

SEPARATING SPARSE AND LOW-DIMENSIONAL SIGNAL SEQUENCES FROM TIME-VARYING UNDERSAMPLED PROJECTIONS OF THEIR SUMS

Jinchun Zhan, Namrata Vaswani*

Ian Atkinson

Dept. of Electrical and Computer Engineering
Iowa State University, Ames, IA, USA
Email: {jzhan,namrata}@iastate.edu

Center for Magnetic Resonance Research
University of Illinois at Chicago, IL, USA
Email:ian@uic.edu

ABSTRACT

The goal of this work is to recover a sequence of sparse vectors, s_t ; and a sequence of dense vectors, ℓ_t , that lie in a “slowly changing” low dimensional subspace, from time-varying undersampled linear projections of their sum. This type of problem typically occurs when the quantity being imaged can be split into a sum of two layers, one of which is sparse and the other is low-dimensional. A key application where this problem occurs is in undersampled functional magnetic resonance imaging (fMRI) to detect brain activation patterns in response to a stimulus. The brain image at time t can be modeled as being a sum of the active region image, s_t , (equal to the activation in the active region and zero everywhere else) and the background brain image, ℓ_t , which can be accurately modeled as lying in a slowly changing low dimensional subspace.

We introduce a novel solution approach called matrix completion projected compressive sensing or MatComProCS. Significantly improved performance of MatComProCS over existing work is shown for the undersampled fMRI based brain active region detection problem.

Index Terms— matrix completion, compressive sensing, fMRI

1. INTRODUCTION

The goal of this work is to recover a sequence of sparse vectors, s_t ; and a sequence of dense vectors, ℓ_t , that lie in low dimensional subspace which is either fixed or can “slowly change” over time, from time-varying undersampled linear projections of their sum, i.e. from measurements of the form

$$y_t := A_t(s_t + \ell_t) \quad (1)$$

where $A_t := H_t \Psi$, Ψ is an orthonormal matrix ($m \times m$) representing the measurement basis; $(H_t)_{n \times m} := I_{\omega_t}'$ is the undersampling matrix at time t and $\omega_t \subseteq \{1, 2, \dots, m\}$ is the set of indices of the rows selected at time t . Here I is the identity matrix; and the notation $(B)_T$ refers to the sub-matrix containing columns of the matrix B with indices in the set T .

One way to quantify the “slowly changing low dimensional subspace” assumption as follows. We assume that $\ell_t = P_{(t)} a_t$ with $P_{(t)} = P_j$ for all $t_j \leq t < t_{j+1}$. Here P_j is an *unknown* $m \times r$ matrix with $r \ll m$ orthonormal columns (i.e. $P_j' P_j = I$), and a_t is an *unknown* zero mean $r \times 1$ vector. The change times t_j are also *unknown*. At the changes times, $t = t_j$, some new directions get added to and/or removed from P_j . Let $P_{j,\text{new}}$ denote the newly

added directions. By “slowly changing”, we mean that the projection of ℓ_t along these new directions, $a_{t,\text{new}} := P_{j,\text{new}}' \ell_t$, is small for some time after $t = t_j$, i.e. $\|a_{t,\text{new}}\|_2 = \|(I - P_{j-1} P_{j-1}') \ell_t\|_2$ is small compared to $\|\ell_t\|_2$ for some time after $t = t_j$. Of course, over time, this projection may increase gradually to a large enough value. Here $'$ denotes transpose.

In addition, we also assume that an initial short sequence of measurements of only the low dimensional sequence are available, i.e. we have $y_t := A_t \ell_t$, for all $t = 1, 2, \dots, t_0$.

