

Subspace based low rank & joint sparse matrix recovery

Sampurna Biswas, Sunrita Poddar, Soura Dasgupta, Raghuraman Mudumbai, and Mathews Jacob

Abstract: We consider the recovery of a low rank and jointly sparse matrix from under sampled measurements of its columns. This problem is highly relevant in the recovery of dynamic MRI data with high spatio-temporal resolution, where each column of the matrix corresponds to a frame in the image time series; the matrix is highly low-rank since the frames are highly correlated. Similarly the non-zero locations of the matrix in appropriate transform/frame domains (e.g. wavelet, gradient) are roughly the same in different frame. The superset of the support can be safely assumed to be jointly sparse. Unlike the classical multiple measurement vector (MMV) setup that measures all the snapshots using the same matrix, we consider each snapshot to be measured using a different measurement matrix. We show that this approach reduces the total number of measurements, especially when the rank of the matrix is much smaller than its sparsity. Our experiments in the context of dynamic imaging shows that this approach is very useful in realizing free breathing cardiac MRI.

Index Terms—Joint sparsity, Subspace estimation, ADMM, MMV, MRI reconstruction

I. INTRODUCTION

The recovery of matrices that are simultaneously low-rank and jointly sparse from its few measurements have received a lot of attention in the recent years. [1], [2], [3]. Our interest in this problem is motivated by its application in cardiovascular dynamic MRI, where the image frames are highly correlated. The cardiac signal can be modeled as a low rank matrix \mathbf{X} with each column as the list of n voxels and each row as the voxel time profile over N snapshots. Similarly the non-zero locations of the matrix in appropriate transform/frame domains (e.g. wavelet, gradient) are roughly the same in different frame [4]. Joint sparsity is a special case of group sparsity where only some groups have non zero coefficients. For a slowly varying temporal profile of the transformed data we also have joint sparsity. Our experiments show that the use of joint sparsity and low-rank priors (k-t SLR) provide a considerable reduction in the number of measurements required to recover the matrix. While there is considerable theoretical progress in problems

such as recovering jointly sparse vectors or low-rank matrices, the recovery of matrices that are simultaneously low-rank and jointly sparse have received considerably less attention.

In [5], Golbabe et al., have recently developed theoretical guarantees for the recovery of a matrix of rank r and has only k non-zero columns using low rank and joint sparsity priors from its random Gaussian dense measurements. Unfortunately, the dense measurement scheme, where each measurement is a linear combination of all matrix entries is not practical in dynamic imaging; each measurement can only depend on a single column of the matrix. In contrast, the classical MMV (multiple measurement vector) problem deals with the recovery of $\mathbf{X}^{n \times N}$ (N snapshots of length n) from its undersampled measurements $\mathbf{Y} = \mathbf{A}\mathbf{X}$. The matrix \mathbf{X} is assumed to have a rank of r , while the columns are assumed to be jointly sparse with a sparsity of k . The spark requirement on the measurement matrix \mathbf{A} to uniquely identify \mathbf{X} is $\text{spark}(\mathbf{A}) > 2k - r + 1$ [6]. While this condition is more relaxed than the identifiability conditions in the single measurement vector (SMV) setting ($\text{spark}(\mathbf{A}) > 2k$), the gains over SMV are marginal when $r \ll k$. This is not desirable since the columns of the matrix to be highly correlated and hence more compact; one would expect higher gains in this case. The main problem with this setting is the lack of diversity in the measurements.

We introduce a hybrid acquisition scheme, where each column \mathbf{x}_i is measured using the measurement matrix \mathbf{A}_i . All \mathbf{A}_i 's have a common submatrix Φ , while the remaining rows of \mathbf{A}_i specified by \mathbf{B}_i are different. We also introduce a novel two-step reconstruction strategy. In the first step, we estimate the row-subspace using the common measurements Φ . Once the subspace is estimated, the subspace aware recovery of the matrix simplifies to a least square problem. This approach is considerably simpler than convex optimization schemes using simultaneously structured priors. We introduce perfect recovery guarantees under simple conditions on the measurement matrices and the signal matrix. Our results show that perfect recovery of the matrix is possible when the number of measurements are of the order of the degrees of freedom in the low-rank and sparse matrix. Considering that the dense measurement scheme is not feasible in dynamic imaging applications, this work is quite significant.

