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

46th International Conference on Acoustics, Speech, and Signal Processing

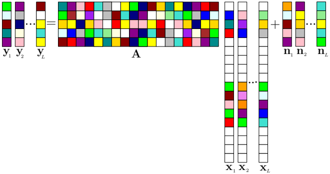
Overview

Block-sparse Signal Recovery

Block-sparse signal recovery: Overview

- Applications in, e.g., mmWave comm., audio and image processing
- Block **sizes and boundaries unknown**: Number of possible blocks grows exponentially in signal size

MMV block-sparse signal model:



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

- $\mathbf{y}_l \in \mathbb{C}^M$: measurement at time l, $\mathbf{A} \in \mathbb{C}^{M \times N}$: dictionary, & $\mathbf{n}_l$: AWGN at time l
- Source and noise vectors are i.i.d
- Signal ensemble $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_L]$ is *block-row-sparse* with block **sizes and locations unknown**

Our aim: Develop algorithm robust to different signal block-sparsity structures

Sparse Bayesian Learning (SBL) for MMV Sparse Signal Recovery

Advantages of using the SBL framework for MMV sparse recovery

- The M-SBL [Wipf, Rao, 2007] abstracts row sparsity of $\mathbf{X}$ by a single parameter (γ_i), **reducing parameters** from NL to N
- It falls under the class of methods that are **correlation-aware** which have shown superior ability to find sparse solutions [Pal, Vaidyanathan, 2014]
- SBL shows great promise for sparse signal recovery under **correlated sources** and ill-conditioned dictionaries [Pote, Rao, 2020]

SBL framework

- *Parametric Gaussian prior* for each signal as

$$p(\mathbf{x}_l; \boldsymbol{\gamma}) = \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}) = \frac{1}{|\pi\boldsymbol{\Gamma}|}\exp\left(-\mathbf{x}_l^{\mathsf{H}}\boldsymbol{\Gamma}^{-1}\mathbf{x}_l\right), \quad \boldsymbol{\Gamma} \triangleq \text{diag}(\boldsymbol{\gamma})$$

- $\boldsymbol{\gamma} = [\gamma_1 \cdots \gamma_N]^{\mathsf{T}} \in \mathbb{R}_+^N$ is a vector of *hyperparameters* **reflecting sparsity**

SBL solves sparse recovery by estimation of the hyperparameters γ, instead of each sparse row of $\mathbf{X}$ individually

SBL Inference: Hyperparameter Estimation

SBL inference: Type-II maximum a posteriori (MAP) estimation of the posterior $p(\boldsymbol{\gamma}|\mathbf{y}_1, \ldots, \mathbf{y}_L)$ over $\boldsymbol{\gamma}$

- General optimization equation:

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

  where $\boldsymbol{\Sigma}_{\mathbf{y}} = \lambda \mathbf{I} + \mathbf{A} \boldsymbol{\Gamma} \mathbf{A}^{\mathsf{H}}$ and $\log p(\boldsymbol{\gamma})$ is the *hyperprior* on $\boldsymbol{\gamma}$

- Sparse signal: MAP estimate over posterior density $p(\mathbf{x}_l|\mathbf{y}_l; \boldsymbol{\gamma})$

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

Role of hyperprior: $\log p(\boldsymbol{\gamma})$

- Used to *enforce additional structure* on signal $\mathbf{X}$
- Improve recovery when the number of snapshots are limited

Our approach: We introduce a robust block-sparsity hyperprior $\log p(\boldsymbol{\gamma})$

Block-sparse Regularization Using Total Variation

Limitations of existing approaches:

- Conventional methods (e.g., Block-OMP) *need block size and boundaries*
- SBL-based methods (e.g., PC-SBL) impose *stringent explicit hyperparameter coupling* and are sensitive to any model mismatch

Our approach: SBL with Total Variation (TV) regularization for hyperparameters, i.e., **TV-SBL**

- Inspired by an **ideal block-counting measure** $T(\boldsymbol{\gamma}) = \sum_i I(|\gamma_i - \gamma_{i-1}|)$
- **Imposed on SBL hyperparameters**, as opposed to the signal; important distinction and also key to the success of our approach
- Assumes equal variances (SBL hyperparameter $\boldsymbol{\gamma}$) of the entries within a block, optimally counting the edges
- Tractable measures developed in CS for $I(|\gamma_i - \gamma_{i-1}|)$; we investigated **two such measures**

