EUSIPCO 2013 1569746703

HIERARCHICAL BAYESIAN KALMAN FILTERS FOR WIRELESS SENSOR NETWORKS

Evripidis Karseras, Kin Leung, and Wei Dai

Department of Electrical and Electronic Engineering
Imperial College, London, UK
{e.karseras11, kin.leung, wei.dai1}@imperial.ac.uk

ABSTRACT

We are interested in reconstructing signals sensed by wireless sensor networks over large areas with a great number of nodes. Numerous constraints usually exist regarding power consumption, communication bandwidth, computational capacity and data volume. By assuming that the sensed signal is sparse in some transform domain we are able to exploit structures in the signal towards achieving coarser and easier-to-deploy sensing grids. By tracking the statistics of the sensed signal it is possible to achieve better performance. Unfortunately traditional approaches like the Kalman filter are known to fail when the signal is sparse. We propose the employment of a *hierarchical Bayesian model* in the tracking process which succeeds in modelling sparsity. It is then possible to achieve further reduction in the number of active sensors by exploiting the temporal correlation of consecutive samples. The theoretical analysis provided solidifies the proposed approach regarding convergence and also provides the conditions under which all sparse signals are recovered exactly. Simulations for synthetic and real-life scenarios show that the proposed method succeeds in recovering time-varying sparse signals with greater accuracy than traditional approaches.

Index Terms**—** dynamic sparse signals, Hierarchical Bayesian Kalman filter

1. INTRODUCTION

When dealing with wireless sensor networks (WSN) spread over a large area we come across some vital constraints. Power consumption of each node rules the communication rates between nodes, and also the local computational capacity. Moreover, the number of nodes needed to achieve the required performance places topological constraints on the grid on which the nodes will be placed. Equally important is the fact that an increasing number of sensors eventually leads to an enormous volume of data for each time that the network is sampled. Keeping a small number of active sensors is of primary importance. By assuming that the sensed signal is sparse in some transform domain then we are able to exploit the structure of the signal and achieve significantly smaller numbers of active sensors compared to the traditional sampling regime. By employing statistical information from past time instants it is possible to accomplish further reduction of the number of measurements by exploiting the temporal relationship between consecutive samples.

If we focus on one time snapshot of the WSN then the problem of sparse signal reconstruction can be solved by proposed approaches in the field of compressed sensing (CS). By following the Sparse Bayesian Learning (SBL) approach, work in [1] is applied for CS in [2], and for basis selection in [3]. A Bayesian network was introduced which promotes sparsity namely the Relevance Vector Machine (RVM). Sparse solutions are promoted by a hierarchy of distributions. This technique also provides estimates of full distributions. The resulting statistical information can be used to make sparsity-aware predictions. Most of the existing sparse recovery algorithms do not exploit statistical information of the signal. There is also no need for a predetermined level of sparsity as is usually the case with CS algorithms. The RVM framework allows for automatic determination of the active components and to be implemented with practical and computationally efficient algorithms [4, 5]. These are key issues for a WSN data fusion platform.

As we mentioned, in order to fully exploit the attributes of the samples (temporal correlation as well as spatial sparsity) we need to track the sensed signal. For this task traditional approaches can be employed like the Kalman filter. Unfortunately this method is not fit for sparse signals. The Kalman filter requires that the mean value of the predicted states to be equal with the mean value of the actual state while minimising variation. The Bayesian optimal estimator must be retained and extended. By examining the formulation of the filter it becomes evident that the model must be revisited. Approaches like the ones in [6, 7] attempt to enforce sparsity in a conservative way without caring about the statistics either. Something crucial to tracking. They also require specifying a number of sensitive parameters. Something not always easy to implement in real-world applications.

We propose the extension of the state-space model behind

The authors would like to acknowledge the European Commission for funding SmartEN ITN (Grant No. 238726) under the Marie Curie ITN FP7 programme. This work is partially supported by the University Defence Research Centre (UDRC), UK.

the Kalman filter with the hierarchy of distributions drawn from the RVM. This results in a sparsity-favouring model of the dynamic system. Since inference in such a dynamic system will be based on the RVM inference procedure, it will not suffer from external parameters that affect sparsity or application dependent care. This internal modification results in useful statistical information on the sparsity of the signal which can be safely used in making safe predictions. The theoretical analysis provides the necessary backing for this idea by extending previous results. Performance is demonstrated empirically by applying the technique on Ozone data measurements over the entire globe. This dataset is especially interesting because measurements are incomplete.

