Sparsity-fused Kalman Filtering for Reconstruction of Dynamic Sparse Signals

Xin Ding*, Wei Chen†*, Ian Wassell*
*Computer Laboratory, University of Cambridge, Cambridge, UK
†State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China
Email: xd225, wc253, ijw24 @cam.ac.uk

***Abstract*—This article focuses on the problem of reconstructing dynamic sparse signals from a series of noisy compressive sensing measurements using a Kalman Filter (KF). This problem arises in many applications, e.g., Magnetic Resonance Imaging (MRI), Wireless Sensor Networks (WSN) and video reconstruction. The conventional KF does not consider the sparsity structure presented in most practical signals and it is therefore inaccurate when being applied to sparse signal recovery. To deal with this issue, we derive a novel KF procedure which takes the sparsity model into consideration. Furthermore, an algorithm, namely Sparsity-fused KF, is proposed based upon it. The method of iterative soft thresholding is utilized to refine our sparsity model. The superiority of our method is demonstrated by synthetic data and the practical data gathered by a WSN.**

I. Introduction

Most practical signals have a sparse structure in a suitable basis and the sparse signal model has made dramatic improvements on solving linear inverse problems. Based on this, Compressive Sensing (CS) [1], [2], [3], has emerged as a new sampling paradigm, which allows sparse signals to be recovered accurately from far fewer measurements as required by the Nyquist sampling rate. CS for static sparse signal recovery has been heavily investigated, e.g., [4], [5], [6], [7]. Compared to the static case, dynamic sparse signals have more correlated statistics that can be exploited. Improved CS schemes using dynamic correlation features have been successfully developed in many applications, e.g., MRI [8], Wireless Sensor Networks (WSNs) [9] and video [10].

The Kalman Filter (KF) has been one of the most popular approaches in the area of linear dynamic system modeling. However, the KF is not an ideal method when it comes to dynamic sparse signals, as the KF in general does not consider any sparsity constraints during the state estimation. In [11], the first KF based CS method was proposed. They employed CS on the KF residual, identified the support set of the signals and then carried out a second KF corresponding to the obtained set. The signal sparsity is enforced by the support detection, rather than by changing the underlying principle of the KF. In [12], they transformed the KF procedure as a MAP estimator and included the sparse constraint as in CS-type algorithms. By exploiting the inherent statistics of the KF, the authors of [13] proposed to employ a hierarchical Bayesian network to capture the sparsity. However they only focused on the sparsity of the innovations between correlated signals. In [14], they formulated the KF into the optimization framework, to which it is straightforward to add any sparsity constraints. But it lacks the updates on the covariance as in conventional KF. Using the constrained optimization problem, Carmi proposed a stand-alone KF-based reconstruction approach by employing pseudo-measurements [15]. To our best knowledge, all the existing methods are either enforcing sparsity by estimating the signal amplitudes and support separately or pursuing sparsity by introducing the optimization problem into the original KF in various ways. The non-sparse estimating framework underlying the KF has not been altered.

In this article, we propose a novel KF process which has directly fused the signal sparsity model into its filtering process and we then develop an algorithm for the reconstruction of dynamic sparse signals based on the sparsity-fused KF. We pursue the sparsity of signals via Iterative Soft Thresholding (IST), which has been proved to be effective in estimating sparse signals [16]. Our method outperforms the existing algorithms in the following respects: 1) our reconstruction results for CS acquired dynamic sparse signals are more accurate in the practical noisy case; 2) it is more robust when the measurement noise increases; 3) it converges in fewer iterations; 4) it is easier to tune the parameters for our algorithm.

The rest of the paper is organized as follows: Section II provides the mathematical formulation of the problem; section III presents the details of our work, including the sparsity statistic model, the novel sparsity-fused KF process and our proposed algorithm; section IV demonstrates the experimental results for both the synthetic and practical data; section V concludes this article.

II. Problem Formulation

Consider a sequence of signals $\mathbf{x}_t \in \mathbb{R}^N$ $(t = 1, ...T)$. Each of the signals is sparse in some sparsifying basis $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$, i.e., $\mathbf{s}_t = \mathbf{\Psi}^*\mathbf{x}_t$, where $\mathbf{s}_t$ has only K $(K \ll N)$ non-zero coefficients and $\mathbf{\Psi}^*$ denotes the conjugate transpose of $\mathbf{\Psi}$. We sample this sequence of sparse signals via:

$$\mathbf{y}_t = \mathbf{A}_t\mathbf{s}_t + \mathbf{e}_t, \tag{1}$$

where $\mathbf{A}_t = \mathbf{\Phi}_t\mathbf{\Psi}$ is the equivalent projection matrix, $\mathbf{\Phi}_t \in \mathbb{R}^{M\times N}$ is the sensing matrix, $\mathbf{y}_t \in \mathbb{R}^M (M \ll N)$ is the t-th observation vector and $\mathbf{e}_t \in \mathbb{R}^M$ is the measurement noise.

