

COMPRESSED SENSING FOR BLOCK-SPARSE SMOOTH SIGNALS

Shahzad Gishkori and Geert Leus *

Faculty of EEMCS, Delft University of Technology, Delft, The Netherlands

ABSTRACT

We present reconstruction algorithms for smooth signals with block sparsity from their compressed measurements. We tackle the issue of varying group size via the group-sparse least absolute shrinkage selection operator (LASSO) as well as via latent group LASSO regularizations. We achieve smoothness in the signal via fusion. We develop low-complexity solvers for our proposed formulations through the alternating direction method of multipliers.

Index Terms— Compressed sensing, block sparsity, smoothness, signal reconstruction

1. INTRODUCTION

Compressed sensing (CS) [1, 2] is one of the most exciting topics of present-day signal processing. Signal reconstruction from its low-dimensional representation becomes a possibility, given the sparse nature of the signal and, incoherence and/or restricted isometry property (RIP) [2] of the sensing/measurement process. In this regard, a number of approaches can be used, e.g., basis pursuit (BP) [3], least absolute shrinkage and selection operator (LASSO) [4] and greedy algorithms [5]. In order to exploit the structure of the signal being sensed, a number of variants of LASSO have become popular, e.g., group LASSO (G-LASSO) [6], sparse group LASSO (SG-LASSO) [7] and fused LASSO (F-LASSO) [8], etc. In this paper we propose new LASSO formulations to handle block sparse smooth signals. Smooth signals are often encountered in a wide range of engineering and biological fields. In engineering, signals observed in image processing, control systems and environment monitoring are often smooth or piece-wise smooth. In biology, a similar structure is observed, e.g., in protein mass spectroscopy [8]. The goal is to recover such structured signals from noisy and/or under-sampled measurements. A related topic is signal smoothing which deals with removing random outliers. Apart from being smooth, such signals can often be represented as sparse in some basis. This sparsity pattern normally varies in terms of the location and block size of the sparse coefficients. The challenge for signal reconstruction is to exploit the block sparsity with varying block sizes, while keeping smoothness intact and using fewer measurements, but all at low complexity. In the CS domain, signal smoothness has been handled by using a fusion constraint in [8]. The fusion is also known as total variation (TV) in the image processing literature. Apart from fusion, [8] also proposes an ℓ_1 -norm penalty to cater for signal sparsity. However, since most of the signals are block sparse, [8] cannot give efficient results. To cater for the block sparsity, one can replace the ℓ_1 -norm penalty with a group penalty. Although this approach can handle the block sparsity very well, it only offers fixed group sizes and causes complete groups to be zero or nonzero. To avoid elimination of small sets of nonzero elements, a very small group size is

opted but that can make the algorithm inefficient in eliminating large blocks of zero elements. In this regard, we propose to use a moderate group size along with an ℓ_1 -norm penalty over the signal, to create sparsity within the groups. Thus by using fusion in combination with an ℓ_1 -norm penalty and a moderate group size, a smooth signal can be reconstructed with high accuracy. The problem of varying group sizes can also be handled by using the concept of latent groups, see [9] and references therein. These are basically overlapping groups, with recurring signal elements in possibly multiple groups. Thus, an element lost in one group may resurface through another group after reconstruction. So we also propose to use such latent groups in combination with a fusion constraint to recover block sparse smooth signals with varying block sizes. Note that a work on using overlapping groups over the fusion function instead of the signal structure has appeared in [10], which however requires complete signal samples. Instead, we propose overlapping groups and fusion penalties over the actual signal for under-determined systems. Thus, we exploit the actual structure of the signal rather than the difference of elements. Further, in order to solve the proposed formulations, we derive low-complexity algorithms based on the alternating direction method of multipliers (ADMM) [11]. The reason for using this version of the augmented Lagrangian methods is primarily the non-separability of the fusion penalty in terms of the elements of the signal. Thus, the general convergence properties of ADMM can be used to guarantee optimal results for our proposed algorithms.

