edges and enforce local smoothness. We use this convex regularizer to enforce a block structure in the recovered signal. In addition to the signal regions, this penalty is found to denoise the zeros more effectively than the unregularized SBL algorithm.

2) Log TV: CS-based Regularizer

Another widely used regularizer in CS is $\sum_{i=1}^N \log(|x_i| + \epsilon)$, where ϵ is a positive stability parameter. This regularizer employs an iterative reweighted ℓ_1 minimization algorithm and has been shown to yield superior recovery [28, 29]. Utilizing this regularizer for block-sparsity, the Log TV regularizer is given by

$$T(\boldsymbol{\gamma}) = \textstyle\sum_{i=2}^N \log(|\gamma_i - \gamma_{i-1}| + \epsilon). \tag{5}$$

As in the CS, the Log TV based approach is found to be more effective than the Linear TV. This is due to its better resemblance to ℓ_0-norm [29], allowing more signal variance differences within a block and restraining small (faulty) signal estimate components to emerge.

3. OPTIMIZATION APPROACHES FOR TV-SBL

There are many options for minimizing the general SBL objective function. We apply the majorization-minimization (MM) approach and derive an iterative algorithm for minimizing the TV-SBL cost.

3.1. Optimization of TV-SBL with Linear TV

The TV-SBL optimization (3) for the Linear TV regularizer in (4) is

$$\boldsymbol{\gamma}^* = \underset{\boldsymbol{\gamma} \succeq \mathbf{0}}{\mathrm{argmin}} \; L \log|\boldsymbol{\Sigma}_{\mathbf{y}}| + \textstyle\sum_{l=1}^L \mathbf{y}_l^{\mathrm{H}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}_l + \beta \sum_{i=2}^N |\gamma_i - \gamma_{i-1}|. \tag{6}$$

Using the MM technique similar to [30], we majorize the concave term $\log|\boldsymbol{\Sigma}_{\mathbf{y}}|$ and solve iteratively a sequence of convex optimization problems. We majorize (i.e., linearize) $\log|\boldsymbol{\Sigma}_{\mathbf{y}}|$ by its first-order Taylor approximation at point $\boldsymbol{\Gamma}^{(j)}$, i.e.,

$$\begin{aligned} &\log|\lambda \mathbf{I} + \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^{\mathrm{H}}| \leq \log|\lambda \mathbf{I} + \mathbf{A}\boldsymbol{\Gamma}^{(j)}\mathbf{A}^{\mathrm{H}}| + \\ &\mathrm{Tr}\big((\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)})^{-1} \mathbf{A}\mathbf{A}^{\mathrm{H}} [\boldsymbol{\Gamma} - \boldsymbol{\Gamma}^{(j)}]\big), \end{aligned} \tag{7}$$

where the superscript j denotes the MM iteration index. Using (7), at iteration j, we end up with solving the convex problem

$$\begin{aligned} \boldsymbol{\gamma}^{(j+1)} = &\underset{\boldsymbol{\gamma} \succeq \mathbf{0}}{\mathrm{argmin}} \; L \, \mathrm{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}\right)^{-1} \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^{\mathrm{H}}\right) \\ &+ \textstyle\sum_{l=1}^L \mathbf{y}_l^{\mathrm{H}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}_l + \beta \sum_{i=2}^N |\gamma_i - \gamma_{i-1}|, \end{aligned} \tag{8}$$

and then updating $\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}$ using the newly obtained $\boldsymbol{\gamma}^{(j+1)}$.

3.2. Optimization of TV-SBL with Log TV

The TV-SBL optimization (3) for the Log TV regularizer in (5) is

$$\begin{aligned}\boldsymbol{\gamma}^* = &\underset{\boldsymbol{\gamma}\succeq\mathbf{0}}{\operatorname{argmin}}\; L\log|\boldsymbol{\Sigma}_{\mathbf{y}}| + \textstyle\sum_{l=1}^{L}\mathbf{y}_l^{\mathrm{H}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}_l \\ &+\beta\textstyle\sum_{i=2}^{N}\log(|\gamma_i-\gamma_{i-1}|+\epsilon).\end{aligned} \tag{9}$$