This type of problem typically occurs when the quantity being imaged can be split into a sum of two layers, one of which is sparse and the other is low-dimensional. For example, in functional magnetic resonance imaging (fMRI), $\Psi = F$ is the discrete Fourier transform (DFT) matrix, $z_t := s_t + \ell_t$ is the brain image at time t which consists of the active region image, s_t , (equal to the activation in the active region and zero everywhere else) and the background brain image, ℓ_t , which can be modeled as lying in a slowly changing low dimensional subspace. In single-pixel video imaging, Ψ can be a block Hadamard matrix [1] or any other structurally random matrix, $z_t = s_t + \ell_t$ is the video image at time t , ℓ_t is the background layer while s_t is a sparse image that is zero everywhere except where the foreground, e.g. moving object(s), is present. As shown in Fig 1, the “slowly changing” low dimensional subspace assumption indeed holds for the background sequence in fMRI. It has also been verified for the background sequences of video data in earlier work [2]. Moreover, the assumption that a short sequence of measurements of only the low dimensional sequence (background sequence) are available is also easy to satisfy in either application.

Related Work. While there has been a very large amount of recent work on both sparse recovery (more popularly known as compressive sensing (CS)), e.g. [3, 4, 5] and all later work, and on low dimensional signal sequence recovery (low rank matrix completion), e.g. [6, 7, 8], the problem of recovering structured signals from their sums has not received as much attention. The problem of separating linear projections of a sparse signal from outliers which are modeled as large but sparse vectors has been studied in [9] as well as in other later works such as [10, 11, 12, 13].

Recent works such as the robust PCA work of Candes et al, and Chandrasekharan et al [14, 15] focus on separating a sparse and a low-dimensional signal sequence from their sums. In very recent works [16, 17, 18] the compressive robust PCA problem has also been looked at. All these works assume that the low dimensional signal sequence is dense (resulting in a low-rank but dense matrix) and the sparse signal sequence has independent and identically distributed (i.i.d.) support sets (resulting in a sparse but full rank matrix with high probability). However, the assumption of i.i.d. support sets is often not valid in many practical applications where the sparse sig-

*This work was partly supported by NSF grant CCF-1117125.

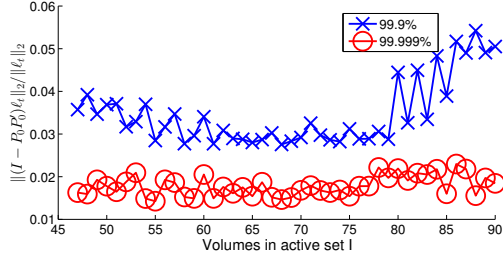


Fig. 1. Verifying slow subspace change in an fMRI background brain image sequence. We used a 90 frame rest brain fMRI sequence (brain fMRI when no stimulus is provided to the subject) where each image was of size 64×64 , i.e. $m = 4096$. We used the first $t_0 = 46$ frames to compute P_0 . We computed P_0 by computing the SVD of $[\ell_1, \ell_2, \dots, \ell_{46}]$ and retaining the eigenvectors containing 99.9% (and 99.999%) of the total energy. We observed that 99.9% of the energy was contained in only $r = 1$ direction while 99.999% of the energy was contained in only 42 directions. For $t > t_0$ we plot $\|(I - P_0 P_0')(\ell_t)\|_2 / \|\ell_t\|_2$. Notice that, for the next 44 frames, this quantity is below 6% in both cases, i.e. slow subspace change does indeed hold.

nal is usually both spatially and temporally correlated. For example, in the brain fMRI problem, the active region consists of one, or a few, contiguous regions of the brain and it changes only a little from one time instant to the next. The same is true for video where the foreground typically consists of one, or a few, moving objects.

In recent work [19, 20], we studied the full sampled version of the above problem and proposed a solution called recursive projected CS that removes the above limitation. In [21], we studied the undersampled version but with a fixed measurement matrix, i.e. the current problem but with $A_t = A$ for all t . Using a time-varying A_t , as done in this paper, has the following advantages. First, with $A_t = A$ one can only recover s_t and $\tilde{\ell}_t := A\ell_t$ (but not ℓ_t) whereas a time-varying A_t makes it possible to recover both s_t and ℓ_t . Second, a time-varying A_t makes sparse recovery much more likely to succeed; the probability that most random matrices A_t are bad (do not satisfy the restricted isometry property (RIP)) is much lower than the probability that a given fixed A is bad. Finally, a time-varying A_t can allow one to use batch sparse recovery algorithms, such as k-t-focuss [22] or [23], which exploit Fourier sparsity along the time axis or blind CS techniques [24, 25], which also exploit sparsity along the time axis, in the projected sparse recovery step (step 2 of Algorithm 1), although we currently do not use these.