Notations. Matrices are in upper case bold while column vectors are in lower case bold, $[\mathbf{X}]_{i,j}$ is the (i, j) th entry of the matrix $\mathbf{X}$, $[\mathbf{x}]_i$ is the i th entry of the vector $\mathbf{x}$, $\mathbf{I}_N$ is the identity matrix of size $N \times N$, $(\cdot)^T$ is transpose, $\hat{\mathbf{x}}$ is the estimate of $\mathbf{x}$, $\triangleq$ defines an entity, $\|\mathbf{x}\|_p = (\sum_{i=0}^{N-1} |[\mathbf{x}]_i|^p)^{1/p}$ is the ℓ_p norm of $\mathbf{x}$, $\text{sign}(x)$ is the sign function which takes values -1 and 1 depending on the polarity of the element x , whereas the function $(x)_+ = x$ if and only if $x > 0$ otherwise $(x)_+ = 0$.

2. SIGNAL RECONSTRUCTION

Let $\mathbf{x}$ be the $N \times 1$ recoverable signal. Given M measurements, the sensed signal can be written as

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{v} \quad (1)$$

where $\mathbf{y}$ is an $M \times 1$ measurement vector, Φ is an $M \times N$ measurement matrix ($M < N$) with compression ratio $\mu \triangleq M/N$ and $\mathbf{v}$ is an $M \times 1$ zero-mean additive white Gaussian noise vector with variance σ_v^2 . To recover the signal from the compressed measurements while keeping the signal structure in tact, we propose below, two LASSO formulations.

2.1. Sparse Group LASSO with Fusion

Through sparse group fused LASSO (SGF-LASSO), we can resolve the issue of signal smoothness, as well as, that of fixed group sizes.

*This work is supported in part by NWO-STW under the VICI program (project 10382).

The optimization problem can be formulated as

$$\begin{aligned} \hat{\mathbf{x}} = \arg \min_{\mathbf{x}} & \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_e \|\mathbf{x}\|_1^1 \\ & + \lambda_g \sum_{i=0}^{G-1} \|\mathbf{x}_i\|_2^1 + \lambda_f \sum_{j=1}^{N-1} \|\mathbf{x}[j] - \mathbf{x}[j-1]\|_1^1 \end{aligned} \quad (2)$$

where $\mathbf{x}_i$ is an $N/G \times 1$ sub-vector of $\mathbf{x}$, representing one of G groups over the elements of $\mathbf{x}$, i.e., $\mathbf{x} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{G-1}^T]^T$. We can see from (2) that $\lambda_g \sum_{i=0}^{G-1} \|\mathbf{x}_i\|_2^1$ accounts for group sparsity, $\lambda_e \|\mathbf{x}\|_1^1$ for element-wise sparsity and $\lambda_f \sum_{j=1}^{N-1} \|\mathbf{x}[j] - \mathbf{x}[j-1]\|_1^1$ accounts for fusion within the elements of $\mathbf{x}$, such that the effect of each penalty becomes severer with increasing penalty parameters, i.e., λ_g , λ_e and λ_f , respectively. For a moderate value of G , the proposed formulation can tackle the varying group size problem by creating sparsity within the group along with fusing consecutive elements. Note that, for $\lambda_g = \lambda_f = 0$, (2) reduces to the standard LASSO problem, for $\lambda_f = 0$, (2) reduces to SG-LASSO, for $\lambda_e = \lambda_f = 0$, (2) takes the shape of G-LASSO and for $\lambda_g = 0$, (2) becomes F-LASSO.

Solver for SGF-LASSO

In order to solve the SGF-LASSO problem via ADMM, we introduce two auxiliary variables $\mathbf{u}$ and $\mathbf{z}$ of size $N \times 1$. Thus, (2) can be written as

$$\begin{aligned} [\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\mathbf{z}}] = \arg \min_{\mathbf{x}, \mathbf{u}, \mathbf{z}} & \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_e \|\mathbf{u}\|_1^1 \\ & + \lambda_g \sum_{i=0}^{G-1} \|\mathbf{u}_i\|_2^1 + \lambda_f \|\mathbf{z}\|_1^1 \\ \text{s.t. } & \mathbf{u}_i = \mathbf{x}_i, \quad 0 \leq i \leq G-1, \quad \mathbf{z} = \mathbf{D}\mathbf{x} \end{aligned} \quad (3)$$

