Iterative Re-weighted Least Squares Algorithms for Non-negative Sparse and Group-sparse Recovery

Angshul Majumdar
IIIT Delhi
angshul@iiitd.ac.in

***Abstract*— Algorithms for non-negative sparse recovery are either based on modifications of orthogonal matching pursuit or are based on thresholding of non-negative least squares. Both are variants of techniques proposed for sparse recovery. This work is based on the iterative re-weighted least squares (IRLS) approach for sparse recovery. IRLS has been found to be a simple yet versatile approach that can handle both l_1-norm and l_p-quasi norm (0<p<1). We extend this approach to handle not only sparse recovery but also group-sparse recovery.**

Keywords*— *non-negative sparse recovery, non-negative group sparsity,

I. Introduction

Recovery of a sparse solution to a linear inverse problem has been widely studied in signal processing. Related to this topic, is the problem of solving an inverse problem where the solution is non-negative and sparse. Our interest lies in the said problem – that of non negative sparse recovery.

The general form of a linear inverse problem is as:

$$y_{m\times 1} = A_{m\times n} x_{n\times 1} \quad (1)$$

For a sparse k-sparse solution x will have k non-zero entries and n-k zeroes. Broadly, there are two approaches to solve the sparse recovery problem – 1. Greedy techniques, and, 2. Optimization based techniques.

Greedy techniques generally start with an all-zero solution and iteratively estimate the support of the solution and the corresponding values. Popular techniques include orthogonal matching pursuit (OMP) [1], a greedier variant of OMP called stage-wise OMP [2] and CoSamp [3].

Optimization based approaches largely fall in the category of iterative thresholding, where they start with a dense solution and progressively threshold it to reach a sparse solution. Iterative soft thresholding for l_1-norm minimization [4] and iterative hard thresholding for l_0-norm minimization [5] are two such examples.

Within the optimization based approaches, there is another comparatively lesser known class of techniques that are based on iterative re-weighted least squares (IRLS) [6-8]. IRLS is versatile, easy to derive, and can solve l_p-norm minimization $0\leq p\leq 1$. Consequently, our proposed approach to solve the non-negative sparse recovery problem will be based on this approach.

Closely related to the problem of sparse recovery is the topic of block-sparse [9] / group-sparse [10] recovery. In this problem, groups of coefficients are always non-zero together. The solution 'x' can be assumed to be formed of several non-overlapping groups. The solution is group-sparse in the sense that the number of groups that are non-zero is very small, but within the non-zero group all the coefficients are non-zeroes.

The aforesaid problem too has been solved in two approaches. One being greedy such as [11, 12] and the other being optimization based [13, 14]. To the best of our knowledge, there has only been a single work [15] that showed how IRLS can be used to solve the group-sparse recovery problem.

In this work our interest is in solving a sparse system of equations where the solution is non-negative. A few greedy solutions to this problem (based on OMP) have been reported in the literature [16, 17]. Thresholding based solutions to the non-negative sparse recovery problem have also been reported [18, 19]. To the best of our knowledge, IRLS techniques have not been employed for this problem. An interesting fact about this problem is that for a particular case, where A is drawn from the Bernoulli distribution, the sparse solution can be obtained without any kind of sparsity constraint – a simple non-negative least squares solver recovers the solution with overwhelming probability [20, 21].

Non-negative sparse recovery is not as widely studied as general sparse recovery. There are a handful of papers on this topic and we have discussed most of them. We could not find papers on the non-negative structured sparse recovery, including block / group sparsity.

IRLS has been a widely used technique in a variety of signal processing topics ranging from filter design [22] to sparse recovery (already discussed) to matrix completion [23]. IRLS has also been used for solving co-sparse recovery problems such as total variation regularized least squares [24]. Given the versatility of the IRLS technique we solve the non-negative sparse and group-sparse recovery problems using it.