The proposed framework to recovery of matrices that are simultaneously jointly-sparse and low-rank is ideally suited to accelerate free breathing cardiac CINE MRI data. The joint sparsity assumption implies that all the snapshots share the same sparsity pattern and thus the data matrix has only some significant rows and other rows are negligible. Very often, the rank r is much smaller than the joint sparsity k .

The authors are with the Department of Electrical and Computer Engineering, Univ. Iowa, IA, USA. (emails: {sampurna-biswas,soura-dasgupta,mathews-jacob}@uiowa.edu).

This work is in part supported by US NSF grants EPS-1101284, ECCS-1150801, CNS-1329657, CCF-1302456, CCF-1116067, NIH 1R21HL109710-01A1, ACS RSG-11-267-01-CCE, and ONR grant N00014-13-1-0202.

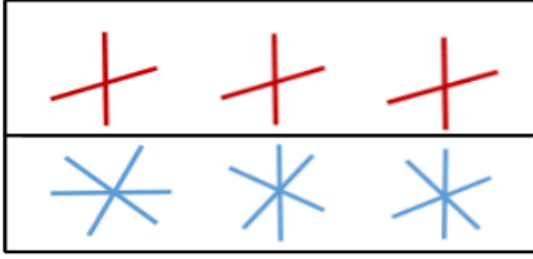


Fig. 1: Common (top) and variable (bottom) sampling lines

The slow nature of MR acquisitions restricts the number of measurements that can be acquired during each time frame. At the same time, the acquisition is flexible enough to allow the use of different measurement matrices to measure each time frame, without any additional penalty. We propose to capitalize this flexibility to considerably accelerate the acquisition. Specifically, we use the sampling pattern shown in Fig. 1 to acquire the data.

II. PROPOSED APPROACH

We propose an acquisition scheme, where the i^{th} frame is measured as

$$\underbrace{\begin{bmatrix} \mathbf{z}_i \\ \mathbf{y}_i \end{bmatrix}}_{\bar{\mathbf{y}}_i} = \underbrace{\begin{bmatrix} \Phi \\ \mathbf{A}_i \end{bmatrix}}_{\mathbf{A}_i} \mathbf{x}_i. \quad (1)$$

Note that the measurement matrix Φ is common for all snapshots. We propose to estimate the right subspace of \mathbf{X} from $\mathbf{Z}^H \mathbf{Z}$, where \mathbf{Z} is obtained from z_i . Once the right subspace of \mathbf{X} is estimated, we propose to perform a right subspace aware recovery of \mathbf{X} from the measurements $\mathbf{y}_i = \mathbf{A}_i \mathbf{x}_i$; $i = 0, \dots, N-1$. These are the variable measurements and they differ across frames.

A. Recovery of the right singular matrix

In [7] the authors look into the preservation of the right singular vectors of a matrix after it is treated by an undersampled matrix. They arrive at a condition for a sensing matrix that satisfies the JL lemma. If the rows of that sensing matrix are of the order of the rank of the matrix for the singular values and vectors to be approximately preserved with a certain failure probability. We look for a similar condition when the sparsity is also coming into picture.

Consider jointly sparse matrix \mathbf{X} and its truncated singular value decomposition, $\mathbf{U}\Sigma\mathbf{V}$ where Σ is diagonal, non singular matrix in $\mathbb{R}^{r \times r}$; $\mathbf{U} \in \mathbb{R}^{n \times r}$, $n > r$, $\mathbf{U}\mathbf{U}^H = \mathbf{I}$; $\mathbf{V} \in \mathbb{R}^{r \times N}$, $N > \text{rand}$, $\mathbf{V}\mathbf{V}^H = \mathbf{I}$.

