

REAL-TIME DYNAMIC MR IMAGE RECONSTRUCTION USING KALMAN FILTERED COMPRESSED SENSING

Chenlu Qiu, Wei Lu and Namrata Vaswani

Dept. of Electrical and Computer Engineering, Iowa State University, Ames, IA,
 {chenlu, luwei, namrata}@iastate.edu

ABSTRACT

In recent work, Kalman Filtered Compressed Sensing (KF-CS) was proposed to causally reconstruct time sequences of sparse signals, from a limited number of “incoherent” measurements. In this work, we develop the KF-CS idea for causal reconstruction of medical image sequences from MR data. This is the first real application of KF-CS and is considerably more difficult than simulation data for a number of reasons, for example, the measurement matrix for MR is not as “incoherent” and the images are only compressible (not sparse). Greatly improved reconstruction results (as compared to CS and its recent modifications) on reconstructing cardiac and brain image sequences from dynamic MR data are shown.

Index Terms/Keywords: Compressed Sensing, Kalman Filtered Compressed Sensing, dynamic MRI

1. INTRODUCTION

In recent work [1], the problem of causally reconstructing time sequences of spatially sparse signals, with unknown and slow time-varying sparsity patterns, from a limited number of linear “incoherent” measurements was studied and a solution called Kalman Filtered Compressed Sensing (KF-CS) was proposed. An important example of this type of problems is real-time medical image sequence reconstruction using MRI, for e.g. dynamic MRI to image the beating heart or functional MRI to image the brain’s neuronal responses to changing stimuli (see Fig. 1). In these examples, the signal (heart or brain image) is approximately sparse (compressible) in the wavelet transform domain [2],[3]. MRI measures the 2D Fourier transform of the image which is known to be “incoherent” w.r.t. the wavelet basis [2]. Because MR data acquisition is sequential, the scan time (time to get enough data to accurately reconstruct one frame) is reduced if fewer measurements are needed for accurate reconstruction and hence there has been a lot of interest in the MRI community to use compressed sensing (CS) to do this [2],[3].

This idea was first demonstrated in [2] for a single MR image or volume. The work of [3] extended the idea to offline dynamic MRI reconstruction, i.e. it used the entire time sequence of measurements to jointly estimate the entire image sequence (treated it as a 3D x-y-t signal, sparse in wavelet domain along the x-y axis and sparse in the Fourier domain along the time axis). But this is a batch solution (needs all measurements first) and also the resulting joint optimization is computationally complex. On the other hand, the solution of [1] is causal and also much faster, and thus can be used to make dynamic MRI real-time. Reduced scan-time and real-time reconstruction are the currently missing abilities that prevent the use of MRI in interventional radiology applications, such as MR-guided surgery[4]. Some other recent work that also targets causal reconstruction of sparse image sequences is [5] and [6].

This research was partially supported by NSF grant ECCS-0725849

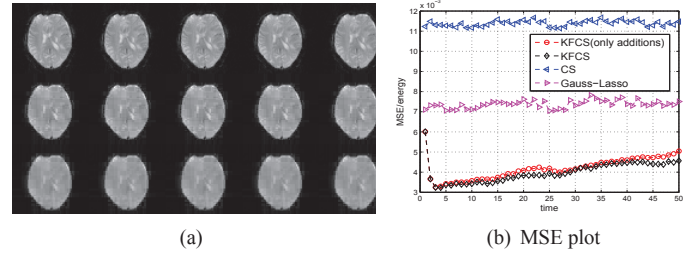


Fig. 1. Fig.1(a) Top: Original image sequence. Middle: MRI reconstructions using KF-CS. Bottom: MRI reconstruction using CS. Fig.1(b) is the corresponding MSE plot. $n = 2049$, $m = 4096$ and $\sigma_{obs}^2 = 100$.

