GAUSSIAN DYNAMIC COMPRESSIVE SENSING

Wei Dai[1]*, Dino Sejdinovic*[2]*, and Olgica Milenkovic*[1]

[1]University of Illinois, Urbana-Champaign, ECE Department
[2] University of Bristol, Department of Mathematics
Emails: {weidai07, milenkov}@uiuc.edu, d.sejdinovic@bristol.ac.uk

ABSTRACT

We consider the problem of estimating a discrete-time sequence of sparse signals with Gaussian innovations. Instances of such problems arise in networking and imaging, in particular, dynamic and interventional MRI imaging. Our approach combines Kalman filtering and compressive sensing (CS) techniques by introducing a sparse MAP estimator for Gaussian signals, and then developing a CS-type algorithm for solving the sparse MAP problem. Despite the underlying assumption that the sequence of sparse signals is Gaussian, our approach also allows for efficient tracking of sparse non-Gaussian signals obtained via non-linear mappings, using only one sample/observation per time instance.

***Keywords*—** Compressive sensing, Kalman filtering, sparse Gaussian signals.

1. INTRODUCTION

In many applications, the signals to be processed can be well approximated by sparse signals. Compressive sensing (CS) is a novel framework for efficient sampling of sparse and approximately sparse signals with provable performance guarantees and polynomial time reconstruction complexity [5, 2].

We consider a special CS scenario where one is concerned with reconstructing a discrete-time sequence of sparse Gaussian signals with Gaussian innovations. Note that Kalman filtering is the classic technique to handle *Gaussian innovation signals*. While both CS and Kalman filtering techniques have been extensively and successfully applied to these problems, a coherent combination of these two techniques in the context of dynamic CS is currently unknown.

We propose a reconstruction paradigm for sparse Gaussian signals using techniques from both Kalman filtering and CS. First, we describe a sparse MAP estimator that is based on the statistics of the signal. Second, we apply CS techniques for solving this sparse MAP estimation problem. Our approach allows for devising 1) computationally efficient algorithms to track sparse Gaussian signals and 2) signal tracking using only *one*, or a small (constant number) of samples per time instance. Furthermore, our analysis presented in Section 2.4 reveals that 1) the formulation of the standard CS problem represents a special instance of the proposed sparse MAP estimator problem. 2) several algorithms for standard CS can be viewed as special instances of algorithms for sparse Gaussian signal reconstruction [3]. The Gaussian Kalman filtering/CS method shows remarkable simulation performance for the case that the sparse signals to be tracked are non-Gaussian, and hence can be applied to a much wider class of sparse signals.

The proposed methods differ significantly from other techniques described in the literature regarding dynamically changing sparse signals. In [6], the authors propose a class of locally competitive algorithms for reconstruction of time varying sparse signals with the goal of simplifying practical implementations of greedy CS algorithms. In [1], the authors propose to develop warm-start algorithms for time varying signals. More recently, the authors in [7, 8] described an approach for causal and recursive reconstruction of a time sequence of sparse signals with slowly changing sparsity patterns. There, one needs to solve a standard CS reconstruction problem at each time instance, and therefore the number of measurements per time instance needs to be sufficiently large to admit sampling RIP conditions. In contrast, the sparse MAP estimator assumes a Gaussian signal statistics and provides reasonably good performance even when only one sample is used per time instance.

2. RECONSTRUCTION OF SPARSE GAUSSIAN SIGNALS

2.1. Problem Formulation

In standard Kalman filtering, one considers a linear dynamic system described by two sets of update matrices, $\boldsymbol{\Phi}_t$, $\boldsymbol{\Psi}_t$ and update equations

$$\boldsymbol{x}_t = \boldsymbol{\Psi}_t \boldsymbol{x}_{t-1} + \boldsymbol{u}_t,$$
$$\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{x}_t + \boldsymbol{v}_t.$$