We should point out that the solution approach needed for the case of time-varying A_t is quite different from that used in [21] when A_t is fixed. The reason is that when $A_t = A$, $\tilde{\ell}_t := A\ell_t$ also lies in a slowly changing low dimensional subspace and this fact can be exploited in trying to recover s_t and $\tilde{\ell}_t$. However this is no longer true for the sequence $A_t \ell_t$.

Contributions. (1) In this work we propose a novel solution approach called Matrix Completion Projected CS (MatComProCS) for separating sparse and low dimensional signal sequences from time-varying undersampled projections of their sums. (2) We show how it can be applied to successfully detect the brain activation pattern from undersampled functional MRI sequences. This problem becomes particularly difficult in the low contrast to noise ratio (CNR) regime, i.e. when the activation is very weak compared with the background. This is exactly the situation where the standard approach of first recovering the entire image sequence from undersampled measurements and then solving the active region detection

problem as in [26] fails. Extensive comparisons with other existing works, e.g. [26] and references therein, for solving this problem show the superiority of MatComProCS.

Notation. For an index set $T \subset \{1, 2, \dots, m\}$, we use T^c to denote its complement w.r.t. $\{1, 2, \dots, m\}$, i.e., $T^c := \{i \in \{1, 2, \dots, m\} : i \notin T\}$. For a matrix B , B' denotes its transpose, and B^\dagger its Moore-Penrose pseudo-inverse. $\|B\|_*$ is the nuclear norm, i.e., sum of singular values. $\|B\|_F$ is the Frobenius norm. B_T denotes the sub-matrix consisting of the columns of B indexed by T . Also \emptyset denotes the empty set. For a vector b , b_T denotes a vector consisting of the entries of b indexed by T .

Define the set of index sets, $\Omega_{t_1, t_2} := \{\omega_{t_1}, \omega_{t_1+1}, \dots, \omega_{t_2}\}$ where $\omega_\tau \subseteq \{1, 2, \dots, m\}$ is a set of cardinality $n < m$. In this work, ω_τ is the set of indices of the rows selected at time τ . For an $m \times t$ matrix Z , define the projector $\mathcal{P}(\Omega_{t_1, t_2}; Z) := [(Z_1)_{\omega_{t_1}}, (Z_2)_{\omega_{t_1+1}}, \dots, (Z_t)_{\omega_{t_2}}]$ where Z_i denotes the i^{th} column of the matrix Z .

2. MATRIX COMPLETION PROJECTED CS (MATCOMPROCS)

The stepwise Matrix Completion Projected CS (MatComProCS) is given in Algorithm 1. The key ideas of our solution approach are as follows. Given an initial short sequence of measurements of the form $y_t = H_t \Psi \ell_t$, for $t = 1, 2, \dots, t_0$, where the ℓ_t 's satisfy $\ell_t = P_0 a_t$ for an $m \times r$ matrix P_0 with $r \ll m$, we estimate P_0 as follows. Let $L := [\ell_1, \ell_2, \dots, \ell_{t_0}]$. The problem of recovering ΨL (and hence $L = \Psi' \Psi L$) from $[y_1, y_2, \dots, y_{t_0}]$ becomes a low-rank matrix completion problem and one can use nuclear norm minimization [6, 7] or one of the greedy approaches in literature, e.g. [8], to solve it. In this work we use nuclear norm minimization, i.e. we solve

$$\min_L \|\Psi L\|_* \text{ s.t. } \mathcal{P}(\Omega_{1, t_0}; \Psi L) = [y_1, y_2, \dots, y_{t_0}].$$

Let \hat{L} denote its solution. As shown in [6], this will result in exact recovery if ΨL satisfies certain incoherence assumptions, the sets ω_t are sampled uniformly at random and independently of all previous sets and $nt_0 \geq Cm^{5/4}r \log m$. We estimate P_0 by computing the left singular vectors of \hat{L} with nonzero (in practice, non-negligible) singular values, i.e. by principal components analysis (PCA).