The next section briefly lists the basic properties of the Kalman filter and shows what are the key properties of SBL. Then follows a discussion on previous work. In Section 3 we outline how this relates to the tracking problem, and the revised prediction and update steps are given. Theoretical results are also presented. Finally we provide simulation results in Section 4 on both synthetic and real-life scenarios.

2. SYSTEM MODEL

The dynamic system model is described by the following equations:

$$\boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \boldsymbol{q}_t, \tag{1}$$

$$\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{x}_t + \boldsymbol{n}_t. \tag{2}$$

Vector $\boldsymbol{x}_t$ denotes the sparse signal to be recovered (or state in the state-space dialect). The spatial samples from the WSN for time instant t are gathered in a single vector $\boldsymbol{y}_t$. Signal $\boldsymbol{x}_t$ innovates according to $\boldsymbol{q}_t$ and measurement noise is modelled by $\boldsymbol{n}_t$. Without loss of generality we also assume that $\boldsymbol{x}_t \in \mathbb{R}^n$ is *s-sparse* and design matrix $\boldsymbol{\Phi}_t \in \mathbb{R}^{m \times n}$.

The classic state-space model in Equations (1) and (2) assumes Gaussian distributions on the model parameters. Generally this does not favour sparse signals. By employing SBL promotes sparse solutions for Equation (2).

2.1. The Kalman filter Bayesian optimal estimator

Kalman filtering is based on the Gaussian assumption so that: $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_{t-1}, \boldsymbol{Q}_t)$ and $p\left(\boldsymbol{y}_t|\boldsymbol{x}_t\right) = \mathcal{N}\left(\boldsymbol{\Phi}\boldsymbol{x}_t, \sigma^2\boldsymbol{I}\right)$ while $p(\boldsymbol{q}_t) = \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{Q}_t\right)$ and $p(\boldsymbol{n}_t) = \mathcal{N}\left(\boldsymbol{0}, \sigma^2\boldsymbol{I}\right)$. The statistics of a dynamic signal are tracked by an iterative two-step procedure: the Kalman filter prediction and update steps. The prediction step calculates the parameters of $p(\boldsymbol{x}_t|\boldsymbol{y}_{t-1})$ while the update step evaluates those of $p(\boldsymbol{x}_t|\boldsymbol{y}_t)$. It can be shown that both distribution functions are Gaussian. As a result, not only one can obtain the globally optimal estimate of $\boldsymbol{x}_t$ in terms of mean squared error (MSE), but also track the full distribution exactly. It can be easily verified that the optimal estimate of the standard Kalman filter is typically not sparse.

There are several approaches in which the Kalman filter framework is externally modified to admit sparse solutions. The essential idea behind [6] and [7] is to apply thresholds that enforce sparsity. Work in [8] adopts a probabilistic model but signal amplitudes and support are estimated separately. Techniques presented in [9] use prior sparsity knowledge into the tracking process. All these approaches typically require a number of parameters to be pre-set. It also remains unclear how these methods perform towards model and parameter mismatch.

2.2. SBL: A hierarchy of distributions

We now focus on a single time instant of Equation (2). Note that time index t is temporarily dropped. SBL was proposed in [2] for compressed sensing. It was underscored in Subsection 2.1 that if $\boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{Q}\right)$ where $\boldsymbol{Q}$ is full rank, then the minimum MSE solution to Equation (2) provided by the Kalman optimal estimator is typically not sparse. In SBL this issue is addressed by introducing individual hyper-parameters α_i to control the variance of each component x_i:

$$p\left(\boldsymbol{x}|\boldsymbol{\alpha}\right) = \prod_{i=1}^{n} \mathcal{N}\left(0, \alpha_i^{-1}\right) = \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{A}^{-1}\right)$$

where matrix $\boldsymbol{A} = \operatorname{diag}\left([\alpha_1, \cdots, \alpha_n]\right)$. Driving $\alpha_i = +\infty$ results in $p\left(x_i|\alpha_i\right) = \mathcal{N}\left(0, 0\right)$ which means that it is *a posteriori* certain that $x_i = 0$. Hence, the reconstruction problem is then changed to finding the maximum likelihood solution of $\boldsymbol{\alpha}$ for the given measurements $\boldsymbol{y}$. The explicit form of the likelihood function $p\left(\boldsymbol{y}|\boldsymbol{\alpha}, \sigma^2\right)$ was derived in [1]. A set of fast algorithms to estimate $\boldsymbol{\alpha}$ and subsequently $\boldsymbol{x}$ are proposed in [5]. Even though the algorithms in [5] are not guaranteed to produce globally optimal solutions, they perform extremely well for many practical scenarios.