II. Proposed Solution

A. General approach to IRLS solution

We will first discuss the IRLS algorithm for general sparse recovery. This affine scaling methodology is derived in [6]. Since it is essential for understanding our contribution, we give the general derivation for this approach. The general problem can be expressed as –

$$\min E(x) \text{ subject to } y = Ax \quad (1)$$

where $E(x)$ is a diversity measure.

The Lagrangian for (4) is

$$L(x,\lambda) = E(x) + \lambda^T (Ax - b) \quad (2)$$

Following the theory of Lagrangian, the stationary point for (5) needs to be obtained by solving the following,

$$\nabla_x L(\hat{x},\hat{\lambda}) = 0 \quad (3a)$$

$$\nabla_\lambda L(\hat{x},\hat{\lambda}) = 0 \quad (3b)$$

Solving (3) requires the gradient of $E(x)$ with respect to x. Please note that, in order to preserve generality, the explicit functional form of $E(x)$ has not been stated. In the diversity measures of interest to the signal processing community, the gradient of $E(x)$ can be expressed as,

$$\nabla_x E(x) = \alpha(x)\Pi(x)x \quad (4)$$

where $\alpha(x)$ is a scalar, and $\Pi(x)$ is a diagonal matrix.

Solving from (3) and (4) yields –

$$\hat{x} = -\frac{1}{\alpha(\hat{x})}\Pi(\hat{x})^{-1}A^T\hat{\lambda} \quad (5a)$$

$$\hat{\lambda} = -\alpha(x)(A\Pi(\hat{x})^{-1}A^T)^{-1}y \quad (5b)$$

Substituting (5b) into (5a) leads to –

$$\hat{x} = \Pi(\hat{x})^{-1}A^T(A\Pi(\hat{x})^{-1}A^T)^{-1}y \quad (6)$$

The expression (6) is an implicit function of x. Therefore it needs to be solved iteratively in the following manner:

$$x^{(k+1)} = \Pi(x^{(k)})^{-1}A^T(A\Pi(x^{(k)})^{-1}A^T)^{-1}y \quad (7)$$

Note that the solution (7) is akin to that of a weighted least squares solution. Since the weights are changing in every iteration, an iterative re-weighting is required.

We have made the notation as similar as possible to the original work [6]. There is a basic problem however with this approach. The theory of linear Lagrangian used to derive this, is valid for convex problems only. However, FOCUSS has been widely used for solving non-convex diversity measures such as l_p-norm minimization ($0 \le p \le 1$) [7]. There is no theoretical guarantee that this formulated approach will lead to a solution to the non-convex problem but in practice, it works very well.

B. IRLS solution to non-negative sparse recovery

$$\min_x \|x\|_p^p \text{ subject to } y = Ax \quad (8)$$

Before going into the solution for the non-negative sparse recovery problem, we briefly discuss the solution to the sparse recovery problem without the non-negativity constraints (8). The solution (albeit following a more heuristic approach) is given in [7]. The diversity measure for sparsity is the l_p-norm ($0 \le p \le 1$).

We skip the derivation (one can refer [6, 7] for it) and directly write the algorithm.

Initialize: $R = Identity$

Repeat till convergence

$\Phi = AR$

$z = \min \|y - \Phi z\|_2^2$

$x_{k+1} = Rz$

$R = diag\left\{\left(abs\left(x^{(k+1)}\right)^{p-2}\right)^{-1/2}\right\}$

Here local convergence is supposed to be achieved when the values of the iterate do not change substantially in subsequent iterations.

$$\min_{x \ge 0} \|x\|_p^p \text{ subject to } y = Ax \quad (9)$$

Our problem is to modify the aforementioned algorithm in such a fashion that the solution x is always positive (9). Note that in every iteration, x is formed as: $x = Rz$. By definition, R is always positive. Therefore, as long as z can be enforced to be non-negative we will achieve the desired solution. The solution to z comes by solving a least squares problem. We propose to replace it with a non-negative least squares (NNLS) solver; in particular, we use the *lsqnonneg* function in Matlab™. Succinctly our algorithm is expressed as follows.