This is an under-determined system and conventionally it is impossible to recover $\mathbf{s}_t$ from $\mathbf{y}_t$. However, under the condition

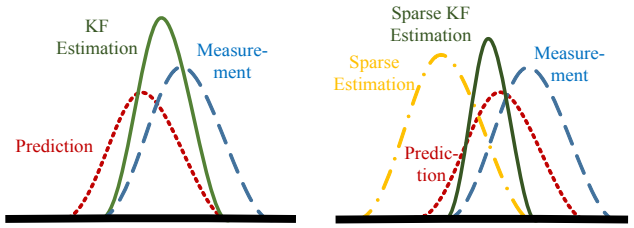

Fig. 1. Fusion process for (a) the conventional KF; red: prediction distribution $p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})$; blue: measurements distribution $p(\mathbf{y}_t|\mathbf{s}_t)$; green: KF estimation $p(\mathbf{s}_t|\mathbf{y}_t,\hat{\mathbf{s}}_{t-1})$; (b) sparsity-fused KF; red and blue curves represent the same distributions as in (a); yellow: distribution of sparse estimation $p(\mathbf{s}_t)$; green: sparsity-fused KF estimation $p(\mathbf{s}_t|\mathbf{y}_t,\hat{\mathbf{s}}_{t-1})$.

that $\mathbf{s}_t$ is sparse, if $\mathbf{A}_t$ satisfies the Restricted Isometry Property (RIP) [17], CS offers a way to solve the problem with an overwhelming probability of success. Conventional CS reconstructs $\mathbf{s}_t$ by solving the following optimization problem:

$$\min_{\mathbf{s}_t} ||\mathbf{s}_t||_1 \; s.t. \; ||\mathbf{y}_t - \mathbf{A}_t\mathbf{s}_t||_2^2 \leq \epsilon, \tag{2}$$

where $||\mathbf{v}||_p = (\sum_{j=1}^{N} \mathbf{v}_j^p)^{1/p}$ is the l_p-norm of $\mathbf{v}$ and ϵ is a tolerance parameter. So reconstructing dynamic sparse signals one by one using the conventional CS as in (2) does not exploit any correlations between the signals.

We model our system as a linear dynamic system which follows:

$$\mathbf{s}_t = \mathbf{F}_t\mathbf{s}_{t-1} + \mathbf{q}_t, \tag{3}$$

where $\mathbf{F}_t$ denotes the state transition matrix from the $(t-1)$-th signal to the t-th signal and $\mathbf{q}_t$ represents the process noise. For the system defined by (1) and (3), if we assume the statistics of $\mathbf{s}_{t-1}$ are known as $p(\mathbf{s}_{t-1}) \sim \mathcal{N}(\hat{\mathbf{s}}_{t-1}, \mathbf{P}_{t-1})$ and we further assume the following Gaussian distributions: $p(\mathbf{e}_t) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$, $p(\mathbf{q}_t) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$, $p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1}) \sim \mathcal{N}(\mathbf{F}_t\hat{\mathbf{s}}_{t-1}, \mathbf{P}_{t|t-1})$ and $p(\mathbf{y}_t|\mathbf{s}_t) = \mathcal{N}(\mathbf{A}_t\mathbf{s}_t, \mathbf{R}_t)$, it becomes a standard KF setting. Under these assumptions, the KF provides a solution that both minimizes the MSE, $E[||\mathbf{s}_t - \hat{\mathbf{s}}_t||_2^2]$, and maximizes a posterior probability, i.e.,

$$\hat{\mathbf{s}}_t = \underset{\mathbf{s}_t}{arg\,max}\, p(\mathbf{s}_t|\mathbf{y}_t, \hat{\mathbf{s}}_{t-1}). \tag{4}$$

See [18] for the original KF equations. The process for the conventional KF includes two stages: predicting the current signal using the previous one and updating the estimate using the measurements. It utilizes the correlations between dynamic signals, but the solution $\hat{\mathbf{s}}_t$ is generally not sparse. This is highly inaccurate when being applied to the recovery of sparse signals. We therefore propose to design a new KF process that does not ignore the sparse structure of the signals.

III. Sparsity-fused Kalman Filtering

From a statistical view, the original KF essentially approximates the probability in (4) as:

$$\begin{aligned} p(\mathbf{s}_t|\mathbf{y}_t,\hat{\mathbf{s}}_{t-1}) &\propto p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})p(\mathbf{s}_t|\mathbf{y}_t) \\ &= p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})[p(\mathbf{y}_t|\mathbf{s}_t)p(\mathbf{s}_t)/p(\mathbf{y}_t)] \\ &\propto p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})p(\mathbf{y}_t|\mathbf{s}_t), \end{aligned} \tag{5}$$

where $p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})$ denotes the predict step and $p(\mathbf{y}_t|\mathbf{s}_t)$ represents the update step. Accordingly, the operation of the KF can be explained by Fig. 1(a). The green curve that represents the distribution of the KF estimation, i.e., $p(\mathbf{s}_t|\mathbf{y}_t,\hat{\mathbf{s}}_{t-1})$, is derived by multiplying the red and blue curves, i.e., $p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})$ and $p(\mathbf{y}_t|\mathbf{s}_t)$, respectively. We refer to this process as "fusion". Obviously, in this fusion process, $p(\mathbf{s}_t)$ is approximated as a uniform distribution, i.e., the conventional fusion process does not include any probability related to the sparsity of the signal.