We need to point out; there are no control parameters to be manually set, like a predetermined sparsity level. A stopping criterion is needed for the optimisation process and this can be set to some safe threshold so as to not affect the validity of the process. This is of great importance for the proposed method since the sparsity level of $\boldsymbol{x}_t$ is unknown and varying. Furthermore, the unknown state vector $p(\boldsymbol{x}|\boldsymbol{\alpha})$ is a Gaussian and this allows for convenient incorporation of the SBL framework into the state-space model.

3. HIERARCHICAL BAYESIAN KALMAN FILTER

The Kalman filter and SBL are put together. The dynamic system is still described by Equations (1) and (2). Measurement noise is assumed to be Gaussian with known covariance matrix, i.e. $\boldsymbol{n} \sim \mathcal{N}\left(0, \sigma^2\boldsymbol{I}\right)$. Differently from the standard Kalman filter where the covariance matrix $\boldsymbol{Q}$ of $\boldsymbol{q}_t$ is given, we assume that $\boldsymbol{q}_t \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{A}_t^{-1}\right)$ where $\boldsymbol{A}_t = \operatorname{diag}\left(\boldsymbol{\alpha}_t\right) = \operatorname{diag}\left([\alpha_1, \cdots, \alpha_n]_t\right)$ and hyper-parameters α_i are not priorly known and will have to be learned from the observation vector $\boldsymbol{y}_t$.

Similar to the standard Kalman filter the two steps of prediction and update, still need to be performed at each time instant. In the prediction step, one has:

$$\begin{aligned}&\boldsymbol{\mu}_{t|t-1} = \boldsymbol{\mu}_{t-1},\ \boldsymbol{\Sigma}_{t|t-1} = \boldsymbol{\Sigma}_{t-1|t-1} + \boldsymbol{A}_t^{-1},\\ &\boldsymbol{y}_{t|t-1} = \boldsymbol{\Phi}_t\boldsymbol{\mu}_{t|t-1},\ \boldsymbol{y}_{e,t} = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}.\end{aligned} \tag{3}$$

where the notation $t|t-1$ means prediction at time instant t for measurements up to time instant $t-1$. In the update step, one computes:

$$\begin{gathered}\boldsymbol{\mu}_{t|t} = \boldsymbol{\mu}_{t|t-1} + \boldsymbol{K}_t\boldsymbol{y}_{e,t}, \boldsymbol{\Sigma}_{t|t} = (\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{\Phi}_t)\boldsymbol{\Sigma}_{t|t-1}\\ \boldsymbol{K}_t = \boldsymbol{\Sigma}_{t|t-1}\boldsymbol{\Phi}_t^T(\sigma^2\boldsymbol{I} + \boldsymbol{\Phi}_t\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{\Phi}_t^T)^{-1}\end{gathered}$$

In addition to these steps is the task of learning hyper-parameters $\boldsymbol{\alpha}_t$. From Equation (3), $\boldsymbol{y}_{e,t} = \boldsymbol{\Phi}_t\boldsymbol{q}_t + \boldsymbol{n}_t$ where a sparse $\boldsymbol{q}_t$ is preferred to produce a sparse $\boldsymbol{x}_t$. As per the analysis in [1, 5], maximising the likelihood $p(\boldsymbol{y}_t|\boldsymbol{\alpha}_t)$ is equivalent to minimising the following cost function:

$$\mathcal{L}(\boldsymbol{\alpha}_t) = \log|\boldsymbol{\Sigma}_\alpha| + \boldsymbol{y}_{e,t}^T\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{-1}\boldsymbol{y}_{e,t}, \tag{4}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = \sigma^2\boldsymbol{I} + \boldsymbol{\Phi}_t\boldsymbol{A}_t^{-1}\boldsymbol{\Phi}_t^T$. The algorithms described in [5] can be applied to estimate $\boldsymbol{\alpha}_t$. Note that the cost function $\mathcal{L}(\boldsymbol{\alpha})$ is not convex. The obtained estimate $\boldsymbol{\alpha}_t$ is generally sub-optimal (details on the estimation of the globally optimal $\boldsymbol{\alpha}_t$ are given in Section 3.1.) Nevertheless, this sub-optimal solution is proved to be very useful in practice.