Similar to the Linear TV case above, we apply the MM approach for (9). Besides majorizing the $\log|\boldsymbol{\Sigma}_{\mathbf{y}}|$ term via (7), we majorize the concave Log TV penalty (5) by its first-order Taylor approximation at points $(\gamma_i^{(j)}-\gamma_{i-1}^{(j)})$, $i=2,\ldots,N$, i.e.,

$$\log(|\gamma_i-\gamma_{i-1}|+\epsilon) \le \log(|\gamma_i^{(j)}-\gamma_{i-1}^{(j)}|+\epsilon) + \frac{|\gamma_i-\gamma_{i-1}|}{|\gamma_i^{(j)}-\gamma_{i-1}^{(j)}|+\epsilon}. \tag{10}$$

Thus, at iteration j, we solve the convex problem

$$\begin{aligned}\boldsymbol{\gamma}^{(j+1)} = &\underset{\boldsymbol{\gamma}\succeq\mathbf{0}}{\operatorname{argmin}}\; L\operatorname{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}\right)^{-1}\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^{\mathrm{H}}\right) \\ &+\textstyle\sum_{l=1}^{L}\mathbf{y}_l^{\mathrm{H}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}_l + \beta\sum_{i=2}^{N}\frac{1}{|\gamma_i^{(j)}-\gamma_{i-1}^{(j)}|+\epsilon}|\gamma_i-\gamma_{i-1}|,\end{aligned} \tag{11}$$

followed by updating $\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}$ using the newly obtained $\boldsymbol{\gamma}^{(j+1)}$.

3.3. Convex Solver Implementation of TV-SBL

Any convex optimization package can be implemented to solve (8) and (11). Algorithm 1 presents the implementation of TV-SBL via the widely used CVX optimization package [31] to facilitate easy adoption and experimentation. One key step is to handle the matrix inverse in $\mathbf{y}_l^{\mathrm{H}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}_l$ through the Schur's complement equivalence [32, Appendix A5.5] by introducing the Hermitian symmetric matrix variables $\mathbf{Z}_{\mathrm{cvx},l}\in\mathbb{S}^{M\times M}$, $l=1,\ldots,L$.

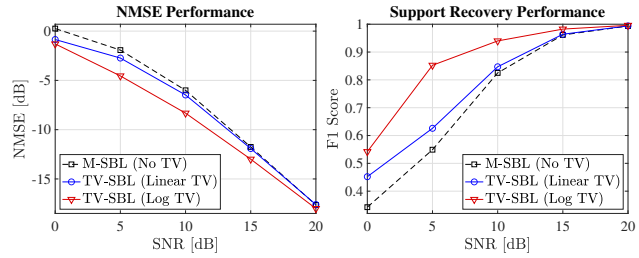
Algorithm 1 CVX Solver for TV-SBL

Input: $\mathbf{A}, \mathbf{Y}, \boldsymbol{\gamma}^{(0)}, \lambda, \beta, \epsilon, j_{\max}$
Output: $\boldsymbol{\mu}_{\mathbf{x}_l|\mathbf{y}_l;\boldsymbol{\gamma}}$ $(\forall\, l\in\{1,\ldots,L\})$