In our method, we take into consideration the signal sparsity constraint, which gives:

$$p(\mathbf{s}_t|\mathbf{y}_t,\hat{\mathbf{s}}_{t-1}) \propto p(\mathbf{s}_t|\hat{\mathbf{s}}_{t-1})p(\mathbf{y}_t|\mathbf{s}_t)p(\mathbf{s}_t), \tag{6}$$

where $p(\mathbf{s}_t)$ is the distribution modeling the sparsity of the signal. As illustrated in Fig. 1(b), we derive our sparsity-fused KF estimation (the green curve) by fusing the distributions of the prediction (the red curve), measurements (the blue curve) and the sparse estimation (the yellow curve).

A. Sparsity Distribution Model

We model the sparse signal distribution by:

$$p(\mathbf{s}_t) \sim \mathcal{N}(\mathbf{0}, \mathbf{T}_t), \; \mathbf{T}_t = \mathrm{diag}(\boldsymbol{\lambda}) = \mathrm{diag}([\lambda_1, ..., \lambda_N]), \tag{7}$$

where the vector $\boldsymbol{\lambda}$ denotes the variance of the signal. As noticed, the sparsity is transferred from the amplitude of the signal to the variance of it. If we denote $\mathbf{s}_t = [\mathrm{s}_1, \mathrm{s}_2, ..., \mathrm{s}_N]^T$, the corresponding signal amplitude at each index is

$$\mathrm{s}_i = \underset{\lambda_i}{arg\,max} \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{\mathrm{s}_i^2}{2\lambda_i}}, \tag{8}$$

where $i = 1, ..., N$. Thus, we can derive that the relationship between the signal amplitudes and the variance in our model is:

$$\mathrm{s}_i^2 = \lambda_i. \tag{9}$$

This relationship paves the way to apply any conventional sparsity pursuing method that functions on signal amplitudes, rather than the variance, into our system.

B. Sparsity-fused KF update steps

Because of the difference between (5) and (6), the original KF update steps are no longer applicable to our model. We derive the following novel KF updating equations based on (6):

$$\mathbf{M}_t = \mathbf{P}_t(\mathbf{P}_t + \mathbf{T}_t)^{-1}, \tag{10}$$

$$\tilde{\mathbf{s}}_t = (\mathbf{I} - \mathbf{M}_t)\hat{\mathbf{s}}_t, \tag{11}$$

$$\tilde{\mathbf{P}}_t = (\mathbf{I} - \mathbf{M}_t)\mathbf{P}_t, \tag{12}$$

where $\mathbf{P}_t$ and $\hat{\mathbf{s}}_t$ are, respectively, the covariance matrix and the state estimate in the original KF, $\mathbf{T}_t$ is the variance matrix for the sparsity model as defined in (7) and $\mathbf{I}$ is an identity matrix. The resulting $\mathbf{M}_t$, $\tilde{\mathbf{s}}_t$, $\tilde{\mathbf{P}}_t$ are the Sparsity-fused KF gain, the sparse state estimate and the corresponding covariance matrix, respectively.

Derivation: Define the green Gaussian function in Fig. 1(a) as $y_1(\mathrm{s}; \mu_1, \sigma_1) = (1/\sqrt{2\pi\sigma_1^2})exp[-(\mathrm{s}-\mu_1)^2/2\sigma_1^2]$ *and the yellow Gaussian function in Fig. 1(b) as* $y_2(\mathrm{s}; 0, \sigma_2) = (1/\sqrt{2\pi\sigma_2^2})exp(-\mathrm{s}^2/2\sigma_1^2)$. *Note that the process of deriving the green Gaussian distribution in Fig. 1(b) is equivalent to*

Algorithm 1 Sparsity-fused KF

Input: $\mathbf{y}_t(t=1,..,T), \mathbf{A}, \mathbf{F}_t, \mathbf{Q}, \mathbf{R}=\sigma_0^2\mathbf{I}, \hat{s}_0, \mathbf{P}_0, \gamma, N, N_\tau$.
Output: $\hat{\mathbf{s}}_t(t=1,...,T)$.

for $t=1$ to T **do**
  Run the original KF:

$$\hat{\mathbf{s}}_{t|t-1}=\mathbf{F}_t\hat{\mathbf{s}}_{t-1}, \mathbf{P}_{t|t-1}=\mathbf{F}_t\mathbf{P}_{t-1}\mathbf{F}_t^*+\mathbf{Q}, \quad (14)$$

$$\mathbf{K}_t=\mathbf{P}_{t|t-1}\mathbf{A}^*(\mathbf{A}\mathbf{P}_{t|t-1}\mathbf{A}^*+\mathbf{R})^{-1}, \quad (15)$$

$$\hat{\mathbf{s}}_t=\hat{\mathbf{s}}_{t|t-1}+\mathbf{K}_t(\mathbf{y}_t-\mathbf{A}\hat{\mathbf{s}}_{t|t-1}), \quad (16)$$

$$\mathbf{P}_t=\mathbf{P}_{t|t-1}-\mathbf{K}_t\mathbf{A}\mathbf{P}_{t|t-1}, \quad (17)$$

  Initialization: $\mathbf{s}_t^0=\hat{\mathbf{s}}_t$, $\mathbf{P}^0=\mathbf{P}_t$.
  for $\tau=1,2,...,N_\tau$ **do**
    Thresholding:

$$\sigma^\tau=\sqrt{(\mathbf{P}^{\tau-1}(1,1)+...+\mathbf{P}^{\tau-1}(N,N))/N}, \quad (18)$$

$$\mathbf{r}_t^{\tau-1}=\mathbf{y}_t-\mathbf{A}\mathbf{s}_t^{\tau-1}, \quad (19)$$

$$\check{\mathbf{s}}_t^\tau=\eta^\tau(\mathbf{A}^*\mathbf{r}_t^{\tau-1}+\mathbf{s}_t^{\tau-1};\gamma\sigma^\tau). \quad (20)$$