The HB-Kalman demonstrates several advantages over previously proposed methods. Probably the most important being the fact that one is able to have statistical information on the *sparsity* of the signal. This retains the original nature of the Kalman filter which follows the statistics of a signal but acts on the sparsity of the signal. This is done through the employment of hyper-parameters to model the innovations of the signal. The HB-Kalman inherits the practicality of SBL in determining the level of sparsity and is free from tedious parameters like previous approaches. One only needs prior knowledge of σ^2 since in many real-world applications such as WSNs the noise floor of the sensors is labelled. Also, in Section 3.1 we demonstrate that this model allows us to draw certain conclusions on global optimality.

3.1. Performance Guarantees

Again for convenience we drop subscript t and focus on Equation (2) where $\boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{A}^{-1}\right)$. In [3] SBL was analysed for basis selection. It had been proven that a maximally sparse solution of $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$ is attained at the global minimum of the cost function. However, the analysis did not specify the conditions to avoid local minima. By contrast, we provide a more refined analysis and derive the conditions under which the original inference algorithm in [5] converges to the global minimum. Due to space constraints, only the main results are presented and all detailed proofs are delayed to the journal version of this paper.

In the performance analysis, we follow [3] by driving noise variance $\sigma^2 \to 0$, which corresponds to the noiseless setting. The following Theorem specifies the behaviour of cost function $\mathcal{L}(\boldsymbol{\alpha})$.

Theorem 1. *For any given $\boldsymbol{\alpha}$, define the set $\mathcal{I} \triangleq \{1 \le i \le n:\ 0 < \alpha_i < \infty\}$. Then it holds that:*

$$\lim_{\sigma^2\to 0} \sigma^2\mathcal{L}(\boldsymbol{\alpha}) = \left\|\boldsymbol{y} - \boldsymbol{\Phi}_{\mathcal{I}}\boldsymbol{\Phi}_{\mathcal{I}}^\dagger\boldsymbol{y}\right\|_2^2, \tag{5}$$

where $\boldsymbol{\Phi}_{\mathcal{I}}$ is a sub-matrix of $\boldsymbol{\Phi}$ formed by the columns indexed by $\mathcal{I}$, and $\boldsymbol{\Phi}_{\mathcal{I}}^\dagger$ denotes the Moore-Penrose pseudo-inverse of $\boldsymbol{\Phi}$. Furthermore, if $|\mathcal{I}| < m$ and $\boldsymbol{y} \in \operatorname{span}(\boldsymbol{\Phi}_{\mathcal{I}})$, then $\mathcal{L}(\boldsymbol{\alpha}) \to -\infty$ and $\sigma^2\mathcal{L}(\boldsymbol{\alpha}) \to 0$ as $\sigma^2 \to 0$.

One important conclusion is that the cases presented in [3] are limiting cases of Theorem 1 for when $\mathcal{L}(\boldsymbol{\alpha}) \to -\infty$. Also and as can be seen, a proper scaling of the cost function gives the squared ℓ_2-norm of the reconstruction error. This is equivalent to recovering a support set that minimises the reconstruction distortion which is the exact same principle as in many greedy CS algorithms.

This theoretical analysis allows us to make the necessary connections between SBL and already well established sparse recovery algorithms. Actually it can be proven that the inference algorithm in [5] is closely related to the OMP algorithm [10]. Further details are delayed to the journal version of this paper. This firstly means that a global performance guarantee can be drawn. This is stated in Theorem 2.

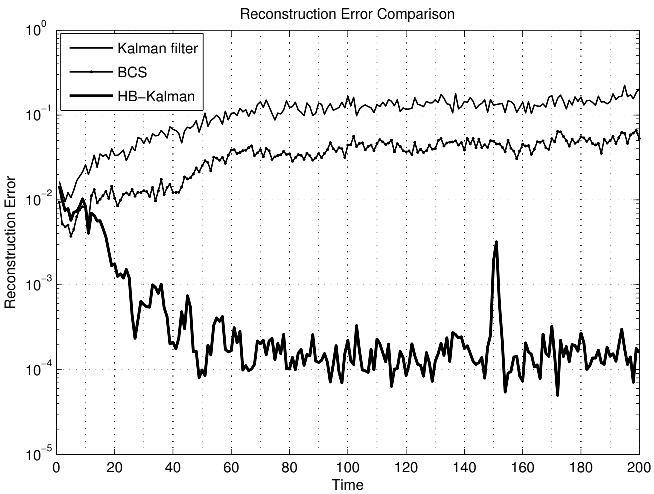
Theorem 2. *Assume the noiseless setting $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$ where the columns of $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ are normalised, i.e. $\phi_i^T\phi_i = 1$ for all $1 \le i \le n$. Suppose that the matrix $\boldsymbol{\Phi}$ satisfies the incoherence condition that $\left|\phi_i^T\phi_j\right| < 0.375/s$ for all $1 \le i \ne j \le n$. Then the sequential algorithm to solve the RVM inference problem [5] reconstructs all s-sparse signals exactly.*