In this work we use [1] to develop a KF-CS algorithm to causally reconstruct image sequences using MR data. There are some key differences in our current problem from the simplistic model used in [1] and these require some practical modifications to the algorithm of KF-CS (described in section 2). Additionally, in this work, (i) we develop a method for estimating the prior model parameters from training data (described in section 3.1) and (ii) we use the results of [7] to develop a method for selecting the number of observations required and the parameters used by the CS step of KF-CS (described in section 3.2). Results on reconstructing a cardiac sequence and a brain sequence are discussed in section 4 and they show greatly reduced mean squared error (MSE) when compared to performing CS at each time as in [2], as well as to some other modifications of CS. For e.g. in Fig 1b, the CS error is more twice that of KF-CS.

1.1. Problem Formulation

Let $(Z_t)_{m_1 \times m_1}$ denote the image at time t and let $m := m_1^2$ be its dimension. Let X_t denote the 2D discrete wavelet transform (DWT) of Z_t , i.e. $X_t := W Z_t W'$. Let F denote the discrete Fourier transform (DFT) matrix and $Y_{full,t} = F Z_t F' = F W' X_t W F'$ denote the 2D-DFT of Z_t . All of this can be transformed to a 1D problem by using Kronecker product denoted by \otimes . Let $y_{full,t} := \text{vec}(Y_{full,t})$ and $x_t := \text{vec}(X_t)$. Then $y_{full,t} = F_{1D} W_{1D} x_t$ where $F_{1D} = F \otimes F$ and $W_{1D} = W \otimes W$. Here, $\text{vec}(X_t)$ denotes the vectorization of the matrix X_t formed by stacking the columns of X_t into a single column vector. In MR imaging, we capture a set of n , ($n < m$), Fourier coefficients corrupted by white noise. This can be modeled by applying a $n \times m$ mask, M (which contains a single 1 at a different location in each row and all other entries are zero) to $y_{full,t}$ followed by adding Gaussian noise.

The above can be rewritten using the notation of [1] as

$$y_t = A x_t + w_t, A := H \Phi, H := M F_{1D}, \Phi := W_{1D} \quad (9)$$

with $w_t \sim \mathcal{N}(0, \sigma_{obs}^2)$ is i.i.d. Gaussian measurement noise. Let

Algorithm 1 Kalman Filtered Compressive Sensing for compressible or less-sparse signal

Initialization: At $t = 0$, compute $x^* = \arg\min_x \frac{1}{2} \|y - Ax\|_2^2 + \gamma \|x\|_1, \gamma = \gamma_{cs} = 2\sqrt{2 \log_2 m \sigma_{obs}}, T_0 = \{i \in [1 : m] : |x^*|_i > \alpha_{init}\}$.
Set $P_0 = 100 * Q, \hat{x}_0 = 0$
For $t > 0$, do

1. Temporary KF using $T \leftarrow T_{t-1}$:

$$(P_{t|t-1})_{T,T} = (P_{t-1})_{T,T} + (Q)_{T,T} \quad (1)$$

$$K_{t,tmp,T} = (\sigma_{obs}^2 (P_{t|t-1})_{T,T}^{-1} + A'_T A_T)^{-1} A'_T \quad (2)$$

$$(\hat{x}_{t,tmp})_T = (\hat{x}_{t-1})_T + K_{t,tmp,T} (y_t - A_T (\hat{x}_{t-1})_T), (\hat{x}_{t,tmp})_{T^c} = \mathbf{0} \quad (3)$$

2. Run CS on KF error to update nonzero set(Addition/Deletion)

- (a) Compute the filtering error, $\tilde{y}_{t,f} = y_t - A_T \hat{x}_{t,tmp}$, run CS on $\tilde{y}_{t,f}$, i.e. estimate $\hat{\beta} = \arg\min_{\beta} \frac{1}{2} \|y_{t,f} - A\beta\|_2^2 + \gamma \|\beta\|_1$ with $\gamma = \gamma_{ERC}$. Set $\hat{x}_{t,CSFE} = \hat{x}_{t,tmp} + \hat{\beta}$.
- (b) Addition/Deletion: estimate the nonzero set at t as: $T_c = \{i \in [1 : m] : |\hat{x}_{t,CSFE}|_i > \alpha_{add}\}$, set $T_t \leftarrow T_c$.