For $t > t_0$, we have measurements of the form $y_t := A_t s_t + A_t \ell_t$ with $A_t := H_t \Psi$. Suppose that $\hat{P}_{(t-1)}$ is available and is such that $\text{span}(\hat{P}_{(t-1)}) \approx \text{span}(P_{(t)})$. By projecting perpendicular to the span of $A_t \hat{P}_{(t-1)}$, we can approximately nullify $A_t \ell_t$ and get projected measurements of $A_t s_t$. To be precise, we compute $\tilde{y}_t := \Phi_t y_t$ where $\Phi_t := (I - A_t \hat{P}_{(t-1)} (A_t \hat{P}_{(t-1)})^\dagger)$. Clearly, \tilde{y}_t can be rewritten as $\tilde{y}_t = (\Phi_t A_t) s_t + \beta_t$, where $\beta_t := \Phi_t A_t \ell_t$ can be interpreted as small noise resulting from incomplete nullification¹ of $A_t \ell_t$. The problem of recovering s_t from \tilde{y}_t now becomes a standard sparse recovery problem in small noise. One can use any of the sparse recovery / compressive sensing (CS) approaches to recover s_t . In this work we use noisy ℓ_1 minimization at the first time instant and we use modified-CS [27] at all other time instants, i.e. for $t > t_0$, we compute $\hat{s}_{t, \text{modcs}}$ as the solution of

$$\min_x \|x_{T_{t-1}^c}\|_1 \text{ s.t. } \|\tilde{y}_t - (\Phi_t A_t)x\|_2 \leq \xi_t$$

with $\hat{T}_{t_0} = \emptyset$. We can pick ξ_t using the following idea which was suggested in [2] for another similar problem. Notice that ξ_t is the

¹incomplete because $\text{span}(\hat{P}_{(t-1)}) \neq \text{span}(P_{(t)})$

Algorithm 1 MatComProCS

Input: y_t 's, Output: \hat{P}_t 's, \hat{s}_t 's, $\hat{\ell}_t$'s, Parameters: $\xi_t, r, C, \alpha, \epsilon$
Initialization:

- Compute \hat{L} as the solution of

$$\min_L \|L\|_* \text{ s.t. } \mathcal{P}(\Omega_{1,t_0}; \Psi L) = [y_1, y_2 \dots y_{t_0}]$$

- Compute the SVD $\hat{L} \stackrel{SVD}{=} U\Lambda V'$ and compute $\hat{P}_0 = (U)_{[1,2,\dots,\hat{r}]}$, where \hat{r} is the number of singular values larger than the 99.9% energy threshold of Λ .
- Let $\hat{P}_{(t_0)} \leftarrow \hat{P}_0$
- Let $\hat{T}_{t_0} \leftarrow \emptyset$.

For $t > t_0$, do the following:

1. Nullify ℓ_t : compute $\Phi_t \leftarrow (I - A_t \hat{P}_{(t-1)} (A_t \hat{P}_{(t-1)})^\dagger)$ and $\tilde{y}_t = \Phi_t y_t$.
2. Sparse Recovery by Modified-CS: compute $\hat{s}_{t,\text{modes}}$ as the solution of

$$\min_x \|x\|_{\hat{T}_{t-1}^c} \text{ s.t. } \|\tilde{y}_t - (\Phi_t A_t)x\|_2 \leq \xi_t$$

with $\xi_t = 1.01 \|\Phi_{t-1} \hat{\ell}_{t-1}\|_2$

3. Support Estimate: $\hat{T}_t = \{i : |(\hat{s}_{t,\text{modes}})_i| > C\}$ where C is the larger one of the 100th largest absolute value of $\hat{s}_{t,\text{modes}}$ and the 99% energy threshold of $\hat{s}_{t,\text{modes}}$ [27].
4. LS estimate: $(\hat{s}_t)_{\hat{T}_t} = ((\Phi_t A_t)_{\hat{T}_t})^\dagger y_t, (\hat{s}_t)_{\hat{T}_t^c} = 0$
5. Estimate $\tilde{\ell}_t$: compute $\hat{\ell}_t = y_t - A_t \hat{s}_t$
6. Recover ℓ_t 's and estimate $\hat{P}_{(t)}$ every α frames
If $\text{mod}(t - t_0, \alpha) \neq 0, \hat{P}_{(t+1)} \leftarrow \hat{P}_{(t)}$.
If $\text{mod}(t - t_0, \alpha) = 0,$