where $\mathbf{u}_i$ is an $N/G \times 1$ sub-vector of $\mathbf{u}$, i.e., $\mathbf{u} = [\mathbf{u}_0^T, \mathbf{u}_1^T, \dots, \mathbf{u}_{G-1}^T]^T$, and $\mathbf{D}$ is the difference matrix with $[\mathbf{D}]_{j,j} = -1$, $[\mathbf{D}]_{j,j+1} = 1$, for $0 \leq j \leq N-2$ and $[\mathbf{D}]_{N-1,N-1} = 1$, such that $\|\mathbf{D}\mathbf{x}\|_1^1$ equals the element-wise fusion. From (3), the Lagrangian problem can be written as

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{z}, \boldsymbol{\rho}_u, \boldsymbol{\rho}_z) = & \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_e \|\mathbf{u}\|_1^1 \\ & + \lambda_g \sum_{i=0}^{G-1} \|\mathbf{u}_i\|_2^1 + \lambda_f \sum_{j=2}^N \|\mathbf{z}\|_1^1 \\ & + \sum_{i=0}^{G-1} \boldsymbol{\rho}_{u_i}^T (\mathbf{u}_i - \mathbf{x}_i) + \frac{c_u}{2} \sum_{i=0}^{G-1} \|\mathbf{u}_i - \mathbf{x}_i\|_2^2 \\ & + \boldsymbol{\rho}_z^T (\mathbf{z} - \mathbf{D}\mathbf{x}) + \frac{c_z}{2} \|\mathbf{z} - \mathbf{D}\mathbf{x}\|_2^2 \end{aligned} \quad (4)$$

where $\boldsymbol{\rho}_u$ (with sub-vectors $\boldsymbol{\rho}_{u_i}$, for $0 \leq i \leq G-1$) and $\boldsymbol{\rho}_z$ are Lagrange multipliers and, c_u and c_z are positive constants. The solution of (3) is generated by the following successive approximations

$$\mathbf{x}^{[n]} = \arg \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{u}^{[n-1]}, \mathbf{z}^{[n-1]}, \boldsymbol{\rho}_u^{[n-1]}, \boldsymbol{\rho}_z^{[n-1]}) \quad (5)$$

$$\mathbf{u}^{[n]} = \arg \min_{\mathbf{u}} \mathcal{L}(\mathbf{x}^{[n-1]}, \mathbf{u}, \boldsymbol{\rho}_u^{[n-1]}) \quad (6)$$

$$\mathbf{z}^{[n]} = \arg \min_{\mathbf{z}} \mathcal{L}(\mathbf{x}^{[n-1]}, \mathbf{z}, \boldsymbol{\rho}_z^{[n-1]}) \quad (7)$$

and the multipliers are updated as

$$\boldsymbol{\rho}_u^{[n]} = \boldsymbol{\rho}_u^{[n-1]} + c_u (\mathbf{x}^{[n]} - \mathbf{u}^{[n]}) \quad (8)$$

$$\boldsymbol{\rho}_z^{[n]} = \boldsymbol{\rho}_z^{[n-1]} + c_z (\mathbf{D}\mathbf{x}^{[n]} - \mathbf{z}^{[n]}). \quad (9)$$

The closed-form solution for (5) at the n th iteration can be derived to be

$$\begin{aligned} \mathbf{x}^{[n]} = & \left(\Phi^T \Phi + c_z \mathbf{D}^T \mathbf{D} + c_u \mathbf{I}_N \right)^{-1} \\ & \times \left(\Phi^T \mathbf{y} - \mathbf{D}^T \boldsymbol{\rho}_z^{[n-1]} + c_z \mathbf{D}^T \mathbf{z}^{[n-1]} - \boldsymbol{\rho}_u^{[n-1]} + c_u \mathbf{u}^{[n-1]} \right). \end{aligned} \quad (10)$$