3. Run KF on the current nonzero set T_c :

$$(\hat{x}_{t|t-1})_{T_c} = \begin{pmatrix} (\hat{x}_{t-1})_{T \cap T_c, T \cap T_c} \\ \mathbf{0}_{T_c \setminus T, T_c \setminus T} \end{pmatrix}, (\hat{x}_{t|t-1})_{T_c^c} = \mathbf{0} \quad (4)$$

$$(P_{t|t-1})_{T_c, T_c} = \begin{bmatrix} (P_{t-1|t-1})_{T \cap T_c, T \cap T_c} & \\ & (P_0)_{T_c \setminus T, T_c \setminus T} \end{bmatrix} + (Q)_{T_c, T_c} \quad (5)$$

$$K_{t,T} = (\sigma_{obs}^2 (P_{t|t-1})_{T_c, T_c}^{-1} + A'_{T_c} A_{T_c})^{-1} A'_{T_c} \quad (6)$$

$$(P_t)_{T_c, T_c} = (I - K_{t,T} A_{T_c}) (P_{t|t-1})_{T_c, T_c} \quad (7)$$

$$(\hat{x}_{t|t})_{T_c} = (\hat{x}_{t|t-1})_{T_c} + K_{t,T} [y_t - A(\hat{x}_{t|t-1})_{T_c}], (\hat{x}_{t|t})_{T_c^c} = \mathbf{0} \quad (8)$$

4. Output T_t, \hat{x}_t and $\hat{x}_{t,CSFE}$. Compute signal estimation $\hat{z}_t = W_{1D} \hat{x}_t$ or $\hat{z}_{t,CSFE} = W_{1D} \hat{x}_{t,CSFE}$.
-

N_t denote the current set of nonzero coefficients (or significantly nonzero coefficients in case of compressible sequences). For $(x_t)_{N_t}$, we assume a spatially independent (but not identically distributed) Gaussian random walk model, i.e.

$$\begin{aligned} x_t &= x_{t-1} + \nu_t, \nu_t \sim \mathcal{N}(0, Q_t), \\ (Q_t)_{N_t, N_t} &= (Q)_{N_t, N_t} \\ (Q_t)_{N_t^c, [1:m]} &= 0, (Q_t)_{[1:m], N_t^c} = 0 \end{aligned} \quad (10)$$

where Q is a diagonal matrix estimated as explained in section 3.1. The diagonal assumption on Q is a valid one because the wavelet transform is well-known to be a decorrelating transform for natural images [8]. The set, N_t of (significantly) nonzero elements of x_t changes slowly over time, for e.g. for a 32×32 block of a cardiac image, $|N_t| \approx 130$ and $|N_t \setminus N_{t-1}| \leq 20$.

2. KF-CS FOR REAL-TIME DYNAMIC MRI

The overall idea of Kalman Filtered Compressed Sensing (KF-CS) for a sparse signal sequence is as follows [1]. Let T_t denote the KF-CS estimated set of (significantly) nonzero coefficients at time t , i.e. $T_t = \hat{N}_t$. At $t = 0$, we perform CS on y_0 followed by thresholding, to estimate T_0 . At any t , we first run a reduced order KF for the elements of T_{t-1} . Denote its output by $\hat{x}_{t,tmp}$. We use this to compute the filtering error in the observation, $\tilde{y}_{t,f} := y_t - A\hat{x}_{t,tmp}$. If $\tilde{y}_{t,f}$ is larger than usual (its weighted norm is greater than a threshold), it indicates that more coefficients have become nonzero. At this time, we perform CS on $\tilde{y}_{t,f}$ followed by thresholding the output to find new additions to T_{t-1} . This is followed by a KF prediction and update step for the current set of nonzero coefficients, T_t . Coefficients

that got falsely added in the past or that became zero over time are removed from T_t by thresholding the KF output \hat{x}_t .

