Particle Filtered Modified Compressive Sensing (PaFiMoCS) for Tracking Signal Sequences

Samarjit Das and Namrata Vaswani
samarjit@iastate.edu, namrata@iastate.edu

Iowa State University, Ames, IA 50010

ABSTRACT

In this paper, we propose a novel algorithm for recursive reconstruction of a time sequence of sparse signals from highly under-sampled random linear measurements. In our method, the idea of recently proposed regularized modified compressive sensing (reg-mod-CS) is merged with sequential monte carlo techniques like particle filtering. Reg-mod-CS facilitates sequential reconstruction by utilizing a partial knowledge of the support and previous signal estimates. Under the assumption of a dynamical model on the support, the sequential Monte Carlo step renders various possibilities of the current support and choose the most likely support given the current observation. The algorithm is similar in sprit to particle filter with mode tracker (PF-MT), where, the support could be considered to be the effective basis whereas the signal values on the current support could be considered to be the residual space; the difference being the fact that in our algorithm, the mode tracking step is replaced by the reg-mod-CS for each particle and the support is re-estimated from the residual part. We compare our algorithm with other techniques like traditional particle filtering, particle filter with mode tracker (PF-MT), static compressive sensing, modified compressive sensing, regularized modified compressive sensing and weighted ℓ_1. We demonstrate that our algorithm outperforms all the other methods in terms of reconstruction accuracy for a simulated sequence of sparse signals.

Index Terms— Sequential estimation, sparse reconstruction, compressive sensing, regularized modified compressed sensing, particle filtering, mode tracker

1. INTRODUCTION

Our primary goal is causal and recursive reconstruction of a time sequence of sparse signals. We assume a slowly changing sparsity pattern over time. Also, we want our algorithm to use as few linear measurements at each time as possible. The "recursive" aspect of the algorithm is important i.e. use the current measurements and the previous reconstruction to get the current reconstruction. Potential applications of such an algorithm could be real-time dynamic MRI, single-pixel imaging etc. with faster acquisition and faster reconstruction.

The theory of compressive sensing (CS) states that a large dimensional sparse signal can be reconstructed from a set of highly under-sampled random measurements by using convex optimization techniques [1, 2]. Mathematically, the sparse reconstruction problem can be posed as follows. Say, we want to reconstruct a sparse signal x, with support N, from $y := \Phi x$, when n = length(y) $< m$ = length(x). This problem can be solved under certain conditions [1, 2] if we can find the sparsest vector satisfying $y = \Phi x$. The solution can be found via exhaustive search; but that leads to exponential complexity. Many practical approaches (polynomial complexity in m) have been proposed in recent years to solve this problem [10, 2, 1, 19, 11, 12].

In recent literature, [3, 5, 6] proposes that if we have a partial knowledge of the support of the signal, a modified version of compressive sensing (called modified CS or mod-CS) can reconstruct the signal from even lesser number of measurements as compared to traditional CS. The key idea of mod-CS is: given a partial but partly erroneous support "knowledge": T, we can rewrite $N :=$ support(x) as $N = T \cup \Delta \setminus \Delta_e$. Here, $\Delta := N \setminus T$, is the unknown set of misses in T and $\Delta_e := T \setminus N$, is the unknown set of extras in T. If $N = T \cup \Delta$, above problem is equivalent to finding the signal that is sparsest on T^c such that data constraint is satisfied. Modified-CS [3, 4] attempts to solve the problem as follows,

$$\min_{\beta} ||(\beta)_{T^c}||_1 \text{ s.t. } y = \Phi\beta \tag{1}$$

Exact reconstruction conditions for mod-CS are much weaker than that of CS when $|\Delta| << |N|$ and $|\Delta_e| \ll |N|$. A similar approach (called the weighted ℓ_1) was proposed in parallel work by [17]. For making the algorithm robust to large observation noises, more constraints can be put in the above optimization problem. Regularized-modified-CS [7] does exactly that. It assumes that the signal values in the known part of the support changes slowly over time (which is a reasonable assumption for many practical applications) and the corresponding optimization problem is modified as,

$$\begin{aligned}\hat{x}_t = \arg\min_x \lambda||x_{T^c}||_1 \quad &+ \quad \gamma||(x)_T - (\hat{x}_{t-1})_T||_2^2 \\ &+ \quad ||y_t - \Phi x||_2^2 \end{aligned} \tag{2}$$