Here, $\boldsymbol{x}_t \in \mathbb{R}^n$ represents the state vector of the system, $\boldsymbol{y}_t \in \mathbb{R}^m$ denotes the measurement vector, $\boldsymbol{u}_t \in \mathbb{R}^n$ and $\boldsymbol{v}_t \in \mathbb{R}^m$ are Gaussian innovation vectors with $\boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$ and $\boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_v)$, respectively. The subscript $t = 1, 2, \cdots$ describes the time instances at which the signal is observed. Suppose that the statistics of $\boldsymbol{x}_{t-1}$ are known, and given by $\boldsymbol{x}_{t-1} \sim \mathcal{N}(\hat{\boldsymbol{x}}_{t-1}, \boldsymbol{\Sigma}_{t-1})$. Given $\boldsymbol{y}_t$, $\boldsymbol{\Psi}_t$ and $\boldsymbol{\Phi}_t$, the Kalman filter estimates $\boldsymbol{x}_t$, denoted by $\hat{\boldsymbol{x}}_t$, such that the mean square error (MSE)

This work was supported by the NSF grants CCF 0821910 and CCF STCs-2010 Class.

$\mathrm{E}\left[\|\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t\|_2^2\right]$ is minimized. It also holds that the estimate $\hat{\boldsymbol{x}}_t$ represents the maximum a posteriori probability (MAP) estimate, i.e.,

$$\hat{\boldsymbol{x}}_t = \arg\max_{\boldsymbol{x}} p_{\boldsymbol{X}_t|\boldsymbol{Y}_t,\boldsymbol{X}_{t-1}}\left(\boldsymbol{x}|\boldsymbol{y}_t,\hat{\boldsymbol{x}}_{t-1}\right), \tag{1}$$

In general, the estimate $\hat{\boldsymbol{x}}_t$ does not have a sparse structure.

Now suppose that one has prior information that $\boldsymbol{x}_t$ is K-sparse. The MAP estimator that takes the sparsity assumption into consideration is given by

$$\hat{\boldsymbol{x}}_t = \arg\max_{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0 \le K} p_{\boldsymbol{X}_t|\boldsymbol{Y}_t,\boldsymbol{X}_{t-1}}\left(\boldsymbol{x}|\boldsymbol{y}_t,\hat{\boldsymbol{x}}_{t-1}\right), \tag{2}$$

where the pseudo-norm $\|\cdot\|_0$ counts the number of non-zero entries of its argument. It is clear that

$$p_{\boldsymbol{X}_t|\boldsymbol{Y}_t,\boldsymbol{X}_{t-1}}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}_{t-1},\boldsymbol{y}_t\right) \propto p_{\boldsymbol{X}_t|\boldsymbol{X}_{t-1}}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}_{t-1}\right) p_{\boldsymbol{Y}_t|\boldsymbol{X}_t}\left(\boldsymbol{y}_t|\boldsymbol{x}_t\right),$$

$p_{\boldsymbol{X}_t|\boldsymbol{X}_{t-1}}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}_{t-1}\right) = p_{\boldsymbol{U}_t}\left(\boldsymbol{x} - \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$ and $p_{\boldsymbol{Y}_t|\boldsymbol{X}_t}\left(\boldsymbol{y}_t|\boldsymbol{x}\right) = p_{\boldsymbol{V}_t}\left(\boldsymbol{y}_t - \boldsymbol{\Phi}_t\boldsymbol{x}\right)$, where $p_{\boldsymbol{U}_t}$ and $p_{\boldsymbol{V}_t}$ denote the probability density functions (PDF) of $\boldsymbol{U}_t$ and $\boldsymbol{V}_t$, respectively. The sparse MAP estimator in (2) can be written as

$$\hat{\boldsymbol{x}}_t = \arg\min_{\|\boldsymbol{x}\|_0 \le K} -\log p_{\boldsymbol{U}_t}\left(\boldsymbol{x} - \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right) - \log p_{\boldsymbol{V}_t}\left(\boldsymbol{y}_t - \boldsymbol{\Phi}_t\boldsymbol{x}\right). \tag{3}$$