Cost function in (5) suggests that it is possible to design inference techniques based on more advanced CS reconstruction algorithms like the SP. The detailed algorithm design and the corresponding performance guarantees are discussed in the journal version of this paper. To summarise; it is possible to track the temporal evolution of a sparse signal with an efficient algorithm. This is possible by adopting a proper model for the dynamic system. Our theoretical analysis shows that further improvements can be achieved by studying the connections between this algorithm and other CS methods.

4. EMPIRICAL RESULTS

4.1. Synthetic Data

We compare the original Kalman filter and the proposed method, namely the Hierarchical Bayesian Kalman filter



(a) Number of active sensors $m = 100$

(b) Number of active sensors $m = 28$.

Fig. 1. Tracking performance comparison.

(HB-Kalman). We also compare against the case where i.i.d samples are assumed, i.e apply Bayesian compressed sensing (BCS) [2] separately at each time instant.

Signal $\boldsymbol{x}_t \in \mathbb{R}^n$ is assumed to be sparse in its natural basis with support set $\mathcal{S}$ chosen uniformly at random from $[1, n]$ where $n = 256$. The non-zero entries of $\boldsymbol{x}_t$ evolve according to Equation 1 with $Q_i = \sigma_q^2 I$ with $\sigma_q^2 = 0.1$. The simulation time for this experiment was $T = 200$ time instants. At two randomly chosen time instants; $t = 50$ and $t = 150$, a change in the support of $\boldsymbol{x}_t$ is introduced. A non-zero component is added to the support of $\boldsymbol{x}_{50}$ and a non-zero component is removed from the support of $\boldsymbol{x}_{150}$. Apart from these two time instants the support of $\boldsymbol{x}_t$ remains stationary. At $t = 1$ the support is initialised with $K = 30$ non-zero components. Measurement noise variance is set to $\sigma_n^2 = 0.01$ for the entire simulation time.

Initially, sensors take noisy measurements $\boldsymbol{y}_t$ by choosing the entries of matrix $\boldsymbol{\Phi}_t \in \mathbb{R}^{128\times256}$ to be drawn from $\mathcal{N}(0, \frac{1}{m})$ and to be re-sampled at each time instant. The number of active sensors at each time instant remains constant. In Figure 1(a) the reconstruction error is plotted against time for each of the three reconstruction methods. It is evident that the error levels are much lower for the HB-Kalman filter when compared to the conventional Kalman filter, direct consequence of the assumed sparse model. By comparing to the repeated application of the BCS method, i.e assuming i.i.d samples, we see that incorporating statistical information from previous estimates results in lower reconstruction error.