Notice that all the above mod-CS type techniques are particularly suited for sequential reconstruction of sparse signals where, the support at the previous instant can be used as the predicted support. Some other related work on sequential sparse reconstruction include Kalman Filtered compressive sensing (KF-CS) [13], dynamic Lasso [14], RLS-Lasso [15] and static Bayesian approaches - [18, 19].

Consider a situation that we have a Markov model on the support change as well as non-zero signal value change. The question is - can we do better than mod-CS by devising a mechanism to make a 'more judicious' prediction about the support ? Our algorithm, Particle Filtered Modified Compressive Sensing (PaFiMoCS) attempts to answer this question. Under the assumption that the sparsity pattern variations might evolve from a dynamical model, PaFiMoCS uses the idea of sequential importance sampling from the particle filtering literature, and provides the modified compressive sensing with the support information. This enables mod-CS style algorithms to work more efficiently i.e. use lesser number of observations.

2. PROBLEM FORMULATION AND NOTATION

Say, $\{x_t\}_{t\geq 0}$ is a sparse signal sequence with $x_t \in \mathcal{R}^m$ and $m \ll K, \ \forall t = 0, 1, \ldots$ where K is the number of non-zero entries in x_t. The observations $\{y_t\}_{t\geq 0}, y_t \in \mathcal{R}^n$ are generated from the following observation model,

$$y_t = \Phi x_t + n_t \tag{3}$$

where Φ is a $n \times m$ measurement matrix with $n < m$ and the observation noise is normal distributed as, $n_t \sim \mathcal{N}(\mathbf{0}, \Sigma_o)$. Under this setup, our primary objective is to sequentially reconstruct the signal x_t from the observations $y_0, \ldots y_t$. In order to do that we develop a recursive algorithm which causally reconstructs the signal x_t from the observation y_t and signal estimate from the previous instant i.e. $\hat{x}_{t-1}$. Also, it is assumed that the sparsity pattern of x_t follows a certain dynamical model as described in the next subsection. But first, we give some of the basic notations that we use in the sections to follow.

Denote the sparse signal at time t as x_t. Define the support of the signal x_t as, $N_t = \{i : (x_t)_i \neq 0\}, \ \ (x_t)_i \triangleq i^{th}$ component of x_t. The term $(x_t)_{N_t}$ denotes a vector comprising of the components of x_t corresponding to the indices belonging to the set N_t. The set of indices of x_t that do not belong to N_t is denoted as N_t^c. The symbol $'\backslash'$ denotes $set\ minus$. While going from $t-1$ to t, the set of new elements being added to the support is denoted as A_t, whereas the set of deleted elements is denoted as R_t. Thus $A_t = N_t \setminus N_{t-1}$ and $R_t = N_{t-1} \setminus N_t$. The function $dim(.)$ gives the dimension of a vector. We define the state of the system at t as :

$$X_t = [x_t; N_t] \tag{4}$$

Next, we describe the generative model for the sparse signal sequence.

2.1. Signal Dynamics : The Generative Model

The generative model for the sparse signal sequence is given as follows.

1 At $t = 0$, start with some N_0 with $|N_0| << m$. Get x_0 as follows,

$$\begin{aligned} (x_0)_{N_0} &= \boldsymbol{\nu_0}, \ \ \boldsymbol{\nu_0} \sim \mathcal{N}(\mathbf{m_0}, \Sigma_0) \\ (x_0)_{N_0^c} &= \mathbf{0} \end{aligned}$$

Thus for $t = 0$, obtain $X_0 = [x_0, N_0]$.