We can see from (10) that the matrix inversion part does not change during the iterations so that it can be performed off-line, resulting in reduced complexity. Note that the matrix inversion lemma can be used to further ease the computations involved in the inversion operation. For $\mathbf{u}$, note that the optimization involves two penalties, i.e., apart from penalizing each element of $\mathbf{u}$ for sparsity, we need to optimize on each of its sub-groups as well. Since both penalties are non-differentiable, we utilize the fact that soft thresholding generates a minimizer for the cost function involving $\lambda_e \|\mathbf{u}_i\|_1^1$ [4], and for the cost function involving $\lambda_g \|\mathbf{u}_i\|_2^1$, the minimizer is $\mathbf{s}_u = \mathbf{u}_i / \|\mathbf{u}_i\|_2^2$ in case $\|\mathbf{u}_i\|_2^2 \neq 0$ and the minimizer is a vector $\mathbf{s}_u$ such that $\|\mathbf{s}_u\|_2^1 < 1$ in case $\|\mathbf{u}_i\|_2^2 = 0$ [7]. Thus the closed-form solution of (6) for the i th sub-group at the n th iteration can be written as

$$\begin{aligned} \mathbf{u}_i^{[n]} = & \left(\|\mathcal{S} \left(\mathbf{x}_i^{[n-1]} + \frac{\boldsymbol{\rho}_{u_i}^{[n-1]}}{c_u}, \frac{\lambda_e}{c_u} \right)\|_2^2 - \frac{\lambda_g}{c_u} \right)_+ \\ & \times \frac{\mathcal{S} \left(\mathbf{x}_i^{[n-1]} + \frac{\boldsymbol{\rho}_{u_i}^{[n-1]}}{c_u}, \frac{\lambda_e}{c_u} \right)}{\|\mathcal{S} \left(\mathbf{x}_i^{[n-1]} + \frac{\boldsymbol{\rho}_{u_i}^{[n-1]}}{c_u}, \frac{\lambda_e}{c_u} \right)\|_2^2} \end{aligned} \quad (11)$$

for $0 \leq i \leq G-1$, where $\mathcal{S}(\mathbf{s}, \lambda) \triangleq \text{sign}(\mathbf{x})(\mathbf{x} - \lambda)_+$ is the soft thresholding operator. Thus the estimate of $\mathbf{u}$ can be obtained as

$$\mathbf{u}^{[n]} = [\mathbf{u}_0^{[n]T}, \mathbf{u}_1^{[n]T}, \dots, \mathbf{u}_{G-1}^{[n]T}]^T \quad (12)$$

which along with $\mathbf{x}^{[n]}$ is used to update $\boldsymbol{\rho}_u^{[n]}$ in (8). Now from (7), the closed-form expression for the estimate of $\mathbf{z}$ at the n th iteration can be derived as

$$\mathbf{z}^{[n]} = \mathcal{S} \left(\mathbf{D}\mathbf{x}^{[n-1]} + \frac{\boldsymbol{\rho}_z^{[n-1]}}{c_z}, \frac{\lambda_f}{c_z} \right) \quad (13)$$

which subsequently updates $\boldsymbol{\rho}_z^{[n]}$ in (9).

2.2. Latent Group LASSO with Fusion

For the latent group fused LASSO (LGF-LASSO), the signal is segmented into many overlapping groups of certain sizes¹. In contrast to the disjoint groups, overlapping groups can reselect the elements from other groups. We create G overlapping groups through an $N/G \times N$ sub-selection matrix $\mathbf{W}_i$ which selects N/G rows from an identity matrix $\mathbf{I}_N$. An overlapping group can then be obtained by the relation, $\mathbf{W}_i \mathbf{x}$, for $0 \leq i \leq G-1$, where $\mathbf{W}_i$ is such

¹In this paper, we consider overlapping groups of fixed sizes, but the concept can easily be extended to varying sizes as well.