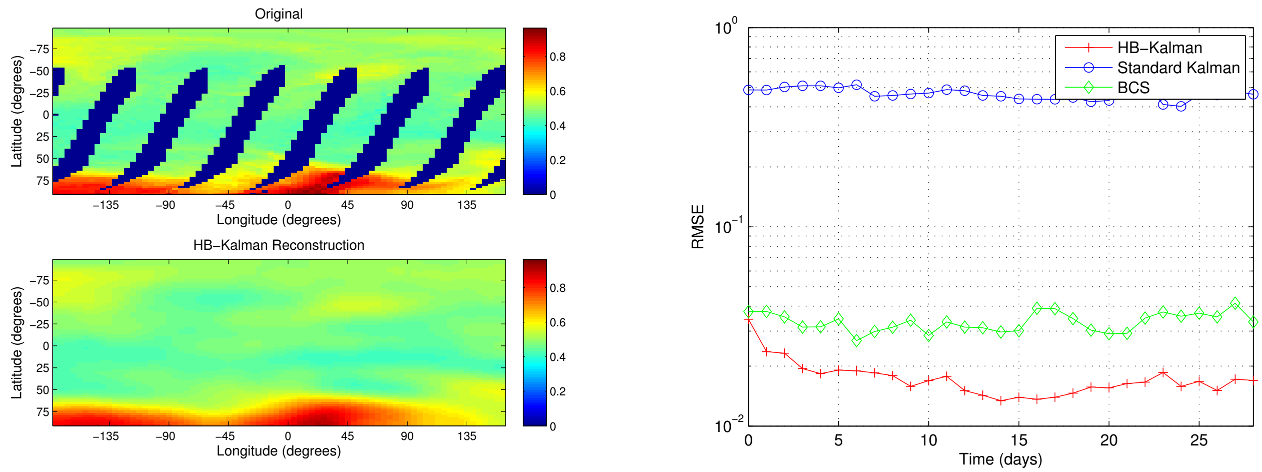
In a more difficult setting we assume that the number of active sensors in the network is reduced but the sparse signal to be sensed continues to evolve with sparsity levels well above this number. More specifically the number of active sensors has been reduced to $m = 28 < 2 \times 30$ which is *less* than twice the number of active components which is 30. This is a particularly difficult case since the number of measurements is less than what is required for exact reconstruction of the sparse signal. We assume that the sparse signal x_0 is known beforehand. It is shown in Figure 1(b) that given statistical information from an earlier time instant the filter manages to retain its performance even-though it might take slightly longer to converge.

4.2. Ozone distribution signal reconstruction

Here we test the proposed method on a real-life scenario. We attempt to track the spatial distribution of Ozone over the globe. The dataset on which we test the proposed method is obtained from the Ozone Monitoring Instrument (OMI) on the NASA Aura spacecraft [12]. This dataset is of particular interest since the measurements for each day of the month are incomplete due to the satellite's trajectory. This can be seen in Figure 2(a) by the blue vertical stripes. Furthermore we assume that the data is sparse in the discrete cosine transform (DCT) domain.

The original images are cropped to a square and undersampled by a factor of 8 so that $m = 4275$ in order to deal with the high dimensionality of the data so as to be able to perform the tests. This however does not hinder the performance and does not affect the generality of the results. Each time instant corresponds to one day of the month for a total of $T = 28$ days. Measurement noise variance is set to $\sigma^2 = 0.1^6$.

As can be seen from Figure 2(b), the standard Kalman filter fails to accurately track the dynamic sparse signal. By contrast, repeated application of the BCS method and the HB-Kalman filter exhibit much lower error levels and accurately reconstruct the missing data. The HB-Kalman outperforms the BCS method as it incorporates statistical information from



(a) Reconstructed signal for one time instant using HB-Kalman.

(b) Reconstruction error.

Fig. 2. (a) Atmospheric ozone distribution measured in (normalised) Dobson units. Original data shown on top. (b) Reconstruction error for the period of one month.

past estimates resulting in lower reconstruction error.

5. CONCLUSIONS

We proposed a method for tracking dynamic sparse signals and reducing the number of active sensors in a WSN by exploiting the spatial and temporal characteristics of the sensed signal; a problem which can not be solved with traditional approaches like the Kalman filter. Theoretical analysis shows that it is possible to improve upon existing techniques and incorporate them in the proposed framework. The robustness of the technique is tested against synthetic and real-life scenarios and is empirically shown to achieve better performance than traditional methods.

6. REFERENCES

[1] Michael E. Tipping, "The relevance vector machine," 2000.

[2] Shihao Ji, Ya Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346 –2356, june 2008.

[3] D.P. Wipf and B.D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153 – 2164, aug. 2004.

[4] Anita C. Faul and Michael E. Tipping, "Analysis of sparse Bayesian learning," in *Advances in Neural Information Processing Systems 14*. 2001, pp. 383–389, MIT Press.

[5] M.E. Tipping, A.C. Faul, et al., "Fast marginal likelihood maximisation for sparse Bayesian models," in *International workshop on artificial intelligence and statistics*. Jan, 2003, vol. 1.

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

[7] A. Carmi, P. Gurfil, and D. Kanevsky, "Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2405 –2409, april 2010.

[8] J. Ziniel, L.C. Potter, and P. Schniter, "Tracking and smoothing of time-varying sparse signals via approximate belief propagation," in *Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2010, pp. 808–812.

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

[10] J.A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231 – 2242, oct. 2004.

[11] Wei Dai and Olgica Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, pp. 2230–2249, 2009.

[12] NASA, "Space-based Measurements of Ozone and Air Quality," 2012, [Online] http://ozoneaq.gsfc.nasa.gov/.