1: **for** $j=0$ to $j_{\max}-1$ **do**
2: Evaluate $[\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}]^{-1} = (\lambda\mathbf{I}+\mathbf{A}\boldsymbol{\Gamma}^{(j)}\mathbf{A}^{\mathrm{H}})^{-1}$
3: CVX variables: $\boldsymbol{\gamma}_{\mathrm{cvx}}\in\mathbb{R}^N$, $\mathbf{Z}_{\mathrm{cvx},l}\in\mathbb{S}^{M\times M}$, $l=1,\ldots,L$
4: **minimize**: $L\operatorname{Tr}\left[\left(\boldsymbol{\Sigma}_{\mathbf{y}}^{(j)}\right)^{-1}\mathbf{A}\boldsymbol{\Gamma}_{\mathrm{cvx}}\mathbf{A}^{\mathrm{H}}\right]+\sum_{l=1}^{L}\operatorname{Tr}(\mathbf{Z}_{\mathrm{cvx},l})+\beta\, T(\boldsymbol{\gamma}_{\mathrm{cvx}})$ $\quad\{T(\boldsymbol{\gamma}_{\mathrm{cvx}})\to$ Linear TV (4) or Log TV (5)$\}$
5: **subject to**: $\boldsymbol{\gamma}_{\mathrm{cvx}}\succeq\mathbf{0}$, $\begin{bmatrix}\mathbf{Z}_{\mathrm{cvx},l} & (\mathbf{y}_l\mathbf{y}_l^{\mathrm{H}})^{1/2}\\ (\mathbf{y}_l\mathbf{y}_l^{\mathrm{H}})^{1/2} & \lambda\mathbf{I}+\mathbf{A}\boldsymbol{\Gamma}_{\mathrm{cvx}}\mathbf{A}^{\mathrm{H}}\end{bmatrix}\succeq\mathbf{0}$ $(\forall\, l\in\{1,\ldots,L\})$
6: $\boldsymbol{\gamma}^{(j+1)}\leftarrow\boldsymbol{\gamma}_{\mathrm{cvx}}$
7: **end for**
8: Evaluate $\boldsymbol{\mu}_{\mathbf{x}_l|\mathbf{y}_l;\boldsymbol{\gamma}}$ $(\forall\, l\in\{1,\ldots,L\})$ using $\boldsymbol{\gamma}^{(j_{\max})}$ in (2)

4. NUMERICAL RESULTS

This section provides the numerical results for the CVX solver described in Algorithm 1. For the MMV setup (1), we consider a signal length of $N=150$ with $M=20$ measurements and $L=5$ snapshots. We form the dictionary $\mathbf{A}\in\mathbb{R}^{M\times N}$ by first drawing its elements from a Gaussian distribution, and then normalizing the columns as $\|\cdot\|_2=1$. The signal ensemble $\mathbf{X}$ contains $K=10$ non-zero rows and each non-zero element is drawn from $\mathcal{N}(0,1/K)$. We consider three classes of block-sparse signals:
(i) Homogeneous block-sparse signal with 2 blocks of length 5 each;



(a) (b)

Fig. 1. Comparison of Linear TV and Log TV penalties under homogeneous block-sparsity (see Fig. 2(a)): (a) NMSE and (b) F_1-Score.

(ii) Random sparse signal with 10 randomly placed components (which are thus mostly isolated);
(iii) Hybrid sparse signal with 1 block of length 4, 1 block of length 3, and 3 isolated components.

Each noise signal $\mathbf{n}_l$, $l=1,\ldots,L$, is generated from $\mathcal{N}(0,\sigma_n^2)$ with variance σ_n^2 chosen so that the Signal-to-noise ratio (SNR), $10\log_{10}\left(\frac{\mathbb{E}\,||\mathbf{A}\mathbf{x}_l||^2}{\mathbb{E}\,||\mathbf{n}_l||^2}\right)$[1], varies from 0 to 20 dB.

We assess performance via the Normalized mean square error (NMSE) and support recovery. The NMSE is defined as $\mathbb{E}\left[\frac{||\hat{\mathbf{X}}-\mathbf{X}||^2}{||\mathbf{X}||^2}\right]$, where $\hat{\mathbf{X}}$ is the estimated source matrix and the norm used is the Frobenius norm. Support recovery is evaluated using the F_1-Score, defined as [33] $F_1=\mathbb{E}\left[2\,\frac{\text{precision}\times\text{recall}}{\text{precision}+\text{recall}}\right]$, where $\text{precision}=\frac{\text{tp}}{\text{tp}+\text{fa}}$, $\text{recall}=\frac{\text{tp}}{\text{tp}+\text{mis}}$, "tp": number of true positives, "fa": number of false alarms, and "mis": number of misdetections.

Remark 1. For the ease of comparison, we evaluate the support recovery by preserving the K largest rows of $\hat{\mathbf{X}}$ while setting the rest to zero. In practice, the support is estimated using a fixed threshold.