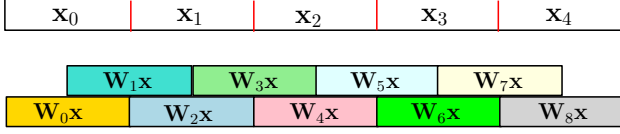


Fig. 1. Above: Disjoint groups. Below: Overlapping groups.

that $\mathbf{W} \triangleq [\mathbf{W}_0^T, \mathbf{W}_1^T, \dots, \mathbf{W}_{\tilde{G}}^T]^T$. Each sub-selection matrix $\mathbf{W}_i$ repeats K rows of $\mathbf{W}_{i-1}$, where K is the overlapping factor and $1 \leq K \leq N - 1$. Figure 1 schematically shows the difference between disjoint ($K = 0$) and overlapping groups (for $K = N/(2G)$). We can see that the overlapping groups can solve the problem of the fixed group size but the price to be paid is in terms of computational complexity which increases excessively with the factor K due to the related increase in $\tilde{G}$. Now, the optimization problem for LGF-LASSO can be formulated as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_g \sum_{i=0}^{\tilde{G}-1} \|\mathbf{W}_i \mathbf{x}\|_2 + \lambda_f \|\mathbf{D} \mathbf{x}\|_1. \quad (14)$$

Solver for LGF-LASSO

To solve the LGF-LASSO problem, we again turn to ADMM. By introducing a new auxiliary variable $\tilde{\mathbf{u}}$ of size $\tilde{G}N/G$, (14) can be written as

$$\begin{aligned} [\hat{\mathbf{x}}, \hat{\tilde{\mathbf{u}}}, \hat{\mathbf{z}}] &= \arg \min_{\mathbf{x}, \tilde{\mathbf{u}}, \mathbf{z}} \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_g \sum_{i=0}^{\tilde{G}-1} \|\tilde{\mathbf{u}}_i\|_2 + \lambda_f \|\mathbf{z}\|_1 \\ \text{s.t. } \tilde{\mathbf{u}}_i &= \mathbf{W}_i \mathbf{x}, \quad 0 \leq i \leq \tilde{G} - 1, \quad \mathbf{z} = \mathbf{D} \mathbf{x} \end{aligned} \quad (15)$$

where $\tilde{\mathbf{u}}_i$ is an $N/G \times 1$ sub-vector of $\tilde{\mathbf{u}}$, i.e., $\tilde{\mathbf{u}} = [\tilde{\mathbf{u}}_0^T, \tilde{\mathbf{u}}_1^T, \dots, \tilde{\mathbf{u}}_{\tilde{G}-1}^T]^T$. Now the Lagrangian for (15) can be written as

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \tilde{\mathbf{u}}, \mathbf{z}, \boldsymbol{\rho}_{\tilde{\mathbf{u}}}, \boldsymbol{\rho}_{\mathbf{z}}) &= \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda_g \sum_{i=0}^{\tilde{G}-1} \|\tilde{\mathbf{u}}_i\|_2 + \lambda_f \sum_{j=2}^N \|\mathbf{z}\|_1 \\ &+ \sum_{i=0}^{\tilde{G}-1} \boldsymbol{\rho}_{\tilde{\mathbf{u}}_i}^T (\tilde{\mathbf{u}}_i - \mathbf{W}_i \mathbf{x}) + \frac{c_u}{2} \sum_{i=0}^{\tilde{G}-1} \|\tilde{\mathbf{u}}_i - \mathbf{W}_i \mathbf{x}\|_2^2 \\ &+ \boldsymbol{\rho}_{\mathbf{z}}^T (\mathbf{z} - \mathbf{D} \mathbf{x}) + \frac{c_z}{2} \|\mathbf{z} - \mathbf{D} \mathbf{x}\|_2^2 \end{aligned} \quad (16)$$

where $\boldsymbol{\rho}_{\tilde{\mathbf{u}}}$ collects the Lagrangian multipliers with sub-vectors $\boldsymbol{\rho}_{\tilde{\mathbf{u}}_i}$ for $0 \leq i \leq \tilde{G} - 1$. Now the successive approximations for the solution of (16) w.r.t. $\mathbf{x}$, $\tilde{\mathbf{u}}$ and $\boldsymbol{\rho}_{\tilde{\mathbf{u}}}$ can easily be obtained by solving