- (a) Compute \hat{L} as the solution of

$$\min_L \|L\|_* \text{ s.t. } \|\hat{\ell}_{t-2\alpha+1}, \hat{\ell}_{t-2\alpha+2} \dots \hat{\ell}_t - \mathcal{P}(\Omega_{t-2\alpha+1,t}; \Psi L)\|_F \leq \epsilon$$

- (b) Compute the SVD $\hat{L} \stackrel{SVD}{=} U\Lambda V'$ and compute $\hat{P}_{(t)} = (U)_{[1,2,\dots,\hat{r}]}$, where \hat{r} is the number of singular values larger than the 99.9% energy threshold of Λ .

7. Update $t \leftarrow t + 1$ and go to step 1.
-

upper bound on $\|\beta_t\|_2$. We do not know β_t . All we can do is estimate it from $t - 1$ as $\hat{\beta}_{t-1} = \Phi_t \hat{\ell}_{t-1}$. We use a value a little larger than $\|\hat{\beta}_{t-1}\|_2$ for ξ_t , i.e. we let $\xi_t = 1.01 \|\hat{\beta}_{t-1}\|_2 = 1.01 \|\Phi_t \hat{\ell}_{t-1}\|_2$.

By thresholding on $\hat{s}_{t,\text{modes}}$, one gets an estimate of its support. As suggested in earlier work [27, 21], the support can be estimated as the 99% energy set of $\hat{s}_{t,\text{modes}}$ or the 100 largest absolute value of $\hat{s}_{t,\text{modes}}$. By computing a least squares (LS) estimate of s_t on the estimated support and setting it to zero everywhere else, we can get a more accurate final estimate, \hat{s}_t , as first suggested in [28].

This \hat{s}_t is used to estimate $\tilde{\ell}_t$ as $\hat{\ell}_t := y_t - A_t \hat{s}_t$. Every α frames, these can then be used in a noisy matrix completion algorithm [7] to

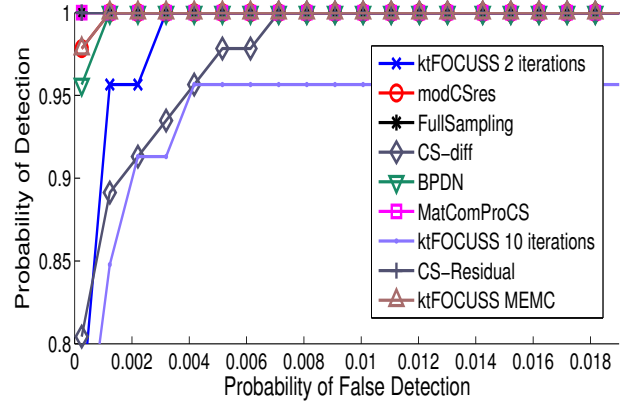


Fig. 2. ROC curves when CNR = 4.

recover the ℓ_t 's. The noisy version is needed because $\hat{\ell}_t = A_t \ell_t + A_t e_t$, where $e_t := s_t - \hat{s}_t$ is the error in estimating s_t . PCA on the last 2α estimates of $\tilde{\ell}_t$ or projection PCA [2] can then be used to compute \hat{P}_t . We summarize the complete algorithm in Algorithm 1.