Let $\Delta_t := N_t \setminus T_{t-1}$. The reason why KF-CS for sparse signals outperforms CS is that CS uses $y_t = Ax_t + w_t$ to estimate the $|T_{t-1} \cup \Delta_t|$ -sparse signal, x_t , while KF-CS uses $\tilde{y}_{t,f} = y_t - A\hat{x}_t = A\beta_t + w_t$ to estimate $\beta_t \triangleq x_t - \hat{x}_{t,tmp}$. As explained in [1], $\beta_t = [(x_t - \hat{x}_t)_{T_{t-1}}, (x_t)_{\Delta_t}, 0_{(T_{t-1} \cup \Delta_t)^c}]$ is also $|T_{t-1} \cup \Delta_t|$ -sparse but is only $|\Delta_t|$ -compressible (assuming $(x_t - \hat{x}_{t,tmp})_{T_{t-1}}$ is small).

Key Differences. KF-CS was designed for estimating a highly sparse signal sequence, with slowly changing nonzero elements' set, from a small number of random Gaussian projections corrupted in Gaussian noise. All nonzero coefficients at a given time were assumed to follow a random walk with equal variance. In the MRI reconstruction problem, there are a few key differences.

1. The matrix, $A = H\Phi$ is no longer random Gaussian. Φ is the DWT matrix and H contains randomly selected rows from the DFT matrix. The resulting A matrix is not as incoherent (incoherence can be quantified by $\mu = \max_{i \neq j} |A'_i A_j|$) as a random Gaussian matrix of the same size.
2. The wavelet transform coefficients' vector, x_t , of a real medical image (e.g. cardiac or brain) is only compressible (not sparse) and the number, S_t , of "significantly" nonzero coefficients (as a percentage of the signal size, m), is much larger than what was used in the simulations in [1].
3. Different wavelet coefficients have different variances, and in fact the nonzero coefficients that get added/deleted over time are typically the smaller variance ones.
4. The problem dimension, m is much larger (e.g. $m = 4096$).

Modifications. We describe below our modifications to address the above issues. The final algorithm is summarized in Algorithm 1.

Use of Basis Pursuit De-Noising (BPDN). Because of issue 2), the Dantzig selector, which was designed for very sparse signals, estimated from highly incoherent observations does not work well. We replace it by the BPDN (see step 2a of Algorithm 1), which was also used in [2].

Dealing with false additions and misses. Even with BPDN, because of issue 3), either many of the (significantly) nonzero coefficients never get added to T_t , i.e. $T_t \setminus N_t$ is large or there are too many false additions, $\Delta = T_t \setminus N_t$. In the latter case, the increase in $|T_t|$ may result in a singular $A_{T_t}^T A_{T_t}$ (making the KF unobservable). Because of issue 1), this begins to occur for smaller values of $|T_t|/n$ than if A were random Gaussian (and of course occurs more often when n is smaller, e.g. when $n = 308$ for a 32×32 cardiac image sequence).

To prevent this, the addition/deletion threshold, α_{add} , needs to be large, but this results in larger $\Delta = N_t \setminus T_t$. Note that the estimation error along $T = T_t$, $(x_t - \hat{x}_t)_T$, depends linearly on $A_T^T A_\Delta$. Because of larger sized Δ and because of 1), $A_T^T A_\Delta$ is no longer very small and thus $(x_t - \hat{x}_t)_T$ is also not very small. By using the output of the CS step, $\hat{x}_{t,CSFE} = \hat{x}_{t,tmp} + \hat{\beta}_t$, as the final output (instead of using the KF output, \hat{x}_t), we ensure that at least the CS estimate of $(\beta_t)_T = (x_t - \hat{x}_{t,tmp})_T$ is included in the final estimate. For the same reason, we also use $\hat{x}_{t,CSFE}$ for deletion, in fact we combine addition and deletion into a single step (see step 2b of Algorithm 1).

Initial covariance. Since there is an unknown delay in detecting new additions to the nonzero coefficients' set, the initial error covariance of newly added coefficients is never correctly known. We use an arbitrarily large value, $P_0 = 100Q$ for it (the estimate for the new coefficients will thus be closer to an LS solution).