2 For $t > 0$ compute,

$$N_t = (N_{t-1} \cup A_t) \setminus R_t$$

where the sets A_t and R_t are generated as follows,

$$\begin{aligned} A_t &\sim \text{Unif}_p(N_{t-1}^c) \\ R_t &\sim \text{Unif}_p(N_{t-1}^*) \\ \text{with } N_{t-1}^* &= \{j : |(x_{t-1})_j| < \Delta, j \in N_{t-1}\} \end{aligned} \tag{5}$$

where the operator $\text{Unif}_p(N)$ selects any p unique entries of N at random, $p << |N_{t-1}|$ and Δ is a heuristically chosen threshold. Finally, generate the signal x_t as,

$$\begin{aligned} (x_t)_{N_t} &= (x_{t-1})_{N_t} + \boldsymbol{\nu}, \text{ where, } \boldsymbol{\nu} \sim \mathcal{N}(\mathbf{0}, \Sigma_\nu) \\ (x_t)_{N_t^c} &= \mathbf{0} \end{aligned}$$

Thus, for $t > 0$ get $X_t = [x_t, N_t]$

3 Set $t \leftarrow t + 1$ and go to step 2.

This dynamical model on x_t and in turn, on X_t, serves as the system model in a tracking framework. The signal dynamics is demonstrated in Fig. 1. Now, a dynamical model on sparsity pattern change is a reasonable assumption as long as it is *generic* in nature. A closer look at equation (5) reveals that the new additions to the support can come anywhere from N_t^c. The second equation comes from the fact that whenever a component of the support is going to be deleted, it is very likely that it would have a small signal amplitude corresponding to that component. In other words, no component of the support, with 'significant' signal amplitude, gets abruptly deleted from the support. If any component were to go, it fades away gradually rather than abruptly. This assumption is quite practical for signals evolved from natural processes for which *continuity* is an inherent characteristics.

3. PARTICLE FILTERING AND PARTICLE FILTERED MODIFIED-CS (PAFIMOCS)

In this section, we briefly discuss sequential Monte Carlo based algorithms for estimating x_t at each instant form the observations $y_0, \ldots, y_t$. The key question here is : given a

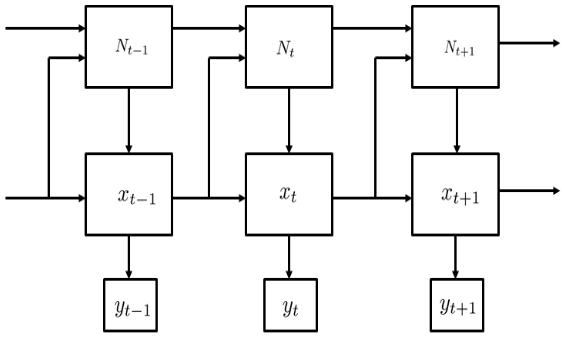

Fig. 1. The dynamical model corresponding to the sparse signal sequence.

Markov model on slow support change, and on slow nonzero signal value change, what is the "best" way to use it? One idea is Particle Filtering(PF) framework i.e. sequential importance sampling from the dynamic prior model for MAP estimates of the support and the signal.