Block-sparsity enforced through block counting in SBL hyperparameter space

TV-SBL Regularizers

Original intractable block-counting measure: $T(\boldsymbol{\gamma}) = \sum_{i=2}^{N} I(|\gamma_i - \gamma_{i-1}|)$

- *Block-extension* to the elementary sparsity-promoting ℓ_0-"norm"
- Tractability introduced through surrogate measures for $I(\cdot)$

1) Linear TV: $T(\boldsymbol{\gamma}) = \sum_{i=2}^{N} |\gamma_i - \gamma_{i-1}|$

- Equivalent to the ℓ_1-norm regularization in CS
- Penalty in this form: **TV-regularizer** used in signal processing to preserve edges and enforce local smoothness

2) Log TV: $T(\boldsymbol{\gamma}) = \sum_{i=2}^{N} \log(|\gamma_i - \gamma_{i-1}| + \epsilon)$

- Extending the CS regularizer $\sum_{i=1}^{N} \log(|x_i| + \epsilon)$ [Candés,Wakin,Boyd,2008]
- Here too we employ *iterative reweighted* ℓ_1 *minimization* for inference
- Analogous to CS, shows better performance than Linear TV

These regularizers enable formulation of TV-SBL via **convex optimization**

Convex Optimization Formulation of TV-SBL

Solving unregularized SBL as a convex optimization [Wipf,Nagarajan,2008]

Using the majorization minimization (MM) on the concave term $\log|\mathbf{\Sigma_y}|$ as

$$\log|\lambda\mathbf{I} + \mathbf{A\Gamma A}^{\mathsf{H}}| \leq \log|\lambda\mathbf{I} + \mathbf{A\Gamma}^{(j)}\mathbf{A}^{\mathsf{H}}| + \mathrm{Tr}\big((\mathbf{\Sigma}_{\mathbf{y}}^{(j)})^{-1}\mathbf{AA}^{\mathsf{H}}[\mathbf{\Gamma} - \mathbf{\Gamma}^{(j)}]\big),$$

enables solving the **unregularized** SBL cost function as a convex optimization

1) Linear TV:

Block-sparse TV regularizer is convex, hence the optimization is

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

2) Log TV:

Analogous to CS, the Log TV penalty is majorized, giving the optimization

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

This enables a flexible choice of numerical solvers and optimization techniques;
Our optimization technique: CVX toolbox for MATLAB®

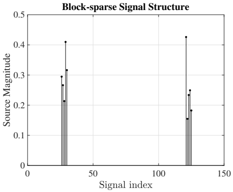
Simulation Setup and Block Distributions

Measurement Setup and Algorithm evaluation:

- *Measurement:* $L = 5$ snapshots of length $M = 20$
- *NMSE:* Defined as $\mathbb{E}[\ ||\hat{\mathbf{X}} - \mathbf{X}||^2/||\mathbf{X}||^2\ ]$, for the estimated $\hat{\mathbf{X}}$
- *Support recovery:* Evaluated using $\mathrm{F}_1 = \mathbb{E}\left[2\,\frac{(\text{precision}\times\text{recall})}{\text{precision}+\text{recall}}\right]$[1]

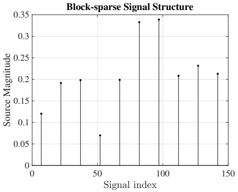
Block distributions used: (Total signal length: 150 with sparsity: 10)

1) Homogeneous block-sparsity



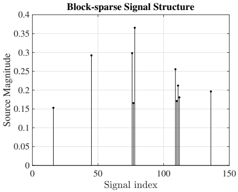
Signal: 2 blocks of size 5 each

2) Random sparsity



Signal: 10 isolated components

3) Hybrid sparsity

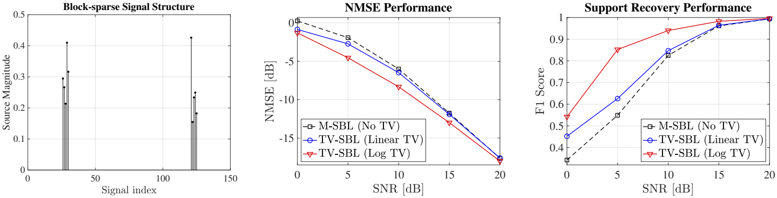


Signal: 2 blocks of size 4 & 3, & 3 isolated components

[1] $\text{precision} = \text{tp}/(\text{tp} + \text{fa})$, $\text{recall} = \text{tp}/(\text{tp} + \text{mis})$, "tp": true positives, "fa": false alarms, and "mis": misdetections