Since $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$ are Gaussian distributed, there exist explicit expressions for $p_{\boldsymbol{U}_t}\left(\boldsymbol{x} - \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$ and $p_{\boldsymbol{V}_t}\left(\boldsymbol{y}_t - \boldsymbol{\Phi}_t\boldsymbol{x}\right)$. Here and henceforth, assume that both covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_v$ have full rank, i.e., that both $\boldsymbol{\Sigma}_u^{-1}$ and $\boldsymbol{\Sigma}_v^{-1}$ exist. Let $\boldsymbol{A}_t = 2\left(\boldsymbol{\Sigma}_u^{-1} + \boldsymbol{\Phi}_t^T\boldsymbol{\Sigma}_v^{-1}\boldsymbol{\Phi}_t\right)$, $\boldsymbol{b}_t = -2\left(\boldsymbol{\Sigma}_u^{-1}\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{\Phi}_t^T\boldsymbol{\Sigma}_v^{-1}\boldsymbol{y}_t\right)$, and

$$f_t\left(\boldsymbol{x}\right) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}_t\boldsymbol{x} + \boldsymbol{b}_t^T\boldsymbol{x}. \tag{4}$$

It can be verified that the sparse MAP estimator in (3) is equivalent to

$$\hat{\boldsymbol{x}}_t = \arg\min_{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0 \le K} f_t\left(\boldsymbol{x}\right). \tag{5}$$

This is the optimization problem that one needs to solve for each time instance t.

Note that it is NP-hard to solve the optimization problem (5). We therefore propose two practical methods for solving this optimization problem approximately in subsections 2.2 and 2.3. Due to space limitations, the corresponding performance analysis is postponed to the companion paper [4].

2.2. Convex Relaxation

One way to solve (5) is via convex relaxation. More precisely, one can solve an unconstrained optimization problem given by

$$\hat{\boldsymbol{x}}_t = \arg\min_{\boldsymbol{x}} f_t\left(\boldsymbol{x}\right) + \mu\left\|\boldsymbol{x}\right\|_1, \tag{6}$$

where $\|\boldsymbol{x}\|_1 = \sum|\boldsymbol{x}_i|$ denotes the ℓ_1-norm. We refer to this method as *ℓ_1-MAP method*. One issue related to the implementation of this method is how to choose μ, since the resulting estimate $\hat{\boldsymbol{x}}_t$ strongly depends on the parameter μ. One practical solution to this issue involves offline training but may not be suitable for online computing.

2.3. Greedy Algorithms

Another way to solve (5) is to perform a greedy search for a K-sparse signal. There are several greedy algorithms developed for standard CS reconstruction, including OMP, CoSaMP and SP. While these greedy algorithms exhibit low complexity and provable performance guarantees, they cannot be directly applied to reconstruction of dynamically changing sparse Gaussian signals as they do not take temporal correlation of signals into consideration. Motivated by the SP algorithm for greedy CS reconstruction [3], we propose *an SP algorithm for sparse MAP estimation (SP-MAP)*. The steps of this algorithm are described in Algorithm 1.

Algorithm 1 The SP-MAP Algorithm

Let $\ell_{\max}$ be the maximum iterations at each time instance. Let $\hat{\boldsymbol{x}}_0 = \boldsymbol{0}$. At time instance t, perform the following operations.
Initialization:

1. Define $\boldsymbol{x}_t' = \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}$, $\boldsymbol{A} = 2\left(\boldsymbol{\Sigma}_u^{-1} + \boldsymbol{\Phi}_t^T\boldsymbol{\Sigma}_v^{-1}\boldsymbol{\Phi}_t\right)$ and $\boldsymbol{b} = -2\left(\boldsymbol{\Sigma}_u^{-1}\boldsymbol{x}_t' + \boldsymbol{\Phi}_t^T\boldsymbol{\Sigma}_v^{-1}\boldsymbol{y}_t\right)$.
2. Let $\ell = 0$. Let $\hat{\boldsymbol{x}}_t = -\boldsymbol{A}^{-1}\boldsymbol{b}$. Let $\mathcal{K}$ be the set of the K indices corresponding to the largest $\boldsymbol{A}_{i,i}\left|\hat{\boldsymbol{x}}_{t,i}\right|^2$'s, $i \in [n]$. Define $\hat{\boldsymbol{x}}_t^{(\ell)}$ such that $\hat{\boldsymbol{x}}_{t,\mathcal{K}^c}^{(\ell)} = \boldsymbol{0}$ and $\hat{\boldsymbol{x}}_{t,\mathcal{K}}^{(\ell)} = -\boldsymbol{A}_{\mathcal{K},\mathcal{K}}^{-1}\boldsymbol{b}_{\mathcal{K}}$.
3. Let $\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_t^{(\ell)}$. Compute $f^{(\ell)} = \frac{1}{2}\hat{\boldsymbol{x}}_t^T\boldsymbol{A}\hat{\boldsymbol{x}}_t + \boldsymbol{b}^T\hat{\boldsymbol{x}}_t$.