Performance of Different TV Penalties

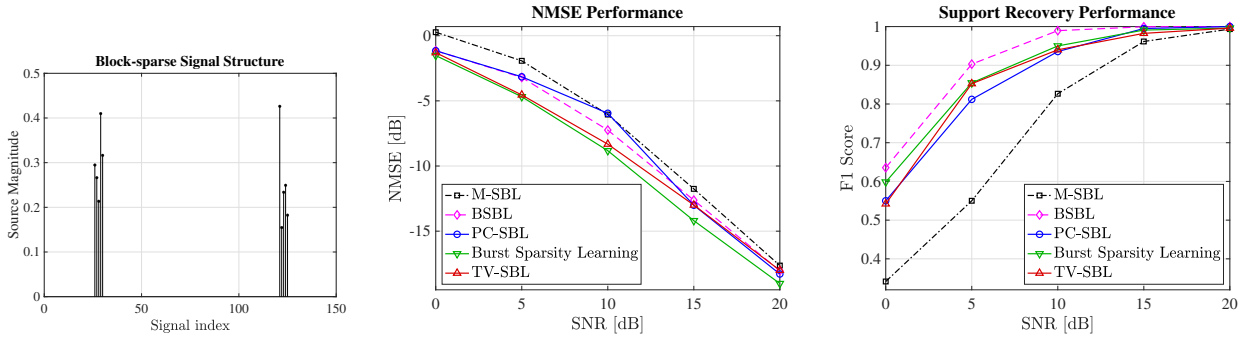
Fig. 1 compares the performance of the Linear and Log TV regularizer in (4) and (5), respectively, for the homogeneous block-sparse signal. Both regularizers improve the performance from that of M-SBL. We observe an improved performance for the Log TV penalty, consistent with our claim in Sec. 2.2, showing that it is more adept at identifying block structures and denoising the zero rows of $\mathbf{X}$.

Comparison with Benchmark Algorithms

We study all three block-sparsity classes and compare the performance of our TV-SBL (Log TV) algorithm to SBL-based block-sparse recovery algorithms: *(i)* BSBL [12], *(ii)* PC-SBL [14], and *(iii)* Burst Sparsity Learning [15]. The M-SBL algorithm [19] is used as a reference to show recovery without regularization. In order to assess the robustness of each algorithm to changes in block patterns, the parameters of each algorithm were empirically tuned for the homogeneous block-sparse signal over the SNR range, and then left unchanged for random and hybrid sparse signals.

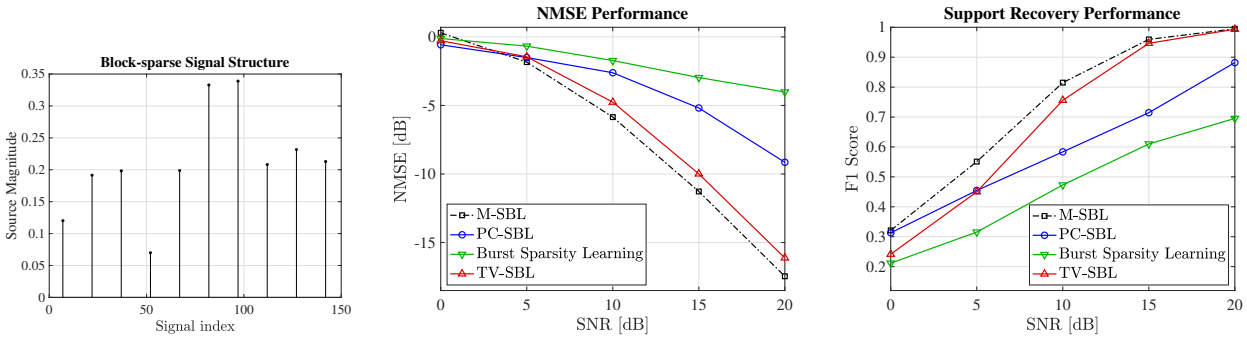
1) Homogeneous block-sparse signals: As seen in Fig. 2, all algorithms, unsurprisingly, outperform M-SBL. BSBL, being provided block size and boundary information *apriori*, attains the best F_1-Score (Fig. 2(c)). Even without partition knowledge, the regularized SBL algorithms fare comparably in F_1-Score and even exceed BSBL in NMSE. Only Burst Sparsity Learning, with its optimal coupling-based inference, exceeds TV-SBL which uses a softer prior. This illustrates that *pure* block-sparse recovery requires explicit coupling of parameters for the best performance.