Initialize: $R = Identity$

Repeat till convergence

$\Phi = AR$

$z = NNLS(\Phi, y)$

$x_{k+1} = Rz$

$R = diag\left\{\left(abs\left(x^{(k+1)}\right)^{p-2}\right)^{-1/2} + \varepsilon\right\}$

Note that we have used a slightly different definition of R in the algorithm. The optimization problem being non-convex, there is a chance that the algorithm converges to local minima. To reduce the chances of converging to a local minimum, a simple modification has been proposed in [7]; that of adding a damping factor. Following this suggestion, we apply a damping factor 'δ'. The value of the damping factor is progressively reduced in subsequent iterations.

Since the goal is to obtain a sparse solution, it is likely that there will be positions in the iterate that have (approximately) zero values. The damping factor bounds the values of the diagonal matrix and does not allow its inverse to become excessively large.

C. IRLS solution to non-negative group-sparse recovery

The convex diversity measure for the group-sparse optimization problem is the $l_{2,1}$-norm; the vector $\underline{x}$ is assumed to be composed of C groups $x = [x_1, x_2, ..., x_C]^T$ and each group is defined as $x_i = [x_{i,1}, ..., x_{i,n_i}]^T$. On this, the $l_{2,1}$-norm is defined as,

$$\|x\|_{2,1} = \sum_{i=1}^{C} \|x_i\|_2 \quad (10)$$

The geometrical justification for employing such a mixed norm $\|.\|_{2,1}$ for recovering group-sparse solutions is well known. The inner l_2-norm in $\|\cdot\|_{2,1}$ enforces selection of all the coefficients within a group while the outer l_1-norm promotes sparsity in the number of selected groups.

$$\min_x \|x\|_{2,1} \text{ subject to } y = Ax \tag{11}$$

The standard formulation to solve the group-sparse recovery problem is (11). In this work, we generalize it and define the group-sparse diversity measure as follows –

$$\|x\|_{m,p} = \left(\sum_{i=1}^{C} \|x_i\|_m^p \right)^{1/p} \tag{12}$$

The minimizer for (12) is the same as the minimizer of

$$\|x\|_{m,p}^p = \left(\sum_{i=1}^{C} \|x_i\|_m^p \right) \tag{13}$$

$$\min_x \|x\|_{m,p}^p \text{ subject to } y = Ax \tag{14}$$

For mathematical convenience we will use the latter form as the divergence measure to solve the group-sparse problem (14). For each of the terms in (14), the derivative is calculated to be

$$\frac{\partial \|x\|_{m,p}^p}{\partial x_{i,j}} = p\|x_i\|_m^{p-m} \left(\left|x_{i,j}\right|^{m-2} \cdot x_{i,j} \right) \tag{15}$$

Here '.' implies element-wise product.

Comparing (12) with (7), we have $\alpha(x) = p$ and $\Pi(x) = diag\left(\|x_i\|_m^{p-m} \left(\left|x_{i,j}\right|^{m-2} \cdot x_{i,j} \right) \right)$

From a practical perspective, we add the damping factor to the group-sparse solution as well, i.e.

$$\Pi(x) = diag\left(\|x_i\|_m^{p-m} \left(\left|x_{i,j}\right|^{m-2} \cdot x_{i,j} \right) + \delta \right) \tag{16}$$

In a succinct form, the algorithm for group-sparse recovery (GSR) can be written as follow.