$$\mathbf{x}^{[n]} = \arg \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \tilde{\mathbf{u}}^{[n-1]}, \mathbf{z}^{[n-1]}, \boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[n-1]}, \boldsymbol{\rho}_{\mathbf{z}}^{[n-1]}) \quad (17)$$

$$\tilde{\mathbf{u}}^{[n]} = \arg \min_{\tilde{\mathbf{u}}} \mathcal{L}(\mathbf{x}^{[n-1]}, \tilde{\mathbf{u}}, \boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[n-1]}) \quad (18)$$

$$\boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[n]} = \boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[n-1]} + c_u (\mathbf{x}^{[n]} - \tilde{\mathbf{u}}^{[n]}) \quad (19)$$

whereas, the estimates of $\mathbf{z}$ and $\boldsymbol{\rho}_{\mathbf{z}}$ are the same as in (7) and (9), respectively.

3. SIMULATIONS

In this section, we present some simulation results to compare the performance of our proposed algorithms. We compare the performance of SGF-LASSO, LGF-LASSO and G-LASSO. We consider

a cloud reflectivity data from the Earth System Research Laboratory (ESRL) [12]. This data basically shows variations in cloud reflectivity over time (around 12 hours) for different cloud heights above ground level (AGL). We consider it to be the ground truth and try to estimate it in the presence of noise of variance $\sigma^2 = 0.25$. There are $N = 425$ reflectivity samples corresponding to each height. We limit ourselves to the data of $N_{CH} = 50$ levels of cloud heights. A signal $\mathbf{x}$ is sensed separately per height, through the same measurement matrix Φ (where each row may correspond to a sensor), which has been drawn from a zero-mean Gaussian distribution with variance $1/M$. We have further orthogonalized the rows of the measurement matrix Φ .

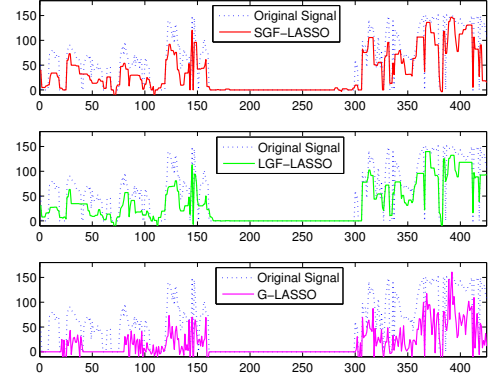


Fig. 2. Comparison of SGF-LASSO, LGF-LASSO and G-LASSO

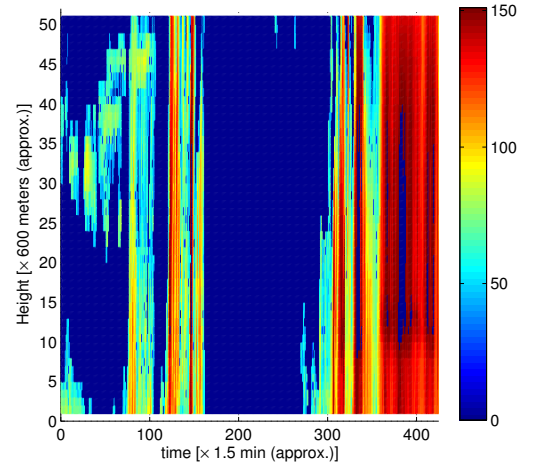


Fig. 3. Original Signal

The penalty parameters for the simulations have been considered as $\lambda_e = 5$, $\lambda_g = 35$ and $\lambda_f = 10$. In general, these parameters can be selected from a given range in a cross-validation manner, by varying one of the parameters and keeping others fixed [7]. Further, since all of these parameters are sparsity promoting, and can possibly affect each other, it is expected that the search of the optimal set of parameters would be restricted to a smaller range. The parameters c_u and c_z are positive numbers and may affect the convergence rate. We take them as $c_u = c_z = 2$. As initial conditions, the vectors $\mathbf{x}^{[0]}$, $\mathbf{u}^{[0]}$, $\mathbf{z}^{[0]}$, $\boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[0]}$, $\boldsymbol{\rho}_{\mathbf{z}}^{[0]}$, $\tilde{\mathbf{u}}^{[0]}$ and $\boldsymbol{\rho}_{\tilde{\mathbf{u}}}^{[0]}$ have all been considered as zero