[1] All expectations $\mathbb{E}[\cdot]$ are evaluated over 200 Monte Carlo trials.



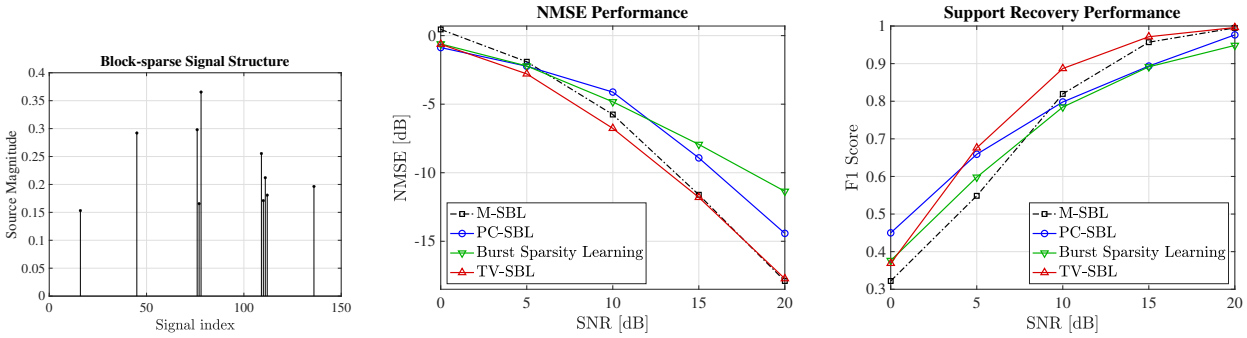
(a) (b) (c)

Fig. 2. Recovery performance for homogeneous block-sparsity: (a) 2 blocks of length 5, (b) NMSE, and (c) F_1-Score.



(a) (b) (c)

Fig. 3. Recovery performance for random sparsity: (a) 10 blocks of length 1, (b) NMSE, and (c) F_1-Score.



(a) (b) (c)

Fig. 4. Recovery performance for hybrid sparsity: (a) 1 block of len. 4, 1 block of len. 3, and 3 blocks of len. 1, (b) NMSE, and (c) F_1-Score.

We now demonstrate that a softer prior for block-sparsity in TV-SBL gains in increased flexibility to block structure.

2) Sparse signals: Fig. 3(a) represents the extreme scenario for the block-sparse algorithms, i.e., the block size is one. TV-SBL outperforms the coupling-based algorithms, being comparable to M-SBL. Explicit hyperparameter coupling biases the algorithms to block structures, and thus renders them ineffective for isolated sparsity. Using a softer prior, TV-SBL supports block-sparsity without such excessive bias; it is remarkably adept at isolated sparsity as well.

3) Hybrid sparse signals: The hybrid block structure in Fig. 4(a) is representative of a practical scenario for, e.g., MIMO wireless channel models, with varying angular spreads due to uneven scattering. As seen in Fig. 4, TV-SBL outperforms all other algorithms in this setting. The soft prior introduced by TV-SBL accommodates blocks as well as isolated components. To summarize, TV-SBL shows itself to be a robust block-sparse recovery algorithm.

5. CONCLUSION

We proposed a TV-regularized SBL method for recovering signal blocks of unknown sizes and boundaries from compressive measurements. As a fresh idea, the method imposes a soft TV prior on the SBL hyperparameters to encourage block-sparse solutions. The developed iterative majorization-minimization algorithm necessitates only convex optimization tools to solve the problem, enabling a use of numerous efficient solvers. The numerical results showed that TV-SBL obtains superior trade-off between recovering block-sparse and random sparse signals. Such robustness has great utility in a diversity of practical sparse signal estimation scenarios.