Efficient Implementation. When m is large, a direct implementation of KF-CS becomes very slow. We make it faster using some simple reformulations such as solving a 2D version of BPDN and using the algorithm of [2] or using standard matrix algebra tricks to speed up matrix multiplications and inversions in the KF step.

3. PARAMETER ESTIMATION

Section 3.1 discuss how we estimate Q and section 3.2 discuss how we select n and γ using ERC.

3.1. Estimating Q

We consider two models for Q . The first one assumes Q to be a diagonal matrix with different entries while the second assumes Q is diagonal with all equal entries. The second model will have more bias (since finer scale wavelet coefficients always have smaller variance than coarser scale ones) but will have smaller variance and hence is a better idea when limited training data is available.

We can compute an approximate Maximum Likelihood (ML) estimate of Q under either model: we pick a zeroing threshold, set all coefficients below it to zero and use the rest of the coefficients to compute the ML estimate. The algorithm is summarized in Algorithm 2. $|\delta_i|$ in step 2 stands for the times of nonzero occurrence for i th entry of x_t and $|\delta_i| = 0$ implies the i th entry, $x_{t,i}$, is zero or equal at all time. We compare the two models for Q in Fig 2.

3.2. Selecting n and γ using ERC

The BPDN estimator is defined as the solution to

$$\arg \min_{x_t} \frac{1}{2} \|y_t - Ax_t\|_2^2 + \gamma \|x_t\|_1 \quad (11)$$

where γ is a regularization parameter that determines the tradeoff between the data consistency and the sparsity. We use Theorem 8

Algorithm 2 Q estimation

1. Zero out “compressible”(nearly zero) coefficients

(a) Select zeroing threshold

- For $t = 1 : L_{train}$, here, L_{train} denotes the length of training sequence,
 - Compute $x_t = W_{1D}' z_t$. Arrange $|x_{t,i}|$ in decreasing order of magnitude, i.e. $|x_{t,1}| > |x_{t,2}| > \dots > |x_{t,m}|$.
 - Compute the smallest S_t such that $\sum_{j \leq S_t} |x_{t,i}|^2 > 99.9\% x_t' x_t$, set threshold $\alpha_t = |x_{t,S_t}|$.
- Average α_t over t , $\alpha = \frac{\sum_t \alpha_t}{L_{train}}$.

(b) Zero out “compressible” coefficients

- For $t = 1 : L_{train}$, if $|x_{t,i}| < \alpha$, set $x_{t,i} = 0$. Set $N_t = \{i \in [1 : m] : |x_{t,i}| \geq \alpha\}$.
- ##### 2. Estimate Q using nonzero coefficients. Compute $\delta_i := \{t : x_{t,i} - x_{t-1,i} \neq 0\}$
- For Q with different diagonal values,
 - if $|\delta_i| \geq 1$, $(\sigma_{sys}^2)_i = \frac{1}{|\delta_i|} \sum_{t=2}^{L_{train}} (x_{t,i} - x_{t-1,i})^2$.
 - if $|\delta_i| = 0$, $(\sigma_{sys}^2)_i = 0.9 \min_{\{j: |\delta_j| \geq 0\}} (\sigma_{sys}^2)_j$
 - Set $Q_{diff} = \text{diag}((\sigma_{sys}^2)_i)$.
 - For Q with equal diagonal values,
 - Compute $\sigma_{sys}^2 = \frac{1}{\sum_j (|\delta_j|)} \sum_{i=1}^m \sum_{t=2}^{L_{train}} (x_{t,i} - x_{t-1,i})^2$. Set $Q_{same} = \sigma_{sys}^2 I$.

of [7] to develop an algorithm to select γ and n using training data. Let Λ be the nonzero set of x_t which we want to estimate. The exact recovery coefficient, $ERC(\Lambda) := 1 - \|A_\Lambda^\dagger A_{\Lambda^c}\|_{l1} > 0$ is one of the conditions for Theorem 8 to hold. Here, \dagger denotes pseudoinverse.