3.1. Particle Filteres

We are in search of an optimal estimator of x_t under the above problem formulation. One idea could be compute the Maximum A Posterior (MAP) estimate of the signal (i.e. the one which maximizes the posterior distribution $p(X_t|y_{1:t})$). The basic particle filter or PF [16] gives an empirical estimate of this posterior distribution. It is to be noted that from the dynamical model described above, it can be shown that $p(X_t|X_{t-1}) = p(x_t|x_{t-1}, N_t)p(N_t|N_{t-1}, x_{t-1})$ which means that the PF can render various possibilities of the support and the signal values (i.e *Importance Sampling or IS*) and assign importance weights based on the observation likelihood (OL) i.e. $p(y_t|X_t)$. Then the MAP estimate becomes the importance sample with highest importance weight. A posterior mode-tracking (MT) version of the PF (PF-MT) [20] can be implemented as well. Here, the support could be considered to be the effective basis whereas the signal values on the current support could be considered to be the residual space. The main difference between the basic PF algorithm and PF-MT is that in PF-MT, the importance sample for the residual space is generated by posterior mode tracking rather than using the dynamical model alone [20]. Under our current problem setup, the mode tracker for the residual part could be replaced by a regularized modified CS step as follows,

$$x_t^{(i)} = \arg\min_x \lambda||(x - x_{t-1}^{(i)})_T||_2^2 + \gamma||x_{T^c}||_1 + ||y_t - \Phi x||_2^2$$

here the superscript '(i)' stands for the i^{th} particle i.e. each particle for x_t is generated by performing reg-mod-CS with the current observation and the corresponding particle at the previous instant.

The basic PF steps are given in Algorithm. 1. The problem with these algorithms is that both of them do not use the

Algorithm 1 PF : Recursive Estimation of $X_t = [N_t, x_t]$

1 At $t = 0$, $N_0^{(i)} = N_0$ and $x_t^{(i)} = \hat{x}_0$, for $i = 1, 2, ..., N_{PF}$

2 At each time $t > 0$, for $i = 1, 2, ..., N_{PF}$

(a) Importance Sample on signal support : $A_t^{(i)} \sim \text{Unif}_p({N_{t-1}^{(i)}}^c)$ and $R_t^{(i)} \sim \text{Unif}_p({N_{t-1}^{(i)}}^*)$. Get $N_t^{(i)} = (N_{t-1}^{(i)} \cup A_t^{(i)}) \setminus R_t^{(i)}$ and define $T \triangleq N_t^{(i)}$

(b) Importance sampling on the signal values : $x_t^{(i)}$: $(x_t^{(i)})_{T^c} = \mathbf{0}, x_T^{(i)} \sim \mathcal{N}((x_{t-1}^{(i)})_T, \Sigma_\nu)$

(c) Assign importance weights : $w_t^{(i)} \propto p(y_t|X_t) \propto \exp(-\frac{1}{2}(y_t - \Phi x_t^{(i)})^T \Sigma_o^{-1}(y_t - \Phi x_t^{(i)}))$ and resample

3 MAP estimate : output maximum weight particle

4 Resample and set $t \to t + 1$ and go to step 2

observations in the support prediction process. Basic PF [16] does not even utilize the sparsity of the underlying signal. PF-MT uses slow sparsity change to estimate x_t but not N_t. Later, we demonstrate from simulation experiments that the support estimation error of these algorithms accumulate over time, making them unstable.

3.2. Particle Filtered Modified CS (PaFiMoCS)

The primary goal of PaFiMoCS is to facilitate the reg-mod-CS algorithm with a reasonably good estimate of the true support. In prior works, researchers used the previous support estimate as the current predicted support. But, as stated above, this heuristic does not work when there happens to be a large support change over successive time instants. PaFiMoCS, precisely, attacks this problem. It leverages the fact that we can do a better job in support prediction by assuming a dynamical model on the support changes. Now, even if there happens to be a large support change, we still have a better handle of the situation than just using the previous support estimate for the prediction.

PaFiMoCS is developed based on the dynamical model for support change as well as dynamical model on signal value change on the known part of the support (Sec. 2.1). At each instant, it uses the importance sampling of particle filtering to generate various support guesses. Here, the hope is that for a large enough number of samples, at least one of them would be very close to the true support. And then, solve reg-mod-CS for each of these guesses to generate importance samples of the signal values on the corresponding support particle. The support particle is then updated by re-computing it from the reg-mod-CS output. The entire procedure is summarized in Algorithm. 2 .