    Derive $\mathbf{T}$: $\mathbf{T}_t^\tau=diag([\check{\mathbf{s}}_t^\tau(1)^2,\check{\mathbf{s}}_t^\tau(2)^2,...,\check{\mathbf{s}}_t^\tau(N)^2])$.
    Run sparsity-fused KF:

$$\mathbf{M}_t^\tau=\mathbf{P}_t(\mathbf{P}_t+\mathbf{T}_t^\tau)^{-1}, \quad (21)$$

$$\mathbf{s}_t^\tau=(\mathbf{I}-\mathbf{M}_t^\tau)\hat{\mathbf{s}}_t, \quad (22)$$

$$\mathbf{P}^\tau=(\mathbf{I}-\mathbf{M}_t^\tau)\mathbf{P}_t. \quad (23)$$

  end for
  Update: $\hat{\mathbf{s}}_t=\mathbf{s}_t^\tau, \mathbf{P}_t=\mathbf{P}^\tau$.
end for

fusing y_1 and y_2. Follow the scalar derivation process in [18], we can get the expression for the fused function as:

$$y_{fused}(\mathrm{s};\mu_{fused},\sigma_{fused})=\frac{1}{\sqrt{2\pi\sigma_{fused}^2}}e^{-\frac{(s-\mu_{fused})^2}{2\sigma_{fused}^2}}, \quad (13)$$

where $\mu_{fused}=\mu_1-\frac{\sigma_1^2\mu_1}{(\sigma_1^2+\sigma_2^2)}$ and $\sigma_{fused}^2=\sigma_1^2-\frac{\sigma_1^4}{(\sigma_1^2+\sigma_2^2)}$. Substitute the scalars with the KF terms as follows: $\mu_{fused}\rightarrow\tilde{\mathbf{s}}_t$, $\mu_1\rightarrow\hat{\mathbf{s}}_t$, $\sigma_{fused}^2\rightarrow\tilde{\mathbf{P}}_t$, $\sigma_1^2\rightarrow\mathbf{P}_t$, $\sigma_2^2\rightarrow\mathbf{T}_t$, $M=\sigma_1^2/(\sigma_1^2+\sigma_2^2)\rightarrow\mathbf{M}_t$, the sparsity-fused KF update equations (10)-(12) are derived. □

C. Proposed Algorithm for the Recovery of Dynamic Sparse Signals via IST

Based on the sparse signal model and the newly-derived KF process, we propose our Sparsity-fused KF algorithm, as presented in Algorithm 1. The algorithm proceeds as follows. For each of the signals in the sequence, an original KF is performed at first, based on which the thresholding and the sparsity-fused KF steps are initialized. Then we employ soft thresholding (with a threshold function η) to pursue the sparsity of the signal, in which the threshold σ is determined by the KF covariance matrix $\mathbf{P}$. The threshold function is defined as:

$$\eta^\tau(u;\gamma\sigma^\tau)=\begin{cases}u-\gamma\sigma^\tau & if\ u\geq\gamma\sigma^\tau,\\ u+\gamma\sigma^\tau & if\ u\leq-\gamma\sigma^\tau,\\ 0 & elsewhere,\end{cases} \quad (24)$$

where τ is the index for the τ-th iteration, γ is a control parameter, u is a scalar element and the the function η is applied component-wise when applied to a vector. Then, according to the result of thresholding and the relationship we obtained in (9), the variance matrix $\mathbf{T}$ of the sparsity distribution can be derived. It is then utilized by the sparsity-fused KF procedure to obtain the reconstruction result for this iteration. The steps of the thresholding and the sparsity-fused KF are repeated to refine the reconstruction iteratively. Serially, the same process is carried out for the next signal in the sequence.

In the area of static sparse signal recovery, iterative thresholding schemes have been popular for some years [19]. Such schemes are attractive because they have very low computational complexity and low storage requirements. However, there are two issues concerning the implementation of such methods. Firstly, it is essential to choose the thresholds very carefully and they need to vary from iteration to iteration. In [20], they carried out extensive experiments and finally proposed how to optimally tune the parameters. Secondly, these methods behave poorly in the presence of noise.

Fortunately, these issues are not problematic in the context of our algorithm. Specifically, for the soft threshold function η that we employed as in [6], the thresholds σ is defined as $(\sigma^\tau)^2=E[||\mathbf{s}^\tau-\mathbf{s}||]_2^2$, which is the estimate of the current MSE. In conventional IST, an empirical estimate of $(\sigma^\tau)^2$ is used as the true MSE is unknown. But in our case, it is much simpler as the KF process happened to provide us such an estimate in the covariance matrix $\mathbf{P}$. Therefore, we define σ^τ as in (18). Regarding the second issue, we benefit from the correlations between the signals. Our reconstruction depends on both the previous signal and the current measurements, rather than purely rely on the noisy measurements as in the conventional IST. We will show by simulation that compare to the existing algorithms, our algorithm is more robust when the noise presents in Section IV.