Iterations:

1. Let $\ell = \ell + 1$. For every $i \notin \mathcal{K}$, compute $\Delta_i = \left(\langle\hat{\boldsymbol{x}}_{t,\mathcal{K}}, \boldsymbol{A}_{\mathcal{K},i}\rangle + \boldsymbol{b}_i\right)^2 / \boldsymbol{A}_{i,i}$. Let $\mathcal{K}_\Delta$ be the set of the K indices corresponding to the largest Δ_i's, $i \in \mathcal{K}^c$.
2. Let $\tilde{\mathcal{K}} = \mathcal{K}\bigcup\mathcal{K}_\Delta$. Define $\tilde{\boldsymbol{x}}_t$ such that $\tilde{\boldsymbol{x}}_{t,\tilde{\mathcal{K}}^c} = \boldsymbol{0}$ and $\tilde{\boldsymbol{x}}_{t,\tilde{\mathcal{K}}} = -\boldsymbol{A}_{\tilde{\mathcal{K}},\tilde{\mathcal{K}}}^{-1}\boldsymbol{b}_{\tilde{\mathcal{K}}}$. For every $i \in \tilde{\mathcal{K}}$, compute $\Delta_i = \boldsymbol{A}_{i,i}\tilde{\boldsymbol{x}}_{t,i}^2$.
3. Let $\mathcal{K}$ be the set of the K indices corresponding to the largest Δ_i's, $i \in \tilde{\mathcal{K}}$. Define $\hat{\boldsymbol{x}}_t^{(\ell)}$ such that $\hat{\boldsymbol{x}}_{t,\mathcal{K}^c}^{(\ell)} = \boldsymbol{0}$ and $\hat{\boldsymbol{x}}_{t,\mathcal{K}}^{(\ell)} = -\boldsymbol{A}_{\mathcal{K},\mathcal{K}}^{-1}\boldsymbol{b}_{\mathcal{K}}$. Compute $f^{(\ell)} = \frac{1}{2}\hat{\boldsymbol{x}}_t^T\boldsymbol{A}\hat{\boldsymbol{x}}_t + \boldsymbol{b}^T\hat{\boldsymbol{x}}_t$.
4. If $f^{(\ell)} > f^{(\ell-1)}$, quit the iterations.
5. Let $\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_t^{(\ell)}$. If $\ell \ge \ell_{\max}$, quit the iterations. Otherwise, go to Step 1 for the next iteration.

2.4. Connections to Standard CS Techniques

In standard CS, the nonzero components of the sparse signal are arbitrary. Indeed, this model can be viewed as a special case of the sparse Gaussian signal model. Consider the sparse Gaussian signal model where $\boldsymbol{\Sigma}_u = \sigma_u^2\boldsymbol{I}$, $\boldsymbol{\Sigma}_v^2 = \sigma_v^2\boldsymbol{I}$, and $\sigma_u^2 \to \infty$. For any given $\hat{\boldsymbol{x}}_{t-1} \in \mathbb{R}^n$ and $\boldsymbol{\Psi}_t \in \mathbb{R}^{n\times n}$, the probabilities $p_{\boldsymbol{U}_t}\left(\boldsymbol{x}_1 - \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$ and $p_{\boldsymbol{U}_t}\left(\boldsymbol{x}_2 - \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$ are asymptotically identical for any given $\boldsymbol{x}_1 \ne \boldsymbol{x}_2 \in \mathbb{R}^n$. As a

result, one can drop the term $\log p_{\boldsymbol{U}_t}\left(\boldsymbol{x}-\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$ in (3) and the sparse MAP estimator (3) becomes

$$\hat{\boldsymbol{x}}_t = \underset{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0\leq K}{\arg\min}\ \|\boldsymbol{y}_t-\boldsymbol{\Phi}_t\boldsymbol{x}\|_2^2\,,$$

which is the optimization problem arising in standard CS.