Consider $\mathbf{Z} = \Phi\mathbf{X} = \Phi\mathbf{U}\Sigma\mathbf{V}$

Now, $\mathbf{Z}^H \mathbf{Z} = \mathbf{V}^H \Sigma^H \mathbf{U}^H \Phi^H \Phi \mathbf{U} \Sigma \mathbf{V} = \mathbf{V}^H \mathbf{B}^H \mathbf{B} \mathbf{V}$, where $\mathbf{B} = \Phi \mathbf{U} \Sigma$.

Theorem 1. Every square root of $\mathbf{Z}^H \mathbf{Z}$ can be expressed as $\mathbf{R}\mathbf{V}$ where $\mathbf{R} \in \mathbb{R}^{r \times r}$. Further if $\text{spark}(\Phi) \geq k+1$ then \mathbf{R} is non zero. If $\text{spark}(\Phi) < k$ then \exists jointly sparse \mathbf{X} with rank r s.t $\det(\mathbf{R}) = 0$.

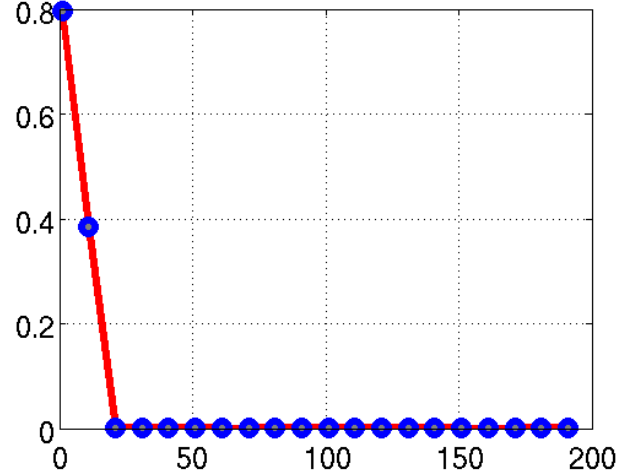


Fig. 2: Subspace estimation error vs # common measurements on Pincat data of rank 20.

If \mathbf{B} is full rank, $\mathbf{B}^H \mathbf{B}$ is positive definite and has a unique symmetric square root.

Thus for $\mathbf{Z}^H \mathbf{Z} = \mathbf{V}^H \mathbf{B}^H \mathbf{B} \mathbf{V} = \mathbf{Q}^H \mathbf{Q}$ where $\mathbf{Q} = \mathbf{R}\mathbf{V}$ and $\mathbf{B}^H \mathbf{B}$ has a unique square root \mathbf{R} then the right singular matrix of \mathbf{V} can be obtained to a factor \mathbf{R} using \mathbf{Z} iff Φ has a spark $\geq k+1$.

From our experiments we found that $\text{spark}(\Phi \geq k+1)$ is strong in practice. The projection error between original and estimated subspace is falling to 0 when the number of measurements or rows in $\Phi = r$. See Fig 2 for illustration. Noting this observation we devise the following, the proof of which comes from a determinant based argument provided in [8].

Theorem 2. The right singular matrix \mathbf{V} of any specific matrix \mathbf{X} can be uniquely recovered from the measurements $\mathbf{Z} = \Phi\mathbf{X}$ for almost all matrices $\Phi \in \mathbb{C}^{r \times n}$.

Theorem 1 is a worst case condition that guarantees the recovery for any k jointly sparse \mathbf{X} , while **Theorem 2** is sufficient for almost all matrices \mathbf{X} .

B. Subspace aware recovery of \mathbf{X}

After the basis vectors of the right subspace specified by the columns of $\mathbf{Q} \in \mathbb{C}^{r \times N}$ are estimated, we can express the matrix \mathbf{X} as $\mathbf{X} = \mathbf{P}\mathbf{Q}$, where $\mathbf{P} \in \mathbb{C}^{n \times r}$ is an appropriate coefficient matrix. From (1) we have, $\mathbf{y}_i = \mathbf{A}_i \mathbf{P} \mathbf{q}_i$ and for the 1st observation,

$$y_1 = \mathbf{A}_1 \mathbf{P} q_{11} + \mathbf{A}_1 \mathbf{P} q_{12} + \dots + \mathbf{A}_1 \mathbf{P} q_{1r}$$

$$y_1 = [q_{11} \mathbf{A}_1 \quad q_{12} \mathbf{A}_1 \quad \dots \quad q_{1r} \mathbf{A}_1] \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_r \end{bmatrix}$$