It is important to note that there is a significant difference between our algorithm and PF-MT. In our algorithm, unlike PF-MT, we actually recompute the effective basis (i.e $N_t^{(i)}$)

Algorithm 2 PaFiMoCS: Tracking Algorithm

Goal: From the observations $\{y_t\}_{t=0}^T$, sequentially estimate the sparse signal $\{\hat{x}_t\}_{t=0}^T$ where $y_t = \Phi x_t$.

1 Assume at $t = 0$, we know $\hat{x}_0$, N_0. For $i = 1 \dots N_{PF}$, assign particle set as,

$$\begin{aligned} N_0^{(i)} &= N_0 \\ x_0^{(i)} &= \hat{x}_0 \end{aligned}$$

2 At $t > 0$, for $i = 1 \dots N_{PF}$:

$$\begin{aligned} (A_t^{(i)})_j &\sim \text{Unif}_p(N_{t-1}^{(i)\,c}) \\ (R_t^{(i)})_j &\sim \text{Unif}_p(N_{t-1}^{(i)\,*}), \\ N_{t-1}^{(i)\,*} &= \{k : |(x_{t-1}^{(i)})_k| < \Delta, k \in N_{t-1}^{(i)}\}, \end{aligned}$$

Importance sampling on the support : $N_t^{(i)} = (N_{t-1}^{(i)} \cup A_t^{(i)}) \setminus R_t^{(i)}$. Then, using current observation y_t, perform Regularized modified CS (sort of substitute of PF-MT) to importance sample on the signal x_t as,

$$x_t^{(i)} = \arg\min_x \lambda ||(x - x_{t-1}^{(i)})_T||_2^2 + \gamma ||x_{T^c}||_1 + ||y_t - \Phi x||_2^2$$

where $T = N_t^{(i)}$. Recompute the support as, $N_t^{(i)} = \{j : |(x_t^{(i)})_j| > \alpha\}$. Thus get, $X_t^{(i)} = [x_t^{(i)}, N_t^{(i)}]$.

3 Assign appropriate importance weight as,

$$\begin{aligned} w_t^{(i)} &\propto \frac{p(y_t|X_t^{(i)}) p(X_t^{(i)}|X_{t-1}^{(i)})}{\pi(X_t^{(i)}|X_{t-1}^{(i)}, y_t)} \\ &\propto \frac{e^{-||y_t - \Phi x_t^{(i)}||_2^2} e^{-\frac{1}{2}((x_t^{(i)} - x_{t-1}^{(i)})_T)^T \Sigma_\nu^{-1}((x_t^{(i)} - x_{t-1}^{(i)})_T)}}{e^{-(\lambda ||(x_t^{(i)} - x_{t-1}^{(i)})_T||_2^2 + \gamma ||(x_t^{(i)})_{T^c}||_1 + ||y_t - \Phi x_t^{(i)}||_2^2)}} \end{aligned}$$

5 MAP estimate : output maximum weight particle

6 Resample and set $t \leftarrow t + 1$ and go to step 2.

---

from the importance sample of the residual part i.e. $x_t^{(i)}$. This step is essential in our algorithm because, the reg-mod-CS step corrects for any residual error in $N_t^{(i)}$ w.r.t the true support N_t. Thus the re-estimated support tends to be more accurate than the one generated by the importance sampling step. Otherwise, this error might accumulate over time and lead to an unstable algorithm with exploding reconstruction error (see performance plot of PF-MT).