D. Heuristics for the Proposed Algorithm

We shall start with the heuristics for the conventional IST algorithm in the static case [6]. Define the operator $\mathbf{H}=\mathbf{A}^*\mathbf{A}-\mathbf{I}$. In the first iteration of the IST, an estimate of the true signal $\mathbf{s}$ is derived as:

$$\mathbf{s}^1=\eta(\mathbf{A}^*\mathbf{y})=\eta(\mathbf{s}+\mathbf{H}\mathbf{s}), \quad (25)$$

where $\mathbf{H}\mathbf{s}$ can be regarded as a noisy random vector with i.i.d. Gaussian entries with variance $N^{-1}||\mathbf{s}||_2^2$. It is well-understood that the soft thresholding results in $\mathbf{s}^1$ being closer to $\mathbf{s}$ than $\mathbf{A}^*\mathbf{y}$ [6]. Then in the next iteration, the estimate is:

$$\mathbf{s}^2=\eta[\mathbf{A}^*(\mathbf{y}-\mathbf{A}\mathbf{s}^1)+\mathbf{s}^1]=\eta[\mathbf{s}+\mathbf{H}(\mathbf{s}-\mathbf{s}^1)], \quad (26)$$

where the noisy vector is again a Gaussian vector with variance $N^{-1}||\mathbf{s}-\mathbf{s}^1||_2^2$. We can see that the noise level is lower than the last iteration, so the thresholding result can be anticipated to be more accurate at this iteration, and so on.

Now consider the proposed algorithm. In the first iteration $\tau=1$ at t, we get an estimate as:

$$\check{\mathbf{s}}_t^1=\eta[\mathbf{A}^*(\mathbf{y}_t-\mathbf{A}\mathbf{s}_t^0)+\mathbf{s}_t^0]=\eta[\mathbf{s}_t+\mathbf{H}(\mathbf{s}_t-\mathbf{s}_t^0)], \quad (27)$$

where the noise vector has a variance of $N^{-1}||\mathbf{s}_t-\mathbf{s}_t^0||_2^2$ and $\mathbf{s}_t^0$ is the initialization by the conventional KF. Therefore, if the

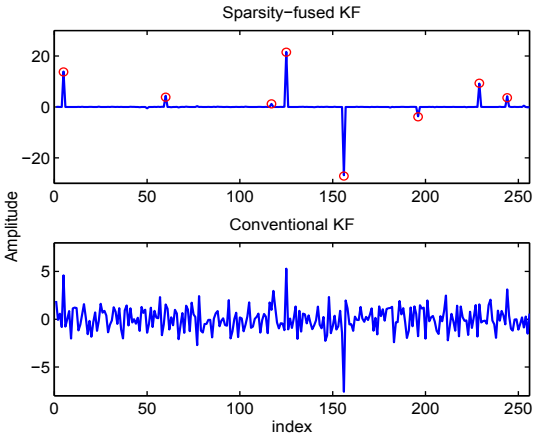

Fig. 2. An example of reconstructed signal using the Sparsity-fused KF algorithm and the conventional KF. The red circles denote the amplitudes of the original signals for all its non-zero coefficients. ($t = 100$, $M = 72$, $N = 256$, $\sigma_0 = 0.2$.)

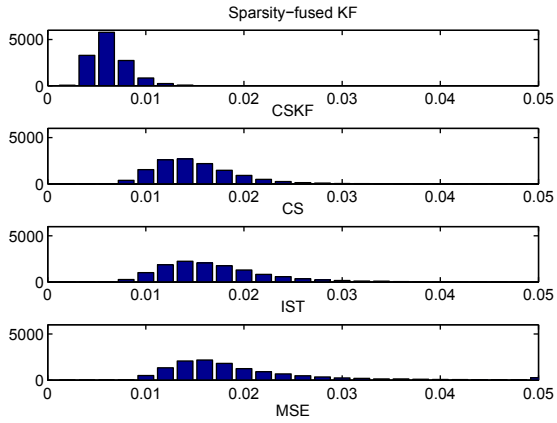

Fig. 3. Histogram of the MSE for all 13000 reconstructions for various algorithms. ($T = 130$, $M = 72$, $N = 256$, $\sigma_0 = 0.2$.)

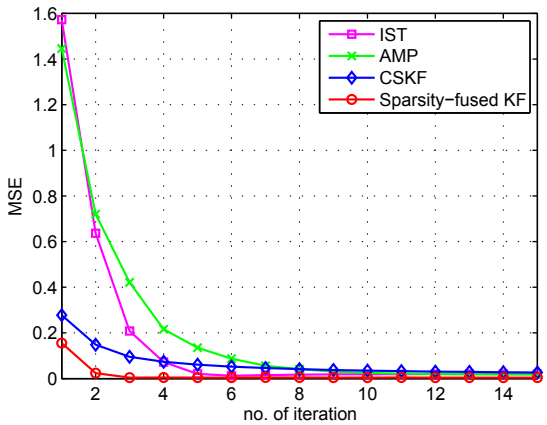

Fig. 4. The track of the MSE at different number of iterations for various algorithms. ($T = 130$, $M = 72$, $N = 256$, $\sigma_0 = 0.2$.)

original KF step produces an acceptable result so that $||\mathbf{s}_t - \mathbf{s}_t^0||_2^2 < ||\mathbf{s}_t||_2^2$, we can anticipate that our estimate $\check{\mathbf{s}}_t^1$ is closer to $\mathbf{s}_t$ than the one in (25). Then in this case, by fusing it with the distribution, which contains the information of both the previous signal and the measurements, that we obtained in the original KF step, the fused estimation $\mathbf{s}_t^1$ is even closer to $\mathbf{s}_t$ than $\check{\mathbf{s}}_t^1$.