3. SIMULATION RESULTS

In this section, we show experiments on real fMRI data with simulated activation. We evaluate the detection performance of MatComProCS and other algorithms using Receiver Operating Characteristic (ROC) curves. To quantify detection performance using ROC curves, we need to know the ground truth for active regions. We captured a rest brain sequence (brain fMRI when no stimulus was provided to the subject) using a real MR scanner, but adding the activation later in software. Rest fMRI (TR/TE=2500/24.3 ms, 90 degrees flip angle, 3 mm slice thickness, 22 cm FOV, 64×64 matrix, 90 volumes) was performed using a 3T whole-body MR scanner and a gradient-echo echo-planar imaging (EPI) acquisition sequence. We added synthetic BOLD contrast with different CNR to pixels corresponding to motor activation on one slice. The 64×64 slice image has 23 active pixels. The BOLD signal was created by convolving a bi-Gamma HDR model (6-s onset delay, 4-s FWHM) with binary-valued function representing a block stimulus (30s active, 30s rest, obtaining 1 volume at every 3 seconds, start/end in rest condition). Stimulus has been added for 4 times at every other 60 seconds starting from 30th second, thus, the active set S_{act} is roughly $\{13, 14, \dots, 23, 33, 34, \dots, 43, 53, 54, \dots, 63, 73, 74, \dots, 83\}$ (this is computed by assuming a 2 frame delay for the activation to become significant), and the other volumes form the rest set S_{rest} . 5 separate sequences were generated by resampling the original rest fMRI data with the wavetrapping technique and adding activation to the motor pixels, as explained in [29]. fMRI measurements were generated by uniformly undersampling the DFT of each image to retain $n = 0.5m$ measurements.

Before we proceed we explain exactly how MatComProCS was implemented for the dataset. In fMRI, one typically does not care about recursive recovery. Since we had a sequence with only 12 initial rest (background-only) frames and these are not sufficient for accurately solving the matrix completion problem, we rearranged the sequence so that all rest frames, i.e. the frames in the set S_{rest} appeared first. There were a total of 46 rest frames, thus we imple-

mented Algorithm 1 with $t_0 = 46$. Also, we picked $\alpha = 46$. Since there were only 44 other frames, this meant that we never actually updated $\hat{P}_{(t)}$. In future work, we will experiment with significantly longer sequences where the update will be essential. Finally, our model assumes that ℓ_t is zero mean. However as with all imaging problems, the background image sequence is not zero mean. Hence, after recovering ℓ_t we computed its mean μ and subtracted it out before the PCA step. The same mean was also subtracted from y_t before step 1 of the algorithm.

We used MatComProCS (Algorithm 1) on these measurements to recover s_t and ℓ_t for each of the 10 sequences. The active region was detected by thresholding on the average of the estimated active images \hat{s}_t , i.e. on $(1/44) \sum_{t=47}^{90} \hat{s}_t$. We varied the threshold in a range and for each threshold computed the detection probability (number of detected pixels divided by total number of active pixels) and the false alarm probability (number of wrongly detected pixels divided by the total number of inactive pixels) and plotted these to get the ROC curve.

We show comparisons with the approach used in [26] for the same problem. In this work modified-CS-residual (which is an algorithm for recursively recovering a time sequence of sparse signals with slowly changing support sets and signal values) was used to reconstruct the entire image sequence. It assumed that each image of the sequence was wavelet sparse. The reconstructed image sequence and the knowledge of the active and rest frames was fed into a t-test based active region detection technique described in [29] to obtain the ROC curves. We also compare with all the works (kt-FOCUSS, basis pursuit denoising (BPDN), CS-residual, CS-diff) compared in [26]. All of these algorithms are also sparse recovery techniques (either recursive or batch) that assume different types of sparsity for the fMRI sequence. CS-residual, CS-diff and BPDN assumed that each image is wavelet sparse. k-t-focuss assumes wavelet sparsity along the x-y axis and Fourier sparsity along the time axis. All of the above algorithms used $n = m$ (100% measurements for the first frame and $n = 0.5m$ for all other frames and thus, in fact used more total number of measurements than MatComProCS. Also the best undersampling scheme was used for each case as explained in [26].

Finally, as baseline, we used a full-sampled fMRI sequence (image sequence recovered by inverse DFT), i.e. $n = m$ for all t , followed by the t-test.