Imposing a TV penalty on the SBL hyperparameters opens up several concave regularization penalties as well as fast numerical solvers. Our future analysis will study further the hyperparameter regularization from a more general perspective of the TV penalty.

6. REFERENCES

[1] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on millimeter-wave channel modeling for 5G cellular system," *IEEE J. Select. Topics Signal Process.*, vol. 10, no. 3, pp. 454–469, April 2016.

[2] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1445–1455, March 2014.

[3] E. J. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[5] Ming Yuan and Yi Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[6] Ewout Van Den Berg and Michael P Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2009.

[7] Richard G Baraniuk, Volkan Cevher, Marco F Duarte, and Chinmay Hegde, "Model-based compressive sensing," *IEEE Transactions on information theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[8] Yonina C Eldar, Patrick Kuppinger, and Helmut Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[9] Junzhou Huang, Tong Zhang, and Dimitris Metaxas, "Learning with structured sparsity," *Journal of Machine Learning Research*, vol. 12, no. 11, 2011.

[10] T. Peleg, Y. C. Eldar, and M. Elad, "Exploiting statistical dependencies in sparse representations for signal recovery," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2286–2303, May 2012.

[11] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," *IEEE J. Select. Topics Signal Process.*, vol. 5, no. 5, pp. 912–926, Sep. 2011.

[12] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2009–2015, April 2013.

[13] Lei Yu, Hong Sun, Jean-Pierre Barbot, and Gang Zheng, "Bayesian compressive sensing for cluster structured sparse signals," *Signal processing*, vol. 92, no. 1, pp. 259–269, 2012.

[14] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 360–372, Jan 2015.

[15] J. Dai, A. Liu, and H. C. So, "Non-uniform burst-sparsity learning for massive MIMO channel estimation," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 1075–1087, Feb 2019.

[16] L. Wang, L. Zhao, S. Rahardja, and G. Bi, "Alternative to extended block sparse Bayesian learning and its relation to pattern-coupled sparse bayesian learning," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2759–2771, May 2018.

[17] Mohammad Shekaramiz, Todd K. Moon, and Jacob H. Gunther, "Bayesian compressive sensing of sparse signals with unknown clustering patterns," *Entropy*, vol. 21, no. 3, 2019.

[18] Lu Wang, Lifan Zhao, Lei Yu, Jingjing Wang, and Guoan Bi, "Structured Bayesian learning for recovery of clustered sparse signal," *Signal Processing*, vol. 166, pp. 107255, 2020.

[19] D. P. Wipf and B. D. Rao, "An empirical bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[20] P. Pal and P. P. Vaidyanathan, "Pushing the limits of sparse support recovery using correlation information," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 711–726, 2015.

[21] R. R. Pote and B. D. Rao, "Robustness of sparse Bayesian learning in correlated environments," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Barcelona, Spain, May 4-8, 2020, pp. 9100–9104.

[22] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[23] D. Wipf, J. Palmer, B. Rao, and K. Kreutz-Delgado, "Performance evaluation of latent variable models with sparse priors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2007, vol. 2, pp. II–453–II–456.

[24] T. A. Srikrishnan and B. D. Rao, "Addressing the noise variance problem in sparse bayesian learning," in *Proc. Asilomar Conf. Signals, Syst., Comp.*, 2018, pp. 1974–1979.

[25] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 460–489, 2005.

[26] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259 – 268, 1992.

[27] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 813–824, 1998.

[28] Y. Shen, J. Fang, and H. Li, "Exact reconstruction analysis of log-sum minimization for compressed sensing," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1223–1226, 2013.

[29] E. J. Candés, M.B. Wakin, and S.P. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.

[30] David P. Wipf and Srikantan S. Nagarajan, "A new view of automatic relevance determination," in *Proc. Int. Conf. Neural Inform. Process. Syst.*, pp. 1625–1632. 2008.

[31] Michael Grant and Stephen Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," `http://cvxr.com/cvx`, Mar. 2014.

[32] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[33] Nancy Chinchor and Ph D, "Muc-4 evaluation metrics," in *In Proceedings of the Fourth Message Understanding Conference*, 1992, pp. 22–29.