Selecting n : Given y_t and the ground truth data x_t . Threshold x_t to define its nonzero set N_t . Find the smallest value of n such that $ERC(N_t \setminus N_{t-1}) > 0$ for at least 90% of times.

Selecting γ : Theorem 8 of [7] guarantees that if γ is large enough (its correlation condition is satisfied), the BPDN estimator will have no falsely nonzero coefficients, but may end up not adding the small coefficients. For sufficiently sparse signal se-

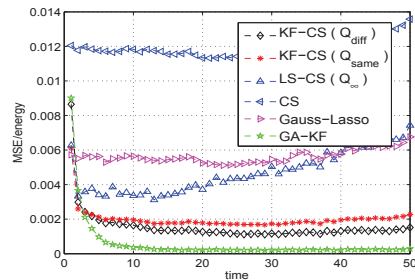


Fig. 2. Comparison of KF-CS with different model of Q . $n = 308, m = 1024$ and $\sigma_{obs}^2 = 25$. Q_∞ denotes $Q = \infty$. Gauss-BPDN is labeled as Gauss-Lasso by mistake.

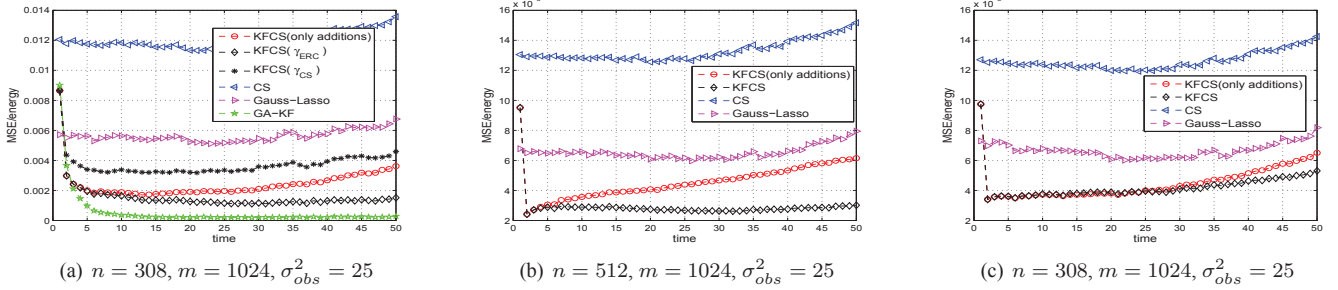


Fig. 3. Fig 3(a) is for a sparsified cardiac sequence. If deletion step is not used false additions become too large (result in unobservable KF and hence larger error). Fig 3(b), 3(c) are for the true cardiac sequence. Gauss-BPDN is labeled as Gauss-Lasso by mistake.

quences, one can use the correlation condition given in this result to find the minimum value of γ required for a given training sequence. This can be done as follows. Starting with $T_0 = N_0$, do the following for $t = 1$ to L_{train} : (a) compute $\Delta_t = N_t \setminus T_{t-1}$; (b) compute $ERC(\Delta_t)$; and (c) run step 2 of Algorithm 1 with $\gamma = \gamma_t = \frac{\|\tilde{y}_{t,f} - A_{\Delta_t} A_{\Delta_t}^\dagger \tilde{y}_{t,f}\|}{ERC(\Delta_t)}$. Use $\gamma_{ERC} = \max_t \gamma_t$ for test data.

In practice, for compressible data, using $\gamma = \gamma_{ERC}$ adds too few coefficients. In this case we compute γ_{ERC} using a sparsified version of the true training sequence, but use a smaller value of γ for the test data. This allows more coefficient additions.