We show the ROC curve comparisons in Fig. 2. As can be seen, MatComProCS significantly outperforms almost all the other algorithms and its ROC curve comes quite close to that of the baseline (full-sampled case) which has a perfect ROC. There are two main reasons for this. The first is that while the wavelet basis may be a valid sparsity basis for the sequence, it is not the best one (we used the same wavelet basis that was used in [26] and was selected after experimenting with many choices). It is also not clear which basis or dictionary is a better one in terms of improving the sparsity. On the other hand, MatComProCS does not assume a sparsity basis for the background sequence. Instead, it only assumes that the sequence is low dimensional and estimates the principal components from the data itself. Moreover, the active region is recovered by first approximately nullifying the background sequence (which is the “noise” for this problem) followed by sparse recovery and this also makes the recovery much better.

We also compare against an adaptation of Dense Error Correction (aDEC) which assumes $\ell_t = Wx_t$ where x_t is a sparse vector, i.e. the background is wavelet sparse and s_t is just sparse. It thus solves $\min_{x,s} \|x\|_1 + \|s\|_1$ s.t. $y_t = A_t Wx + A_t s$. But it turned out that it achieved only 0.01% detection for 5% false alarm with our data, much worse than all the others, thus we didn’t plot it on

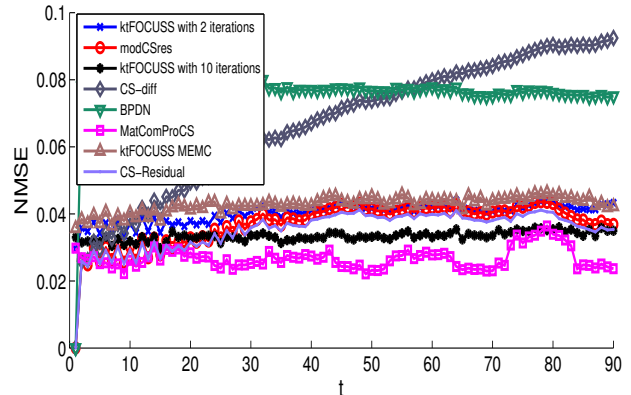


Fig. 3. Background recovery error when $\text{CNR} = 3$. $\text{NMSE}(= \frac{\|(\hat{\ell}_t - \ell_t)_{T^c}\|_2}{\|(\ell_t)_{T^c}\|_2})$, where T is the active area, $|T| = 23$.)

our figures. The reason is that (a) the wavelet basis is not a very good sparsity basis for the background sequence and also that (b) the active region image s_t is itself compressible in the wavelet basis.

We also compared background image recovery performance of MatComProCS and of the other algorithms in Fig 3. As can be seen MatComProCS again significantly outperforms the other algorithms.

4. CONCLUSIONS AND FUTURE WORK

In this work we studied the problem of trying to separate a sequence of sparse vectors, s_t ; and a sequence of dense vectors, ℓ_t , that lie in a “slowly changing” low dimensional subspace from time-varying under-sampled linear projections of their sum. A novel solution approach called MatComProCS was introduced. Significantly improved performance over existing work was shown for undersampled fMRI based brain activity pattern detection.

Future work on this topic includes developing an improved version of MatComProCS that can give accurate detection performance with even fewer number of measurements; and developing an efficient algorithm for recursively solving the matrix completion problem as more data comes in (instead of re-solving the entire problem again). We are also looking at how to obtain performance guarantees for MatComProCS. The starting point will be the approach introduced in [2] for reprocs and the matrix completion results [7].