The ℓ_1-norm regularization and the SP methods for standard CS may also be seen as special instances of the proposed ℓ_1-MAP and SP-MAP methods, respectively. Suppose that $\boldsymbol{\Sigma}_u = \sigma_u^2\boldsymbol{I}$, $\boldsymbol{\Sigma}_v = \boldsymbol{I}$ and $\sigma_u^2\to\infty$. It is straightforward to verify that the optimization problem in (6) becomes

$$\hat{\boldsymbol{x}}_t = \underset{\boldsymbol{x}}{\arg\min}\ \|\boldsymbol{y}_t-\boldsymbol{\Phi}_t\boldsymbol{x}\|_2^2+\mu\,\|\boldsymbol{x}\|_1\,.$$

The connections between standard SP algorithm and the proposed SP-MAP algorithm will be elucidated in our companion paper [4].

Despite this apparent similarity between the sparse MAP approach and standard CS, the fundamental difference between the approaches amount to the sparse Gaussian signal model. In standard CS, when the number of measurements is very small, the solution of the corresponding convex relaxation may not be unique, that is, the intersection of the corresponding ℓ_1-ball and the null space of the measurement matrix may contain multiple points. In our approach, Gaussian statistics play an important role in the reconstruction process. When $\boldsymbol{\Sigma}_u$ has full rank, the Hessian matrix of $f_t\left(\boldsymbol{x}\right)$ in (4) is positive-definite. As a result, the solution of the corresponding convex relaxation (6) is always unique.

This distinguishes our approach from the approach taken in [7, 8]. There, one needs to solve a standard CS problem (sometimes the sparse support set is restricted) at each time instance. A consequence is that sufficiently many measurements have to be taken in each time instance in order to avoid multiple plausible solutions. In contrast, there is no such issue in our approach: *in principle, one may take one measurement at each time instance*.

3. SIMULATIONS

We performed extensive numerical simulations to test our approach on both Gaussian and non-Gaussian, K-sparse dynamical signals. Due to space limitations, we present what we believe to be the most interesting findings. These findings pertain to sparse signals obtained via non-linear mappings of Gaussian signals, as described below. Note that uncented Kalman filtering and related Monte Carlo sampling methods usually applied in this case failed to produce results comparable to our approximate Gaussian CS method.

In our numerical tests, we need a dynamic system where the signal innovation $\boldsymbol{u}_t$ is Gaussian and the signal $\boldsymbol{x}_t$ at each time instance is forced to be sparse. In order to generate such a sparse Gaussian dynamic signal, we use the model

$$\boldsymbol{x}_t = \mathcal{T}_k\left(\boldsymbol{\Phi}_t\boldsymbol{x}_{t-1}+\boldsymbol{u}_t\right), \tag{7}$$

where $\boldsymbol{u}_t$ was defined in Section 2, while the nonlinear mapping $\mathcal{T}_K$ has domain $\mathbb{R}^n$ and range $\mathbb{R}^n$, and when applied to the vector $\boldsymbol{x}$, $\mathcal{T}_K\left(\boldsymbol{x}\right)$ produces a vector that agrees with $\boldsymbol{x}$ in the K largest magnitude entries, and has all other coordinates equal to zero. Note that the resulting signal is non-Gaussian, since it contains the K largest order statistics of non-Gaussian signals and have prohibitively large computational complexity.

To evaluate the performance of our proposed approach, we introduce two benchmark algorithms. The first one is the genie-aided sparse Gaussian signal reconstruction described in Algorithm 2. In this setup, the support set of the sparse signal $\boldsymbol{x}_t$ at each time instance is given as side information. Hence, one only needs to solve an MMSE estimation problem to recover the signal components in the support set. The performance of this algorithm is the best performance that one can hope to achieve without additional side information regarding the problem.