Comparing TV Penalties

Performance of TV-SBL regularizers for Homogeneous Block-sparsity:



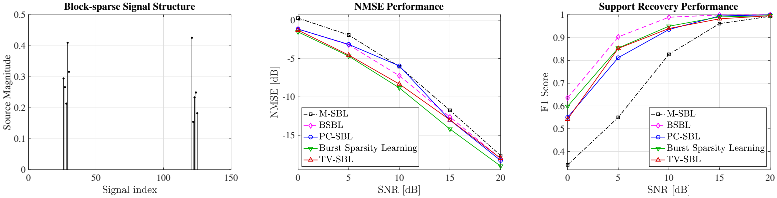
- TV-regularization promotes block-sparsity over unregularized M-SBL
- Log TV more adept at identifying block structures over Linear TV

> Improved performance of Log TV, consistent with CS theory, is because it is a better tractable surrogate for the block-counting measure

Remark: Subsequent tests for TV-SBL use Log TV regularization

Comparing TV-SBL to Benchmark Block-sparse Recovery Algorithms

Performance comparison for Homogeneous Block-sparsity

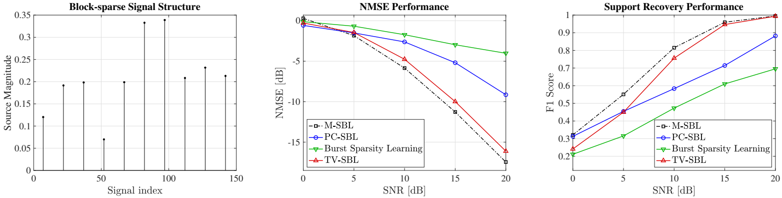


- Superior support recovery: *BSBL [Zhang,Rao,2013]*, due to prior information of block size and boundaries
- Superior NMSE: *Burst Sparsity Learning [Dai,et al.,2019]* due to optimal SBL hyperparameter coupling-based inference for block-sparsity
- TV-SBL performance: Softer TV-based encourages block structure, though not as optimally as explicit coupling

Best strategy for homogeneous block-sparsity: Exact hyperparameter coupling

Comparing TV-SBL to Benchmark Block-sparse Recovery Algorithms

Performance comparison for Random Sparsity (Isolated components)

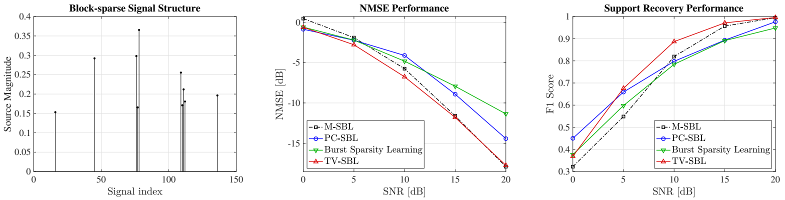


- Represents least favourable scenario for block-sparse recovery algorithms
- Superior Performance: *M-SBL*, due to absence of block-sparse prior
- TV-SBL performance: Outperforms other block-sparse regularizers due to softer prior **without excessive bias** to block-structure

Using a softer prior, TV-SBL supports block-sparsity without excessive bias, hence remarkably adept at isolated sparsity as well

Comparing TV-SBL to Benchmark Block-sparse Recovery Algorithms

Performance comparison for Hybrid Sparsity



- *Representative of a practical scenario* for, e.g., MIMO wireless channels
- TV-SBL performance: Outperforms all algorithms, robustly accommodating both block-sparse and isolated components

Overall, TV-SBL obtains superior trade-off between recovering block-sparse and random sparse signals, hence a very robust algorithm

Conclusions and Future Work

Through the novel TV hyperparmeter regularized SBL framework:

- Introduce an SBL regularizer based on an ideal block-counting metric
- Iterative convex optimization strategy; enables **efficient numerical solvers**
- Soft TV-SBL prior **not biased** to block structure
- **Superior trade-off** between block-sparse and random sparse signals; robust signal recovery

Future work:

- Explore further surrogate measures for the block-counting function
- Generalize perspective of TV regularization beyond the current block-counting function

SBL hyperparameter regularization holds immense potential for block-sparse recovery, both from a conceptual perspective and an efficient algorithm design