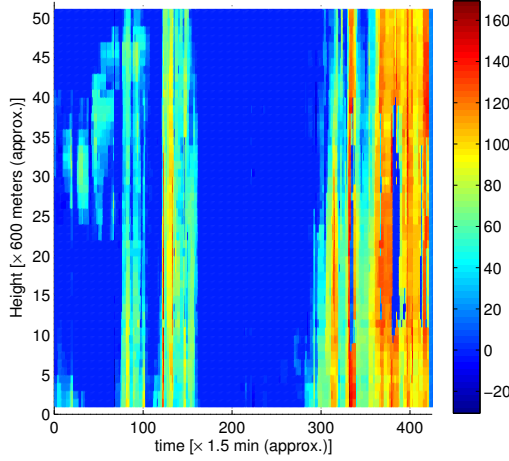


Fig. 4. Reconstruction by SGF-LASSO

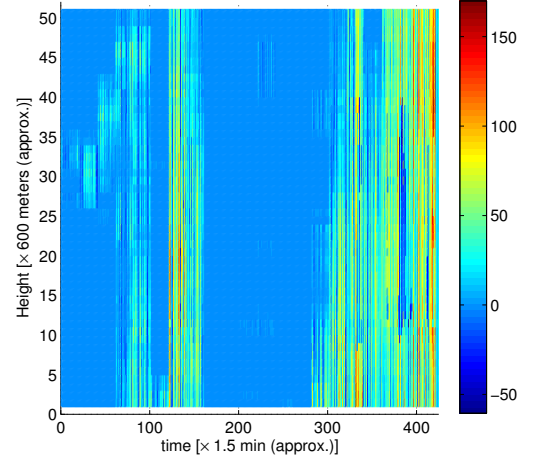


Fig. 6. Reconstruction by G-LASSO

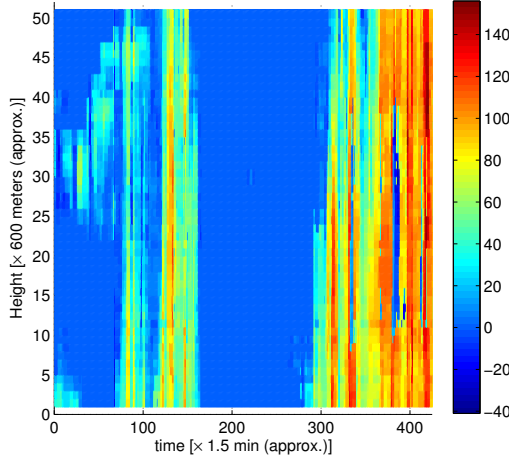


Fig. 5. Reconstruction by LGF-LASSO

vectors, respectively. Note that, a least-squares solution of $\mathbf{x}$, can also be considered as a warm-start to speed up the convergence rate. The group size for SGF-LASSO, LGF-LASSO and G-LASSO has been taken as 20. Therefore, the number of groups in SGF-LASSO and G-LASSO are the same, i.e. 21. For LGF-LASSO, an overlapping factor of $K = 5$ has been used, and therefore the number of overlapping groups of size 20 are $\tilde{G} = 28$. We use a maximum of 250 iterations for each algorithm. We have observed that a tolerance level of 10^{-3} between consecutive updates is reached much earlier than this limit, and therefore we stop the algorithm at this stage. Figure 2 shows the reconstruction performance of SGF-LASSO, LGF-LASSO and G-LASSO for a particular cloud height, when the signal was sensed with a compression ratio $\mu = 0.5$. We can see that the performance of SGF-LASSO and LGF-LASSO is very close to each other and both are able to recover the smooth transitions of the original signal. On the other hand, the performance of G-LASSO deteriorates both on the front of smoothness as well as block size. Note that in contrast to SGF-LASSO and LGF-LASSO, λ_g is the only sparsity creating parameter for G-LASSO. Therefore, we increase its value to 122.5, which is the minimum to recreate the actual zero blocks. Also in case of SGF-LASSO, we take $\lambda_g = 17.5$