Vectorizing all the y_i s we obtain,

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix}}_{\text{vec}(\mathbf{Y})} = \underbrace{\begin{bmatrix} q_{11} \mathbf{A}_1 & \cdots & q_{r1} \mathbf{A}_1 \\ \vdots & & \vdots \\ q_{1N} \mathbf{A}_N & \cdots & q_{rN} \mathbf{A}_N \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_r \end{bmatrix}}_{\text{vec}(\mathbf{P})} \quad (2)$$

Note that the sparsity of $\text{vec}(\mathbf{P})$ is kr . Thus, $\text{vec}(\mathbf{P})$ can be uniquely identified if $\text{spark}(\mathbf{B}) = 2kr$. We conjecture that this is in fact pessimistic and unique recovery can be achieved if $\text{spark}(\mathbf{A}_i) \geq 2kr/N$.

Theorem 3. Assume that $N = (p + 1)r$, where p is an arbitrary integer. Let

$$\begin{aligned} \mathbf{C}_1 &= \mathbf{A}_1 = \mathbf{A}_{2..} = \mathbf{A}_r \\ \mathbf{C}_2 &= \mathbf{A}_{r+1} = \mathbf{A}_{r+2..} = \mathbf{A}_{2r} \\ &\vdots \\ \mathbf{C}_p &= \mathbf{A}_{pr+1} = \mathbf{A}_{pr+2..} = \mathbf{A}_N \end{aligned} \quad (3)$$

Here, $\mathbf{C}_i \in \mathbb{R}^{m_i \times n}$; $i = 1, \dots, p$. \mathbf{P} can be uniquely determined from (2) if $\text{spark}(\mathbf{X}) = r + 1$ and

$$\text{spark} \left(\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_p \end{bmatrix} \right) \geq 2k. \quad (4)$$

We present here a proof for the rank one case. The proof for a general low rank r case is omitted here for brevity.

Proof. Consider $r = 1$. $y_i = \mathbf{A}_i x_i = \mathbf{A}_i u_1 v_i$, where u has only one column as the rank is 1 and v_i s are scalars. We have, $y_i/v_i = \mathbf{A}_i u_1$. From the compressed sensing results on unique recovery of a single unknown sparse vector we need,

$$\text{spark} \left(\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_N \end{bmatrix} \right) \geq 2k. \quad (5)$$

Note that \mathbf{X} and \mathbf{Q} share the same null space. So from the spark condition $\text{spark}(\mathbf{X}) = r + 1$ in Theorem 3, we get that any r columns of \mathbf{Q} forms an non singular matrix. If we consider the clustering in (3), we only need the \mathbf{Q} submatrix formed by picking the rows corresponding to the indices in a group in (3) to be non singular. Simply, we want consecutive r rows of \mathbf{Q}^H to be full rank. This requirement comes from the proof of the general case.

This is only required for the sufficient condition. In practice as shown by our simulation results we do not need that spark condition on \mathbf{X} . It is just to attain the theoretical claim.

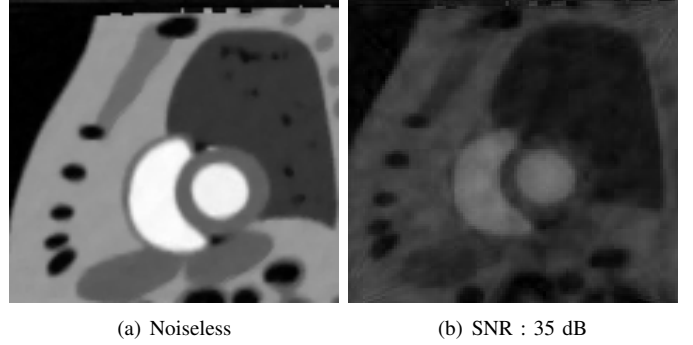


Fig. 3: Reconstructing Pincat data using our two step strategy. Left: in a noiseless setting. Right: with a SNR of 35 dB

C. Effect of sampling pattern

The clustering specified in (3) assumes that every consecutive r frames are observed by same set of variable measurements. We have devised this clustering of sensing matrices to arrive at the sufficient condition in (4). As specified above, our goal is to ensure that the frames which are operated by similar sensing matrices or in other words the frames that are operated by matrices in the same group should form an invertible matrix. The \mathbf{Q} submatrix such formed should be non singular. The reconstruction of simulated data satisfying (3) is shown in Fig 3. We see perfect recovery in the noiseless case on the left.