In the next iteration, the estimate is:

$$\check{\mathbf{s}}_t^2 = \eta[\mathbf{A}^*(\mathbf{y}_t - \mathbf{A}\mathbf{s}_t^1) + \mathbf{s}_t^1] = \eta[\mathbf{s}_t + \mathbf{H}(\mathbf{s}_t - \mathbf{s}_t^1)], \quad (28)$$

where the noise vector has a variance of $N^{-1}||\mathbf{s}_t - \mathbf{s}_t^1||_2^2$. As shown previously, when our estimation in the first iteration, i.e., $\mathbf{s}_t^1$, is closer to $\mathbf{s}_t$ than in the conventional IST, i.e., $\mathbf{s}^1$ in (25), the noise level in (28) is lower than that in (26). In consequence, we can anticipate that our estimate in each iteration is more accurate than that of the original IST and so our algorithm will converge in fewer iterations than the conventional IST. This will be shown by simulation in Section IV. Besides, a better recovery of $\mathbf{s}_t$ will help to create a more accurate $\mathbf{s}_{t+1}^0$, which is crucial to the reconstruction of the $(t+1)$-th signal.

IV. Experimental Results

The performance of the Sparsity-fused KF algorithm is tested using several examples in which dynamic sparse signals are recovered from noisy observations. We first test it using synthetic signals and then employ it to reconstruct practical signals obtained in a WSN.

A. Experiments on Synthetic Signals

We generate a sequence of sparse signals ($T = 130$) which behave as a random walk process as following:

$$\mathbf{s}_{t+1}(i) = \begin{cases} \mathbf{s}_t(i) + \mathbf{q}_t(i), & if,\ \mathbf{s}_t(i) \in supp(\mathbf{s}_t) \\ 0, & otherwise, \end{cases} \quad (29)$$

where i is the index, $\mathbf{s}_t \in \mathbb{R}^{256}$ is assumed to be a sparse vector and its support consists of 8 elements of which the index is uniformly sampled in the interval $[1, 256]$. The process noise obeys $\mathbf{q}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and the covariance matrix $\mathbf{Q}$ is taken as a diagonal matrix with entries equal to 1 on its diagonal. The transition matrix $\mathbf{F}_t$ is taken as an identity matrix for $t = 1, ..., T$. For all the simulations, we employ the i.i.d. Gaussian matrix as the sensing matrix $\mathbf{A}$ and $\mathbf{R} = \sigma_0^2\mathbf{I}$ with various values of σ_0 as the measurement noise covariance matrix.

In the first experiment, we take 72 samples at each t and try to recover the signals when $\sigma_0 = 0.2$. For initialization, we set $\hat{\mathbf{s}}_0 = \mathbf{0}$ and $\mathbf{P}_0$ as a diagonal matrix whose entries are all 9 [11]. Let $\gamma = 2$ [21] and the number of iteration $N_\tau = 100$. Fig. 2 shows an example of the signal reconstructed by our algorithm and the conventional KF method. We can see that our method manages to recover the signal with a sparse structure; while the conventional KF fails to do so as it does not use any information about the sparsity.

Under the same parameter settings, we then compare our algorithm with the CS-embedded KF method for dynamic signals in [15] (denoted as CSKF), as well as the conventional static methods, i.e., CS and IST. We set $R_\epsilon = 200^2$, $N_\tau = 100$ for the CSKF algorithm and the IST algorithm is optimally tuned [20]. 100 trials for each algorithm are carried out. Fig. 3 illustrates the distributions of the MSE for all the $T \times 100 = 13000$ reconstructions. It is clear that the performance of our algorithm surpasses the others. Most of

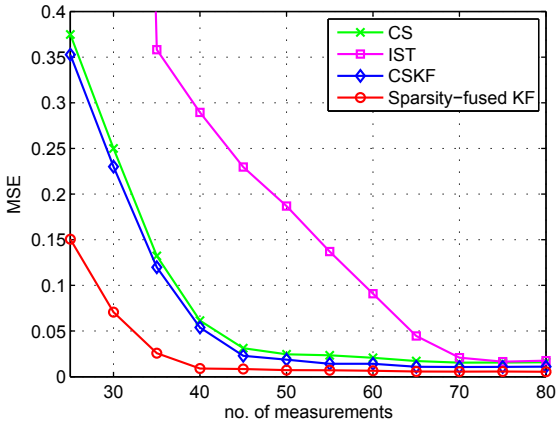

Fig. 5. The MSE for reconstructions with different number of measurements for various algorithms. ($T = 130$, $\sigma_0^2 M = 3$, $N = 256$.)