Initialize: $R = Identity$
Repeat till convergence
$\Phi = AR$
$z = \min \|y - \Phi z\|_2^2$
$x_{k+1} = Rz$
$R = diag\left(\left(\|x_i^{(k+1)}\|_m^{p-m} \left|x_{i,j}^{(k+1)}\right|^{m-2} + \delta \right)^{-1/2} \right)$

$$\min_{x \geq 0} \|x\|_{m,p}^p \text{ subject to } y = Ax \tag{17}$$

So far we have derived the solution for IRLS based group-sparse recovery; our goal is to solve it for the non-negative case (17). As in the case of sparse recovery one notices that R is always positive, therefore to enforce non-negativity on the solution it is enough to have z to be non-negative in each iteration. This is simply achieved by solving the least squares problem in the algorithm by NNLS. Our final algorithm for non-negative group-sparse recovery (NNGSR) is written in a succinct fashion as follows.

Initialize: $R = Identity$
Repeat till convergence
$\Phi = AR$
$z = NNLS(\Phi, y)$
$x_{k+1} = Rz$
$R = diag\left(\left(\|x_i^{(k+1)}\|_m^{p-m} \left|x_{i,j}^{(k+1)}\right|^{m-2} + \delta \right)^{-1/2} \right)$

III. Experimental Results

A. *Sparse Recovery*

In the first set of experiments, we benchmark our technique with existing non-negative sparse recovery algorithms. We compare with one greedy approach called Fast Non-Negative Orthogonal Matching Pursuit (FNNOMP) [15] and one thresholding / shrinkage approach called Thresholded Non-negative Least Squares (TNNLS) [16]. We also compare with the Iterative Re-weighted Least Squares (IRLS) algorithm for sparse recovery (IRLS) [7]; this is because our proposed technique is based on the same.

One standard protocol for evaluating algorithms for sparse recovery is to see how they perform when the number of equations / measurements are progressively reduced. The metric for performance, in this case, is the empirical success rate. In such a case a trial is assumed to be successful if the normalized mean squared error (NMSE) is less than a pre-defined threshold; a standard practice is to have a threshold of 10^{-3}.

For all our experiments, the system of equations 'A' is generated from i.i.d Gaussian random distribution. For all the problems 10% of the variables in x are kept as positive values; the values are drawn from a uniform distribution. The size of x is 1000 (including zero and positive values).

The results for the first set of experiments are shown in Fig. 1. Under-sampling ratio is the ratio of the number of equations to the number of variables. This ratio is progressively reduced and the empirical success rate is plotted. Success rate is the ratio of the number of successful trials (defined before) to the total number of trials. We have conducted 1000 trials for each under-sampling ratio.

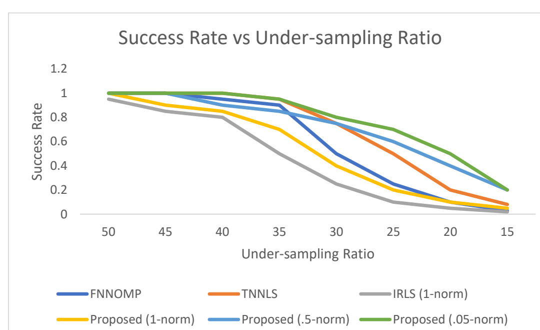

Fig. 1. Plot to evaluate different algorithms for non-negative sparse recovery with varying under-sampling ratios.

In Fig.1 we see that our proposed algorithm for l_1-minimization performs slightly worse than the thresholding and greedy techniques, but is better than the IRLS based l_1-minimization algorithm without non-negativity constraints. However, as we move towards the ideal l_0-norm minimization, by reducing the value of p in l_p-minimization, we see that the results improve. By improvement, we mean that the success rates deteriorate much more gradually for our proposed algorithm compared to the benchmark FNNOMP and the TNNLS. The proposed IRLS technique shows a much higher success rate than existing techniques when the number of equations falls drastically.