Table 1. MSE comparisons w.r.t. compression ratio

μ	SGF-LASSO	LGF-LASSO	G-LASSO
0.1	0.4607	0.4523	0.4953
0.3	0.2589	0.2607	0.4122
0.5	0.1661	0.1607	0.3079
0.7	0.1250	0.1197	0.2576

in order to facilitate the parsimonious effect of λ_e . Figures 3-6 show the reconstruction performance of SGF-LASSO, LGF-LASSO and G-LASSO for the complete range of cloud heights. Again, we can see that the performance of SGF-LASSO and LGF-LASSO is better than G-LASSO and very close to the original. Table 1 shows the performance comparison of the proposed algorithms through the mean squared error (MSE) metric against varying compression ratios, $\text{MSE} \triangleq \mathbb{E}\{\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\|_2^2 / NN_{CH}\}$, where $\tilde{\mathbf{x}}$ is the concatenation of N_{CH} signals $\mathbf{x}$ (i.e., of all cloud heights), and average ($\mathbb{E}\{\cdot\}$) is over different noise realizations. We can see that the performance improves in general with increasing value of μ , for $0.1 \leq \mu \leq 0.7$. Nonetheless, the difference in performance follows the previously observed pattern. Note that the performance of LGF-LASSO can be improved by increasing the overlapping factor but that would cause a subsequent increase in the computational complexity.

4. CONCLUSIONS

In this paper, we have proposed two new LASSO formulations, namely, sparse group fused LASSO and latent group fused LASSO. The former uses element-wise sparsity, group sparsity (over disjoint groups) and fusion penalties, whereas the latter combines the fusion penalty with a latent group penalty. Both formulations can be used to reconstruct smooth signals from their compressed measurements. We also provide low-complexity solvers for the proposed formulations, based on the alternating direction method of multipliers. We compared the performance of our proposed algorithms with standard group LASSO over a smooth test signal. The simulation results confirm the better performance of the proposed algorithms for signal reconstruction against group LASSO. Similar results were obtained for the mean squared error metric, for varying compression ratios.

5. REFERENCES

- [1] David L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [2] E. Candès, J. Romberg and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transaction on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.
- [3] S. S. Chen, D. L. Donoho and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM journal on Scientific Computing*, vol. 43, no. 1, pp. 129–159, 2001.
- [4] R. Tibshirani, “Regression shrinkage and selection via the lasso”, *J. Royal Statist. Soc B.*, vol. 58, no. 1, pp 267–288.
- [5] J. A. Tropp, A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. on Info. Theory*, vol. 53, no. 12, pp. 4655–4666, December 2007.
- [6] M. Yuan, Y. Lin, “Model selection and estimation in regression with grouped variables,” *J. Royal Statist. Soc B.*, vol. 68, no. 1, pp. 49–67.
- [7] J. Friedman, T. Hastie and R. Tibshirani, “A note on the group lasso and a sparse group lasso,” Technical report, Department of Statistics, Stanford University, 2010.
- [8] R. Tibshirani, S. Rosset, J. Zhu and K. Knight, “Sparsity and smoothness via the fused lasso,” *J.R. Statist. Soc. B.*, vol. 67, pp. 91–108, 2005.
- [9] G. Obozinski, L. Jacob and J.-P. Vert, “Group lasso with overlaps: the latent group lasso approach,” Technical report, 2011.
- [10] I. W. Selesnick and Po-Yu Chen, “Total variation denoising with overlapping group sparsity,” *IEEE ICASSP*, 2013.
- [11] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, 1997.
- [12] NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, <http://www.esrl.noaa.gov/psd/>.