Algorithm 2 The Genie-Aided Algorithm

Let $\hat{\boldsymbol{x}}_0=\boldsymbol{0}$.
At time instance t, suppose that the support set $\mathcal{K} = \{i\in[n]:\ \boldsymbol{x}_{t,i}\neq 0\}$ is given. Let

$$\hat{\boldsymbol{x}}_{t,\mathcal{K}} = \left(\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)_{\mathcal{K}} + \boldsymbol{E}_{t,\mathcal{K}}\left(\boldsymbol{y}_t-\left(\boldsymbol{\Phi}_t\right)_{:,\mathcal{K}}\left(\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)_{\mathcal{K}}\right),$$

and $\hat{\boldsymbol{x}}_{t,\mathcal{K}^c}=\boldsymbol{0}$, where

$$\boldsymbol{E}_{t,\mathcal{K}} = \left(\boldsymbol{\Sigma}_u\right)_{\mathcal{K},\mathcal{K}}\left(\boldsymbol{\Phi}_t\right)_{:,\mathcal{K}}^T\left(\left(\boldsymbol{\Phi}_t\right)_{:,\mathcal{K}}\left(\boldsymbol{\Sigma}_u\right)_{\mathcal{K},\mathcal{K}}\left(\boldsymbol{\Phi}_t\right)_{:,\mathcal{K}}^T+\boldsymbol{\Sigma}_v\right)^{-1}.$$

The other benchmark algorithm is a one-step Kalman filter summarized in Algorithm 3. As opposed to the traditional Kalman filter, one does not track the covariance matrix of $\boldsymbol{x}_t$: due to the nonlinear operator $\mathcal{T}_K$, the statistics of the dynamic signal cannot be represented in a closed form. Note that this algorithm ignores the sparsity assumption.

Algorithm 3 The Standard One-Step Kalman Filter

Let $\hat{\boldsymbol{x}}_0=\boldsymbol{0}$.
At time instance t, compute $\hat{\boldsymbol{x}}_t = \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{E}_t\left(\boldsymbol{y}_t-\boldsymbol{\Phi}_t\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$, where $\boldsymbol{E}_t = \boldsymbol{\Sigma}_u\boldsymbol{\Phi}_t^T\left(\boldsymbol{\Phi}_t\boldsymbol{\Sigma}_u\boldsymbol{\Phi}_t^T+\boldsymbol{\Sigma}_v\right)^{-1}$.

As a variation of the standard one-step Kalman filter, we also test a heuristic algorithm described in Algorithm 4. Both Algorithms 3 and 4 are based on standard Kalman filtering, but the latter takes the K-sparse side information into consideration.

Algorithm 4 A Sparsity-Aware Kalman Filter

Let $\hat{\boldsymbol{x}}_0=\boldsymbol{0}$.
At time instance t, compute $\tilde{\boldsymbol{x}}_t = \boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{E}_t\left(\boldsymbol{y}_t-\boldsymbol{\Phi}_t\boldsymbol{\Psi}_t\hat{\boldsymbol{x}}_{t-1}\right)$, where $\boldsymbol{E}_t = \boldsymbol{\Sigma}_u\boldsymbol{\Phi}_t^T\left(\boldsymbol{\Phi}_t\boldsymbol{\Sigma}_u\boldsymbol{\Phi}_t^T+\boldsymbol{\Sigma}_v\right)^{-1}$. Then set $\hat{\boldsymbol{x}}_t=\mathcal{T}_K\left(\tilde{\boldsymbol{x}}_t\right)$.

In our simulations we used $n=256$ and $K=32$. We set $\boldsymbol{\Psi}=\boldsymbol{I}_n$, $\boldsymbol{\Sigma}_u=\sigma_u^2\boldsymbol{I}_n$ and $\boldsymbol{\Sigma}_v=\sigma_v^2\boldsymbol{I}_m$, where $\sigma_u=\sigma_v=0.01$.

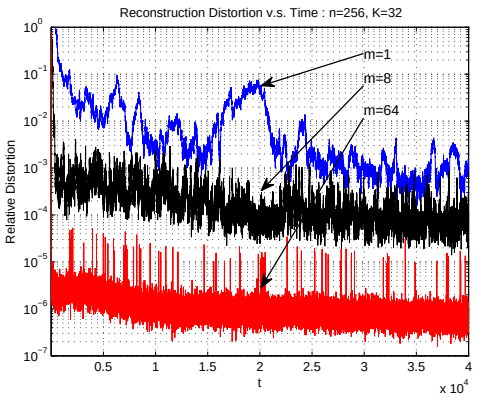

Fig. 1: Distortion versus sample size trade-off.