In the next set of experiments we keep the under-sampling ratio to be constant at 50% but vary the number of positive values in the variable x. The results are shown in Fig. 2. We see that the implications are somewhat similar to that of the previous set of experiments. For l_1-minimization using our proposed method, the results are subpar compared to the existing benchmarks (but better than IRLS). When the value of p is reduced, the results improve drastically; this is especially true when the proportion of positive values in the solution increases. For example, even with 25% of positive values, our $l_{.05}$-minimization has more than 20% chance of succeeding; for existing benchmarks, the chances are close to zero in such a challenging scenario.

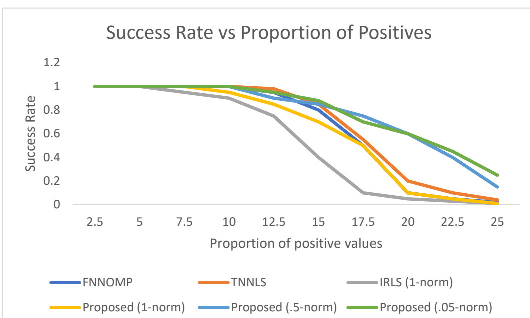

Fig. 2. Plot to evaluate different algorithms for non-negative sparse recovery with varying proportion of positive values.

B. *Group-sparse Recovery*

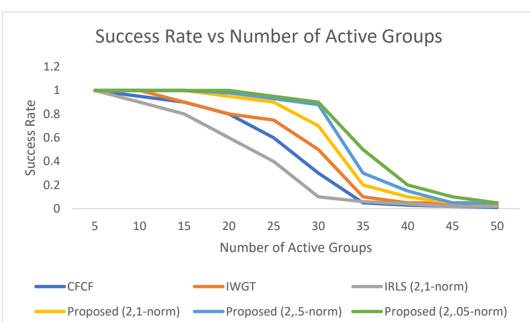

Fig. 3. Plot to evaluate different algorithms for non-negative group-sparse recovery with varying proportions of active groups.

To the best of our knowledge, there are no existing algorithms for recovering non-negative group-sparse solutions. Therefore our algorithm does not have a direct benchmark. What we can do is to compare our proposed GSR and NNGSR algorithms with existing benchmarks for group-sparse recovery. We consider two of the latest techniques as benchmarks; they are – capped folded concave functions (CFCF) [25] and iterative weighted group thresholding (IWGT) [26].

As in the case of sparse recovery, the system of equations '*A*' is drawn from i.i.d Gaussian distribution. As before, the number of variables to be solved is 1000; this is comprised of 100 groups where the size of groups varies between 8-12. In the first set of experiments, the success rate is plotted with a varying number of active groups. A group is said to be active if all the values in the group are positive. The under-sampling ratio is kept fixed at 50%. 1000 trials were conducted for each number of active groups.

Fig. 3 shows that our proposed algorithm yields the best results. This is expected since the existing algorithms are not tailored for non-negative group-sparse recovery. We find that even with $l_{2,1}$-minimization our method outperforms the rest. The results continue to improve when the value of p is reduced in $l_{2,p}$-minimization.

In the final set of experiments, we keep the number of active groups to be fixed at 10. We change the sub-sampling ratio to see how the success rate changes with it. The results are shown in Fig. 4. The results are as expected. Our proposed method yields the best results; improvement is seen when the value of p is reduced to almost 0.

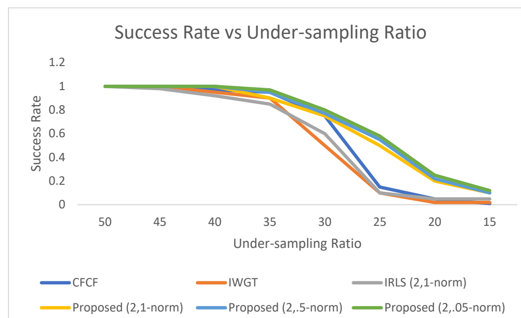

Fig. 4. Plot to evaluate different algorithms for non-negative group-sparse recovery with varying under-sampling ratios.