5. REFERENCES

- [1] T.T. Do, T.D. Tran, and L. Gan, "Fast compressive sampling with structurally random matrices," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2008.
- [2] C. Qiu, N. Vaswani, and L. Hogben, "Recursive robust pca or recursive sparse recovery in large but structured noise," *submitted to IEEE Trans. Info. Th., also at arXiv:1211.3754 [cs.IT]*, 2012.
- [3] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Th.*, vol. 51(12), pp. 4203–4215, Dec. 2005.
- [4] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Th.*, vol. 52(2), pp. 489–509, February 2006.
- [5] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Th.*, vol. 52(4), pp. 1289–1306, April 2006.
- [6] Emmanuel J. Candès and Benjamin Recht, "Exact matrix completion via convex optimization," *Commun. ACM*, vol. 55, no. 6, pp. 111–119, 2012.
- [7] E. Candes and Y. Plan, "Matrix completion with noise," *arXiv:0903.3131*, 2009.
- [8] K. Lee and Y. Bresler, "Admira: Atomic decomposition for minimum rank approximation," *IEEE Transactions on Information Theory*, vol. 56, no. 9, September 2010.
- [9] John Wright and Yi Ma, "Dense error correction via ℓ_1 -minimization," *IEEE Transactions on Information Theory*, 2009.
- [10] Jason N. Laska, Mark A. Davenport, and Richard G. Baraniuk, "Exact signal recovery from sparsely corrupted measurements through the pursuit of justice,".
- [11] Nam H. Nguyen and Trac D. Tran, "Robust lasso with missing and grossly corrupted observations," *To appear in IEEE Transaction on Information Theory*, 2012.
- [12] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *submitted to journal of Applied and Computational Harmonic Analysis*, July 2011.
- [13] G. Pope, A. Bracher, and C. Studer, "Probabilistic recovery guarantees for sparsely corrupted signals," *submitted to IEEE Transactions on Information Theory*, March 2012.
- [14] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM*, vol. 58 (1), pp. 1–37, 2009.
- [15] V. Chandrasekaran, S. Sanghavi, P.A. Parrilo, and A.S. Willsky, "Rank-sparsity incoherence for matrix decomposition," in *ArXiv preprint arXiv:0906.2220*, 2009.
- [16] J. Wright, A. Ganesh, K. Min, and Y. Ma, "Compressive principal component pursuit," 2012.
- [17] M. Mardani, G. Mateos, and G.B. Giannakis, "Exact recovery of low-rank plus compressed sparse matrices," *IEEE Statistical Signal Processing Workshop*, pp. 49–52, 2012.
- [18] Andrew E. Waters, Aswin C. Sankaranarayanan, and Richard G. Baraniuk, "Sparcs: Recovering low-rank and sparse matrices from compressive measurements," in *Neural Information Processing Systems (NIPS)*, 2011.
- [19] C. Qiu and N. Vaswani, "Real-time Robust Principal Components' Pursuit," in *Allerton Conf. Communication, Control, and Computing*, 2010.
- [20] C. Qiu and N. Vaswani, "Support predicted modified-cs for recursive robust principal components' pursuit," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2011.
- [21] C. Qiu and N. Vaswani, "Recursive sparse recovery in large but correlated noise," in *Allerton Conf. on Communication, Control, and Computing*, 2011.
- [22] Hong Jung, Kyunghyun Sung, Krishna S. Nayak, Eung Yeop Kim, and Jong Chul Ye, "k-t focuss: A general compressed sensing framework for high resolution dynamic mri," *Magnetic Resonance in Medicine*, vol. 61, pp. 103–116, 2009.
- [23] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic mri," *Magnetic Resonance in Medicine*, vol. 59(2), pp. 365–373, January 2008.
- [24] S. Gleichman and Y.C. Eldar, "Blind compressed sensing," *Information Theory, IEEE Transactions on*, vol. 57, no. 10, pp. 6958–6975, 2011.
- [25] S. Goud and M. Jacob, "A blind compressive sensing framework for accelerated dynamic mri," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2012.
- [26] W. Lu, T. Li, I. Atkinson, and N. Vaswani, "Modified-cs-residual for recursive reconstruction of highly undersampled functional mri sequences," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2011.
- [27] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Sig. Proc.*, vol. 58(9), pp. 4595–4607, September 2010.
- [28] E. Candes and T. Tao, "The dantzig selector: statistical estimation when p is much larger than n ," *Annals of Statistics*, vol. 35 (6), pp. 2313–2351, 2007.
- [29] I. C. Atkinson, D. L. Jones F. Kamalabadi, and K. Thulborn, "Blind estimation for localized low contrast-to-noise ratio bold signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 350–364, 2008.