4. EXPERIMENT RESULTS

We evaluated our algorithm on cardiac and a brain image sequences. Fig. 1 is the comparison of KF-CS reconstructed image sequence with that of CS for brain data. Notice the white region in the center is much more blurred in CS reconstruction than in KFCS. For the results of Fig 2 and 3, we selected a 32x32 region in the cardiac data and used its sequence as test data. MRI was simulated by taking the 2D-DFT of the given image sequence, selecting a random set of n Fourier coefficients using the variable-density undersampling scheme proposed in [2] and adding i.i.d. Gaussian noise to them. In the plots of Fig 2, 3, we plot MSE/energy $:= \mathbb{E} \|x_t - \hat{x}_t\|_2^2 / \mathbb{E} \|x_t\|_2^2$ which was computed by averaging over 50 Monte Carlo simulations for the cardiac sequence. Fig. 2 compares KF-CS using Q_{diff} , Q_{same} and using $Q = \infty$ (replacing KF by LS). $n = 308$ observations were taken while $|N_t| \approx 130$ and $|\Delta_t| \leq 20$. Since σ_{obs}^2 is large, KF-CS which makes use of prior model knowledge performs much better than LS-CS. Also, KF-CS with Q_{diff} outperforms KFCS with Q_{same} .

In Fig3(a), we show that KF-CS with $\gamma_{ERC} = 2$ (we use a slightly simplified version of the method of section 3.2 for computing this) outperforms KF-CS with a larger choice of γ ($\gamma = \gamma_{cs} = 2\sqrt{2\log_2 m\sigma_{obs}}$), KF-CS (only additions), Gauss-BPDN and CS[2]. The comparison is done for sparsified cardiac sequence (image sequence obtained by zeroing out small wavelet coefficients and computing inverse wavelet transform). KF-CS (only additions) refers to algorithm proposed in [1], in which we kept adding new additions to the nonzero coefficients set at each t without dealing with false additions. For Gauss-BPDN, we first estimated nonzero set by BPDN, followed by LS on it (similar to Gauss-Dantzig[9]). Genie-aided KF (GA-KF) is a KF that knows N_t at each t . This serves as a MSE lower bound. Fig3(b) and Fig3(c) compare KF-CS with KF-CS (only additions), CS and Gauss-BPDN for true (not sparsified) cardiac data. In Fig3(b), we set $\gamma = 2 < \gamma_{ERC}$ while $n = 512$ and $m = 1024$. In Fig3(c), we set $\gamma = 0.005 \ll \gamma_{ERC}$ while $n = 308 \ll 512$. When n is very small, we use smaller γ

and only run KF on "larger" coefficients while use \hat{x}_{CSFE} as final output.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we have developed the KF-CS idea for causal reconstruction of medical image sequences from MR data and have shown greatly improved reconstruction results on cardiac dynamic and brain fMRI data, as compared to CS [2] and its modifications. This is the first real application of KF-CS and is considerably more difficult than simulation data because the measurement matrix for MR is not as incoherent as a random Gaussian matrix and because the different wavelet coefficients have vastly different magnitudes and variances. Future work will involve a rigorous analysis of the proposed algorithmic ideas and using it to propose a novel KF-CS based algorithm for compressible sequences.

6. REFERENCES

- [1] N. Vaswani, "Kalman filtered compressed sensing," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2008.
- [2] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse mri: The application of compressed sensing for rapid mr imaging," *Magnetic Resonance in Medicine*, vol. 58(6), pp. 1182–1195, Dec. 2007.
- [3] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic mri," *Magnetic Resonance in Medicine*, vol. 59(2), pp. 365–373, January 2008.
- [4] A. J. Martin, O. M. Weber, D. Saloner, R. Higashida, M. Wilson, M. Saeed, and C. B. Higgins, "Application of MR Technology to Endovascular Interventions in an XMR Suite," *Medica Mundi*, vol. 46, December 2002.
- [5] C. Rozell, D. Johnson, R. Baraniuk, and B. Olshausen, "Locally competitive algorithms for sparse approximation," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2007.
- [6] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t focuss: a general compressed sensing framework for high resolution dynamic mri," *Magnetic Resonance in Medicine*, vol. 35, No. 6, 2313–2351, Dec. 2007.
- [7] Joel A. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. on Info. Th.*, vol. 52, no. 3, March 2006.
- [8] Anil K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall of India, 1995.
- [9] E. Candes and T. Tao, "The dantzig selector: statistical estimation when p is much larger than n ," *Annals of Statistics*, 2006.