IV. CONCLUSION

This work addresses the problem of non-negative sparse and group-sparse recovery. The problem of non-negative sparse recovery has been addressed before, but non-negative group-sparse recovery is a new problem. Even though solutions exist for the sparse recovery problem, they are either based on greedy algorithms or are based on thresholding. This is the first work, that proposes to solve it via IRLS. The advantage of IRLS is that it can handle both the convex l_1-norm as well as the l_p-(quasi) norms. The problem of non-negative group-sparse recovery has not been addressed before. We provide the first solution to the problem based on the IRLS approach.

Experimental results show that for sparse recovery our work is marginally worse than existing non-negative sparse recovery algorithms. Since there are no algorithms for non-negative sparse we could not compare with any. However, the problem is of utmost importance since it arises in many practical scenarios like classification [27-31] and inverse problems in imaging [32-35].

REFERENCES

[1] Y. C. Pati, R. Rezaiifar and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, 1993, pp. 40-44.

[2] D. L. Donoho, Y. Tsaig, I. Drori and J. Starck, "Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit," IEEE Transactions on Information Theory, vol. 58, no. 2, pp. 1094-1121, 2012.

[3] D. Needell and J. A. Tropp. "CoSaMP: iterative signal recovery from incomplete and inaccurate samples." Communications of the ACM vol. 53, no. 12, pp. 93-100, 2010.

[4] I. Daubechies, M. Defrise, and C. De Mol, "An iterativethresholding algorithm for linear inverse problems with a sparsity constraint,"Comm. Pure Appl. Math., vol. 57, no. 11, pp. 1413–1457, 2004.

[5] T. Blumensath, M. Yaghoobi and M. E. Davies, "Iterative Hard Thresholding and L0 Regularisation," 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, 2007, pp. III-877-III-880.

[6] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," in IEEE Transactions on Signal Processing, vol. 47, no. 1, pp. 187-200, Jan. 1999.

[7] R. Chartrand and Wotao Yin, "Iteratively reweighted algorithms for compressive sensing," 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 2008, pp. 3869-3872.

[8] I. Daubechies, R. DeVore, M. Fornasier and C. S. Güntürk. "Iteratively reweighted least squares minimization for sparse recovery." Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, vol. 63, no. 1, pp. 1-38, 2010.

[9] Y. C. Eldar, P. Kuppinger and H. Bolcskei, "Block-Sparse Signals: Uncertainty Relations and Efficient Recovery," in IEEE Transactions on Signal Processing, vol. 58, no. 6, pp. 3042-3054, June 2010.

[10] N. Simon, J. Friedman, T. Hastie, and Robert Tibshirani. "A sparse-group lasso." Journal of computational and graphical statistics, vol. 22, no. 2, pp. 231-245, 2013.

[11] J. Wen, Z. Zhou, Z. Liu, M.-J. Lai, and X. Tang, "Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonalmatching pursuit,"Appl. Comput. Harmon. Anal., vol. 47, no. 3, pp.948-974 2018.

[12] G. Swirszcz, N. Abe and A. C. Lozano. "Grouped orthogonal matching pursuit for variable selection and prediction." Advances in Neural Information Processing Systems, vol. 22, pp. 1150-1158, 2009.

[13] P. Y. Chen and I. W. Selesnick. "Translation-invariant shrinkage/thresholding of group sparse signals." Signal Processing, vol. 94, pp. 476-489, 2014.

[14] R. Chartrand and B. Wohlberg, "A nonconvex ADMM algorithm for group sparsity with sparse groups," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 6009-6013.

[15] C. Chen, J. Huang, L. He and H. Li, "Preconditioning for Accelerated Iteratively Reweighted Least Squares in Structured Sparsity Reconstruction," 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2713-2720.

[16] M. Yaghoobi, D. Wu and M. E. Davies, "Fast Non-Negative Orthogonal Matching Pursuit," in IEEE Signal Processing Letters, vol. 22, no. 9, pp. 1229-1233, Sept. 2015.