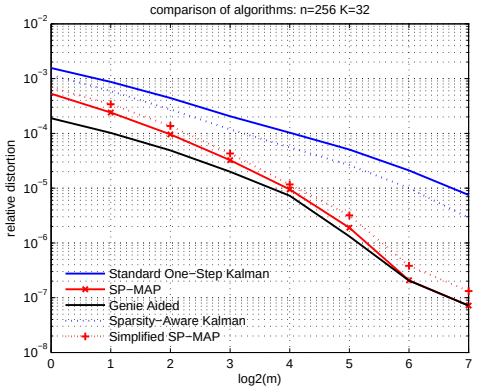

Fig. 2: Comparison of reconstruction algorithms.

For each algorithm, we ran five realizations. For each realization, we used an observation time t of the dynamic system ranging from 1 to 2e5. In each realization, the initial state of the dynamic system was chosen randomly, i.e., the support set $\mathcal{K}$ of $\boldsymbol{x}_0$ was randomly generated from the uniform distribution on all subsets of $[n]$ with cardinality K, $\boldsymbol{x}_{0,\mathcal{K}^c} = \mathbf{0}$, while the values of $\boldsymbol{x}_{0,\mathcal{K}}$ were randomly generated from the Gaussian distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{I}_K)$. Note that we deliberately used $\hat{\boldsymbol{x}}_0 = \mathbf{0}$, which represents a poor guess for $\boldsymbol{x}_0$, in order to understand how fast an algorithm can lock onto an actual dynamic sparse signal. In all our simulations, we use the relative reconstruction distortion, defined as $\|\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t\|_2^2 / \|\boldsymbol{x}_t\|_2^2$, as the performance measure.

Due to space limitations, we present only two sets of simulation results about the SP-MAP algorithm. Since the performance of the ℓ_1-MAP method depends on the choice of the parameter μ, the corresponding numerical tests are omitted here. See the companion paper [4] for more details.

The first results, shown in Figure 1, demonstrate the tradeoffs between the number of samples taken at each time instance and the reconstruction distortion. Notice that as the time index increases, taking even *one sample* per time allows for fairly accurate signal tracking. In this case, the measurements matrix clearly does not satisfy RIP-type conditions, since it consists of a single row.

The second set of results is shown in Figure 2. The figure compares the proposed SP-MAP algorithm and Algorithms 2, 3 and 4. For fair comparison, we accounted only for the relative distortion of these algorithms after they lock onto the dynamical sparse signal: in particular, we average the relative distortions at time instances from 1e5+1 to 2e5. According to the simulation results, the SP-MAP algorithm outperforms the algorithms based on standard Kalman approach. Furthermore, as the number of samples per time instance increases (yet remains a small constant), the SP-MAP algorithm performs close to the genie-aided approach.

At the first glance, the SP-MAP algorithm may require many iterations per time instance and therefore its computational complexity may be an issue. To address this concern, we numerically tested a simplified SP-MAP algorithm that executes only the initial steps of the SP-MAP algorithm and does not loop through any of the iterations. Simulation results in Figure 2 show that the performance loss introduced by this modification is negligible. Hence, when computational resources are restricted, one may skip the iterations in the SP-MAP algorithm. Furthermore, a quick check of the simulations of the SP-MAP algorithm reveals that only one or two iterations are needed for most time instances.

4. REFERENCES

[1] M. Salman Asif and J. Romberg. Dynamic updating for sparse time varying signals. *Proceedings of the CISS, Princeton*, 2009.

[2] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.

[3] Wei Dai and Olgica Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inform. Theory*, 55(5):2230–2249, May 2009.

[4] Wei Dai, Dino Sejdinovic, and Olgica Milenkovic. Sequential compressive sensing for sparse gaussian signals. *In preparation*.

[5] D. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[6] C. Rozell, D.Johnson, R. Baraniuk, and B. Olshausen. Locally competitive algorithms for sparse approximation. *Proceedings of the ICIP*, 2007.

[7] N. Vaswani. Analyzing least squares and kalman filtered compressed sensing. *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Taipei, Taiwan*, 2009.

[8] N. Vaswani and W. Lu. Modified-compresive sensing: Modifying compressive sensing for problems with partially known support. *IEEE Intl. Symp. Info. Theory (ISIT), Korea*, 2009.