This clustering can also be done differently. Possibilities include repeating the \mathbf{A}_i s after every r frame hence generating a periodicity in the observation matrix. Another possibility is taking the given cluster and permuting it across the N frames i.e., randomly shuffling the set of sensing matrices such that a certain r of them are same. We want to look at how the reconstructions look like if we violate from the pattern we have in (3), i.e., for a given number of measurements how the three patterns perform. This depends on the knowledge of the structure of the data we are trying to reconstruct.

We will highlight this effect in Fig 4. The chosen toy Pincat dataset is periodic in nature with a rank 20. So every 20^{th} frame belongs to the same cardiac phase. So the matrix formed by stacking every 20^{th} frame is rank 1. Clearly, the clustering in (3) is favorable here. From the possibilities mentioned above, the periodic clustering will fail as the \mathbf{Q} submatrix formed using every $r = 20^{th}$ frame will be rank deficient or in this case rank 1. Random permutation works a bit better than the periodic clustering as it brings some complimentary information. The rank of the \mathbf{Q} submatrix formed by the random clustering is still higher than 1 but not full rank, hence the recovery isn't perfect.

So the underlying fact remains that we can choose any clustering we want provided the \mathbf{C}_i s formed satisfy (4) and we have the idea of roughly how the time frames behave (periodicity) for a particular dataset. And lastly, this is only to comply with the theory. In practice as we will see in Fig 5, even without satisfying this spark condition $\text{spark}(\mathbf{X}) = r + 1$ we can get a good recovery.

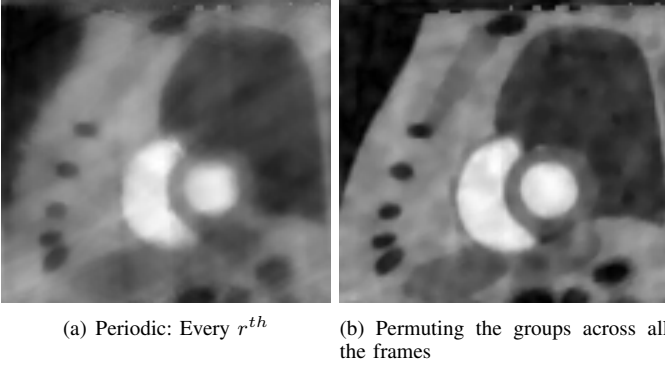


Fig. 4: Effect of clustering the similar sensing matrices on the reconstruction: placing them periodically (left) and permuting them randomly (right)

D. Measurements required

For a matrix of dimension $n \times N$ and rank r , classical MMV scheme requires $s \geq (2k - r + 1)N$ measurements for its unique recovery. Dense measurement scheme requires s of the order of the degrees of freedom in a matrix. Variable and common measurements in the proposed scheme requires $s \geq (2k - r + N)r$ for unique recovery. Combining the above, \mathbf{X} can be uniquely identified if the average number of measurements per snapshot is of the order of $s = r + 2kr/N$. Note that $s \rightarrow r$ as $N \rightarrow \infty$. This is a quite considerable gain over the number required for MMV identifiability, specified by $2k - r + 1$, especially when $r \ll k$.