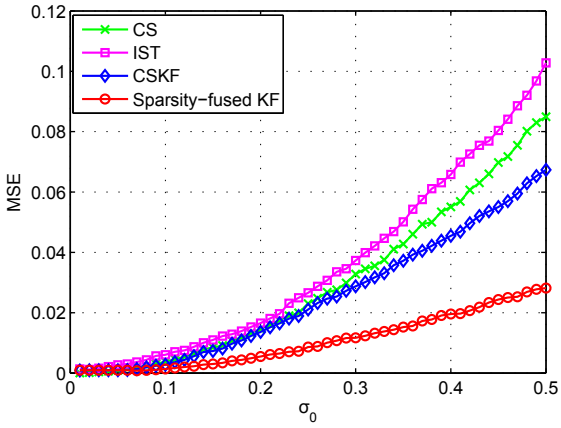

Fig. 6. The MSE for reconstructions under different amount of noise for various algorithms. ($T = 130$, $M = 72$, $N = 256$.)

our reconstructions have an MSE of approximate 0.006; while the mean MSE for the other algorithms are all above 0.014.

Note that the other dynamic algorithm, CSKF, is only slightly better than the static algorithms. This is because the CSKF algorithm is highly dependent on the number and quality of the measurements. In other words, when just a small fraction of measurements are taken and/or they are too noisy, the benefit of the dynamic correlations in the CSKF algorithm is submerged by the error induced by the measurements.

Fig. 4 demonstrates the track of the MSEs at different numbers of iterations for various algorithms. Our algorithm is compared with the iterative algorithms for static reconstruction (IST, AMP [6]) and for dynamic reconstruction (CSKF). Each point in the figure is the average of 130 signals over 100 trials. For clarity, we only show the first 15 iterations. We can see that the dynamic algorithms converge in fewer iterations than the static ones. Our algorithm requires the fewest number of iterations for convergence and its MSE is always the lowest for different number of iterations.

Then we test the performance of our algorithm with different numbers of measurements and under different amounts of measurement noise. The results are shown in Fig. 5 and Fig. 6. All the results are the average of 130 signals over 100 trials.

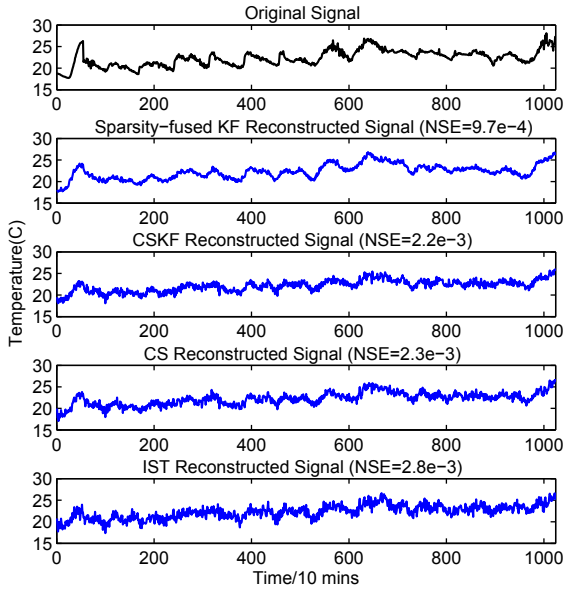

Fig. 7. An example of the reconstructed temperature signals using various algorithms. ($17th\,SN$, $M = 150$, $N = 1024$, $\sigma_0 = 0.2$.)

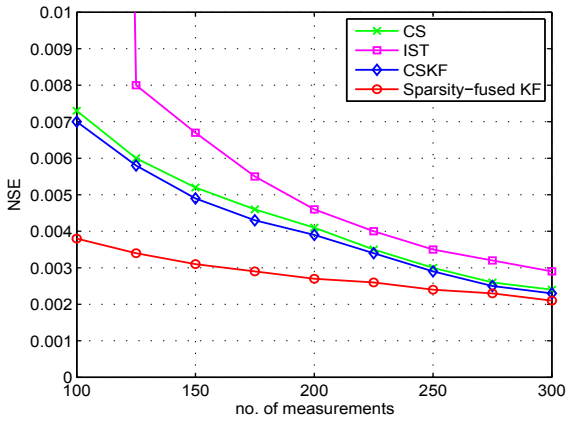

Fig. 8. The NSE for reconstructions with different number of measurements for various algorithms. ($T = 20$, $\sigma_0^2 M = 6$, $N = 1024$.)

In the experiment that yielded the results in Fig. 5, to keep the energy of the measurement noise constant, we set $\sigma_0^2 M = 3$. It is clear that our algorithm is superior to the others, especially when the number of measurement is low. Note that the IST algorithm fails when less than 30 measurements are taken. To obtain Fig. 6, we set $M = 72$ and the rest of the settings are the same as before. We can observe that our algorithm is also more robust than the others when the noise increases.

B. Experiments on Practical Signals

We now employ the proposed algorithm to reconstruct the signals gathered by the WSN located in the Intel Berkeley Research lab [22]. In this WSN, 54 Sensor Nodes (SNs) were deployed at various positions in one floor of the lab building, and each of them has gathered humidity, temperature and light level data for more than one month.

We extract the temperature data during a continuous time period collected by 20 SNs as the dynamic sparse signals, i.e., $T = 20$. As demonstrated in [9], these independently collected

signals are highly correlated. Meanwhile, these signals are highly sparse in the Discrete Cosine Transform (DCT) domain. Therefore, we utilize the DCT matrix as the sparsifying matrix $\mathbf{\Psi}$ and the sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ is set as an i.i.d. Gaussian random matrix. Due to the high correlation presented, we let the transition matrix $\mathbf{F}$ be an identity matrix and use the method in [23] to learn the process noise covariance matrix $\mathbf{Q}$. We use the conventional CS reconstruction result to initialize $\hat{\mathbf{s}}_0$ and $\mathbf{P}_0 = 100\mathbf{Q}$ [11].