[17] T. T. Nguyen, J. Idier, C. Soussen and E. Djermoune, "Non-Negative Orthogonal Greedy Algorithms," in IEEE Transactions on Signal Processing, vol. 67, no. 21, pp. 5643-5658, 1 Nov.1, 2019.

[18] M. Slawski and M. Hein. "Sparse recovery by thresholded non-negative least squares," Advances in Neural Information Processing Systems, vol. 24, pp. 1926-1934, 2011.

[19] M. Yuan and Y. Lin, "On the non-negative garrotte estimator,"J. Roy.Statist. Soc. B, Statist. Methodology, vol. 69, no. 2, pp. 143–161,Apr. 2007.

[20] S. Foucart and D. Koslicki, "Sparse Recovery by Means of Nonnegative Least Squares," in IEEE Signal Processing Letters, vol. 21, no. 4, pp. 498-502, April 2014.

[21] R. Kueng and P. Jung, "Robust Nonnegative Sparse Recovery and the Nullspace Property of 0/1 Measurements," in IEEE Transactions on Information Theory, vol. 64, no. 2, pp. 689-703, Feb. 2018.

[22] C. S. Burrus, J. A. Barreto and I. W. Selesnick, "Iterative reweighted least-squares design of FIR filters," in IEEE Transactions on Signal Processing, vol. 42, no. 11, pp. 2926-2936, Nov. 1994.

[23] M. Fornasier, H. Rauhut, and R. Ward. "Low-rank matrix recovery via iteratively reweighted least squares minimization." SIAM Journal on Optimization, vol. 21, no. 4, pp. 1614-1640, 2011.

[24] P. Rodriguez and B. Wohlberg, "Efficient Minimization Method for a Generalized Total Variation Functional," in IEEE Transactions on Image Processing, vol. 18, no. 2, pp. 322-332, Feb. 2009.

[25] Pan, L., & Chen, X. (2021). Group sparse optimization for images recovery using capped folded concave functions. SIAM Journal on Imaging Sciences, 14(1), 1-25.

[26] L. Jiang and W. Zhu, "Iterative Weighted Group Thresholding Method for Group Sparse Recovery," in IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 63-76, Jan. 2021.

[27] A. Majumdar and R. K. Ward, "Non-linear Sparse and Group-sparse Classifier", IEEE CCECE 2013.

[28] A. Majumdar and R. K. Ward, "Synthesis and Analysis Prior Algorithms for Joint-Sparse Recovery", IEEE ICASSP, pp. 3421-3424, 2012.

[29] A. Majumdar and R. K. Ward, "Classification via Group Sparsity Promoting Regularization", IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 873-876, 2009.

[30] A. Majumdar and R. K. Ward, "Improved Group Sparse Classifier", Pattern Recognition Letters, Vol. 31 (13), pp. 1959-1964, 2010.

[31] A. Majumdar and R. K. Ward, "Fast Group Sparse Classification", IEEE Canadian Journal of Electrical and Computer Engineering, Vol. 34 (4), pp. 136-144, 2009

[32] A. Majumdar and R. K. Ward, "Non-convex Group Sparsity: Application to Color Imaging", IEEE ICASSP, pp. 469-472, 2010.

[33] A. Majumdar and R. K. Ward, "Compressed Sensing of Color Images", Signal Processing, Vol. 90 (12), 3122-3127, 2010.

[34] A. Majumdar and R. K. Ward, "Accelerating Multi-Echo T2 Weighted MR Imaging: Analysis Prior Group Sparse Optimization", Journal of Magnetic Resonance, Vol. 210 (1), pp. 90-97, 2011.

[35] A. Majumdar and R. K. Ward, "Joint Reconstruction of Multi-echo MR Images Using Correlated Sparsity", Magnetic Resonance Imaging, Vol. 29 (7), pp. 899-906, 2011.