III. ALGORITHM

We use Alternating direction method of multipliers (ADMM) [9], [10] which combines the augmented Lagrangian method and variable splitting techniques to solve the constrained optimization problem. The problem is posed as:

$$\begin{aligned} \underset{\mathbf{P}}{\operatorname{argmin}} \quad & \|\mathbf{B} \operatorname{vec}(\mathbf{P}) - \operatorname{vec}(\mathbf{Y})\|_2^2 + \\ & \lambda \sum_{xy} \sqrt{\sum_t \mathbf{P}_x(t)^2 + \mathbf{P}_y(t)^2} \end{aligned} \quad (6)$$

With the \mathbf{V} 's supplied by the 1st step we solve for \mathbf{P} 's by creating an auxiliary variable \mathbf{G} . After the variable splitting we have the Lagrangian,

$$\begin{aligned} \mathcal{L} = & \|\mathbf{B} \operatorname{vec}(\mathbf{P}) - \operatorname{vec}(\mathbf{Y})\|_2^2 + \\ & \lambda \sum_{xy} \sqrt{\sum_t \mathbf{G}_x(t)^2 + \mathbf{G}_y(t)^2} \\ & + \lambda \frac{\beta}{2} \|\nabla_x \mathbf{P} - \mathbf{G}_x\|_2^2 + \lambda \frac{\beta}{2} \|\nabla_y \mathbf{P} - \mathbf{G}_y\|_2^2 \\ & + \langle \Lambda_x, \mathbf{G}_x - \nabla_x \mathbf{P} \rangle + \langle \Lambda_y, \mathbf{G}_y - \nabla_y \mathbf{P} \rangle \end{aligned} \quad (7)$$

where β is the penalty parameter which drives the difference between the auxiliary and original variable to zero, λ is the regularization parameter, Λ_x, Λ_y are the Lagrange multipliers. The \mathbf{P} sub-problem is solved using Preconditioned conjugate gradient and \mathbf{G} sub-problem is solved by l_1 shrinkage. The data

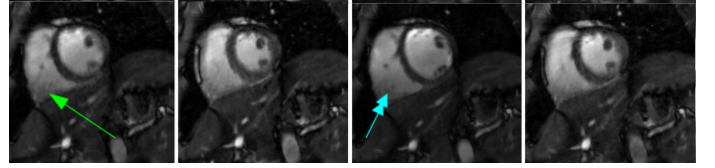


Fig. 5: Reconstruction of CINE MRI data using our algorithm. Few cardiac phases are shown. The dots pointed out by the arrow shows the same phases reappearing.

is transformed to the gradient domain to make it sparse so the problem is solved with a Total variation (TV) regularization. l_1 norm encourages sparsity and l_2 norm encourages diversity [11]. We enforce joint sparsity by penalizing the $l_1 - l_2$ mixed norm of the gradient data.

IV. NUMERICAL RESULTS

We validate our Theorems on toy Pincat [12] phantom data of dimension $128 \times 128 \times 200$ and rank 20. In Fig. 2 we consider a noiseless setting. We plot the projection error between the estimated and actual right subspace vs the number of random samples of the image data in the Fourier domain. Projection error between two subspaces V_1 and V_2 is defined as:

$$\frac{\|(\mathbf{I} - \mathbf{V}_1 \mathbf{V}_1^H) \mathbf{V}_2\|_2^2 + \|(\mathbf{I} - \mathbf{V}_2 \mathbf{V}_2^H) \mathbf{V}_1\|_2^2}{\|\mathbf{V}_1\|_2^2 + \|\mathbf{V}_2\|_2^2}. \quad (8)$$

We see a drop in the normalized error after the the number of samples equals the rank 20.

In Fig. 3 we look at the subspace aware recovery of the Pincat data using the clustering in (3) which is consistent with the formulation of Theorem 4. This was undersampled with 4 common and 5 variable lines for a noiseless (left) and a noisy setting (SNR of 40 db on the right). We get perfect recovery in the noiseless case. All the Pincat performance comparison can be done with the perfect reconstruction Fig. 3a as reference. We see blurring artifacts in the noisy case Fig. 3b.

Fig. 4 shows the effect of clustering the sampling pattern on the reconstruction. On the left figure the similar sensing matrices are distributed periodically to every r^{th} frame and on the right the clusters are shuffled randomly across all the N frames. We see that for this dataset, following the setting of Theorem 4 better reconstructions are obtained for the same number of measurements as against when those conditions were violated. This indicates that the clustering pattern can be modified based on knowledge of the structure in the data. But after that if (4) is satisfied good recovery is ensured. In all the Pincat results, the first frame is shown from the reconstructed time series.