Let $M = 150$, $N = 1024$, $\sigma_0 = 0.2$ and the rest of the parameters are set as in section IV-A. Fig.7 demonstrates the reconstructed temperature signals for the 17th SN using different algorithms. It is observable that our algorithm produces better reconstruction results in term of the Normalized Squared Error (NSE), which is defined as $||\hat{\mathbf{s}}_t - \mathbf{s}_t||_2^2 / ||\mathbf{s}_2||_2^2$.

We then vary the number of the measurements to see the performance of our algorithm, as illustrated in Fig. 8. For each particular number of measurements, we carried out 100 trials and took the average of them to calculate the NSE over all 20 SNs. For clarity, the NSE when $M < 100$ is not shown. But we note that when $M = 50$, both the CS and CSKF algorithms fail, i.e., the NSE values are in excess of 0.01; while that of the Sparsity-fused KF is 0.0073. We can see that our algorithm preserves its reconstruction performance advantage with fewer measurements when evaluated with practical data. With fewer measurements needed, our method can reduce the energy consumed in both sampling and transmission, which is one of the most significant concerns in a practical WSN.

V. Conclusions

In this article, we have derived a novel KF procedure, which has fused the sparsity distribution model into its underlying operational principle. Based on this, a Sparsity-fused KF algorithm for reconstructing dynamic sparse signals from a series of noisy CS measurements has been proposed. In the algorithm, we employ the IST method to refine the sparsity model iteratively, which is then utilized in the sparsity-fused KF step. Our algorithm has solved the problem of the conventional KF, i.e., it does not consider the sparsity of signals and cannot be applied to sparse signal recovery. Experimental results using synthetic signals have shown that our algorithm outperforms the static reconstruction algorithms (CS, IST) and the dynamic algorithm (CSKF) when noise exists. It is also shown that compared to these algorithms, our algorithm converges in fewer iterations, has better performance when the number of measurements is lower and it is more robust when the noise level increases. Using practical temperature data gathered by the WSN in the Berkeley Research lab, we have demonstrated that our algorithm can produce visibly better reconstruction than the aforementioned algorithms and it can help to reduce the energy consumption in a practical WSN.

References

[1] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489 – 509, feb. 2006.

[2] E. Candes, E.s and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21 –30, march 2008.

[3] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 –1306, april 2006.

[4] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655 –4666, dec. 2007.

[5] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1063520308000638

[6] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009. [Online]. Available: http://www.pnas.org/content/106/45/18914.abstract

[7] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4053 –4085, sept. 2011.

[8] 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, pp. 2313–2351, 2007.

[9] W. Chen, M. Rodrigues, and I. Wassell, "A frechet mean approach for compressive sensing date acquisition and reconstruction in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 10, pp. 3598–3606, October 2012.

[10] T. Do, Y. Chen, D. Nguyen, N. Nguyen, L. Gan, and T. Tran, "Distributed compressed video sensing," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009, pp. 1393–1396.

[11] N. Vaswani, "Kalman filtered compressed sensing," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 893–896.

[12] W. Dai, D. Sejdinovic, and O. Milenkovic, "Gaussian dynamic compressive sensing," *presented at the Int. Conf. Sampling Theory Appl.(SampTA)*, May. 2011.

[13] E. Karseras, K. Leung, and W. Dai, "Tracking dynamic sparse signals using hierarchical bayesian kalman filters," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, May 2013, pp. 6546–6550.

[14] A. Charles, M. Asif, J. Romberg, and C. Rozell, "Sparsity penalties in dynamical system estimation," in *Information Sciences and Systems (CISS), 2011 45th Annual Conference on*, March 2011, pp. 1–6.

[15] A. Carmi, P. Gurfil, and D. Kanevsky, "Methods for sparse signal recovery using kalman filtering with embedded pseudo-measurement norms and quasi-norms," *Signal Processing, IEEE Transactions on*, vol. 58, no. 4, pp. 2405–2409, April 2010.

[16] D. Donoho and I. Johnstone, "Minimax risk overl p -balls forl p -error," *Probability Theory and Related Fields*, vol. 99, no. 2, pp. 277–303, 1994. [Online]. Available: http://dx.doi.org/10.1007/BF01199026

[17] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9 - 10, pp. 589 – 592, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1631073X08000964

[18] R. Faragher, "Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]," *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 128–132, 2012.

[19] J. Tropp and S. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.

[20] A. Maleki and D. Donoho, "Optimally tuned iterative reconstruction algorithms for compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 330–341, April 2010.

[21] D. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 1094–1121, Feb 2012.

[22] P. Bodik, W. Hong, C. Guestrin, S. Madden, M. Paskin, and R. Thibaux, "Intel lab data." *Online.*, p. http://db.csail.mit.edu/labdata/labdata.html, 2004.

[23] C. Qiu, W. Lu, and N. Vaswani, "Real-time dynamic mr image reconstruction using kalman filtered compressed sensing," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 393–396.