We apply our algorithm on free breathing CINE data in Fig. 5. A cropped portion of the reconstruction is shown highlighting the heart. The dots indicated by the arrows show the similar cardiac phases. The data was acquired using an SSFP sequence with an 18 channel coil array, with TR/TE of 4.2/2.1 ms, matrix size of 512×512 , FOV of $300\text{mm} \times 300\text{mm}$ and slice thickness of 5mm on 3T Siemens Trio scanner. We considered 12 radial lines of Fourier space to reconstruct each image frame, 4 of which were common measurement

lines. This corresponds to a temporal resolution of 50 ms. The acquisition time was 25 s translates to 500 image frames. Each radial Fourier line is a 512 pixels long line on the 2D Fourier space data passing through the center of the Fourier space as shown in Fig 1. The blurring is optimal and the cardiac phases are clearly visible. Despite the undersampling our algorithm is providing good reconstruction using the two step strategy. This is significant because MR acquisition is slow and we can't fully sample MR data to get the perfect reconstruction. Thus the strategic undersampling to get good recovery is essential when the data has a low rank and joint sparse structure.

V. CONCLUSION

We propose to reconstruct a low rank and sparse matrix from its under sampled measurements. We introduced a two step subspace strategy that recovers the right and left subspaces one after another. Our sufficient conditions show that by using our sampling scheme we are gaining in terms of the number of measurements to uniquely recover the data as compared to the MMV scheme. Our experiments suggest that considerable undersampling when strategically done, gives good reconstruction in the noiseless case for the Pincat data. In the CINE MRI data where the ground truth is unknown considerable undersampling along with our reconstruction algorithm shows good recovery. The reduced number of required measurements due to undersampling will translate to high accelerations in dynamic MRI.

REFERENCES

- [1] H. Pedersen, S. Kozerke, S. Ringgaard, K. Nehrke, and W. Y. Kim, "k-t PCA: temporally constrained k-t BLAST reconstruction using principal component analysis," *Magn Reson Med*, vol. 62, no. 3, pp. 706–716, Sep 2009.
- [2] S. G. Lingala, E. DiBella, M. Jacob *et al.*, "Accelerated imaging of rest and stress myocardial perfusion mri using multi-coil kt slr: a feasibility study," *Journal of Cardiovascular Magnetic Resonance*, vol. 14, no. Suppl 1, p. P239, 2012.
- [3] Z. Liang, "Spatiotemporal imaging with partially separable functions," in *ISBI, 2007*, pp. 181–182.
- [4] B. Zhao, J. Haldar, A. Christodoulou, and Z. Liang, "Further development of image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1593–1596.
- [5] M. Golbabaee and P. Vanderghenst, "Compressed sensing of simultaneous low-rank and joint-sparse matrices," *arXiv preprint arXiv:1211.5058*, 2012.
- [6] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.
- [7] A. C. Gilbert, J. Y. Park, and M. B. Wakin, "Sketched SVD: recovering spectral features from compressive measurements," *Arxiv preprint arXiv:1211.0361*, 2012.
- [8] L. Losaz, "Singular spaces of matrices and their application in combinatorics," *Bulletin of the Brazilian Mathematical Society*, vol. 20, no. 1, pp. 87–89, 1989.
- [9] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *Image Processing, IEEE Transactions on*, vol. 20, no. 3, pp. 681–695, 2011.
- [10] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, no. 5, pp. 303–320, 1969.
- [11] J. Le Roux, P. T. Boufounos, K. Kang, and J. R. Hershey, "Source localization in reverberant environments using sparse optimization," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4310 – 4314.
- [12] B. Sharif and Y. Bresler, "Physiologically improved neat phantom (pincat) enables in-silico study of the effects of beat-to-beat variability on cardiac mr," in *Proceedings of the Annual Meeting of ISMRM, Berlin, 2007*, p. 3418.