4. EXPERIMENTAL RESULTS AND DISCUSSION

We tested our algorithm on a simulated sequence of sparse signals. The dynamical model described in Sec. 2.1 was used for generating a sequence of length $T = 20$. The signal dimension was set at $dim(x_t) = m = 200$ with sparsity $K = dim(N_t) = 20$ (i.e. 10% of the signal dimension). The initial signal ($t = 0$) was generated using $N_0 = \{10, 15, 20, \dots, 105\}$ and $\mathbf{m_0} = 5 * \mathbf{1}_{20}$ and $\Sigma_0 = 3 * \mathbf{I}_{20}$ where $\mathbf{1}_k$ is a k-dimensional column vector with all entries as 1 and $\mathbf{I}_k$ is a $k \times k$ identity matrix. We considered two additions and two deletions over successive time instants i.e. $p = 2$ or $dim(A_t) = 2$ and $dim(R_t) = 2$. The parameter Σ_ν corresponding to the dynamics of signal values was set to be $\mathbf{I}_{20}$. The entries on the observation matrix Φ was generated from a standard normal distribution with $n = 50$ (25% of the signal dimension). The observation noise was generated from a multivariate standard normal distribution i.e. $\mathcal{N}(\mathbf{0}, \mathbf{I})$ or $\sigma_o^2 = 1$ for each dimension. Under all the above parameter settings, we simulated a 200 dimensional, 20-sparse signal sequence $\{x_t\}$ and generated corresponding observation vectors $\{y_t\}$. These observations are fed to the various reconstruction algorithms to study their reconstruction performances.

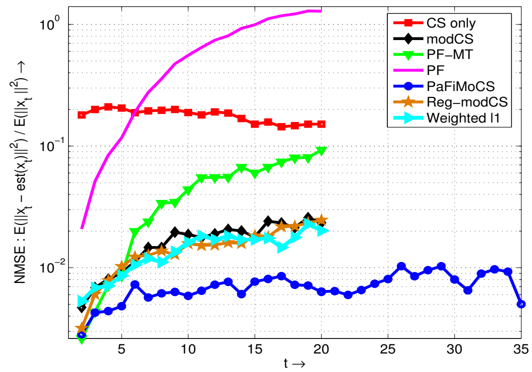

Fig. 2. Performance comparison plots. We compute the Normalized Mean Squared Error for all the methods and compared them over 50 Monte Carlo runs. It can be seen that PaFiMoCS has the best performance among all the methods.

We compared the reconstruction performances of seven different methods including the basic particle filter, traditional compressed sensing, modified/regularized modified compressed sensing without sequential Monte Carlo sampling, weighted ℓ_1, PF-MT and finally PaFiMoCS. For simplicity, it was assumed that we have a perfect knowledge of the initial state of the system i.e. $\hat{X}_0 = X_0$. For regularized mod-CS the parameters λ and γ were set at 0.1 and 0.5 respectively. For basic particle filter and PaFiMoCS the number of particles used was $N_{PF} = 100$. The performance of the algorithms were compared by computing the normalized mean-squared reconstruction error (NMSE) i.e. $\frac{E(||x_t - \hat{x}_t||^2)}{E(||x_t||^2)}$ for $t = 0, ..., 20$. The expectations were computed over 50 Monte Carlo runs of each algorithm. The corresponding NMSE plots for various algorithms are shown in Fig. 2. The re-computation of the support for each particle made sure that each support particle reflects the support error correction after the reg-mod-CS step. That is why this algorithm has a low and stable NMSE of the order of 10^{-2}. The other variant, PF-MT which does not recompute the support for each particle (more similar to a PF-MT approach) gradually loses track of the signal as the support error accumulates over time. Other algorithms like mod-CS, reg-mod-CS and weighted ℓ_1 without the importance sampling step, performs worse as they do not have a support prediction as good as PaFiMoCS (remember that all of them used the previous support estimate as the current predicted support). Traditional compressed sensing fails with 25% observations as it does not utilize the dynamics of sparsity pattern change as well as the constraints on the signal value changes. Basic particle filter primarily fails due to the lack of a mechanism to utilize the sparsity of the underlying signal.

Thus it turns out that by facilitating a regularized compressed sensing approach with a good support estimate from a PF framework, we can actually come up with a very efficient algorithm for sequential reconstruction of sparse signals from highly under-sampled random linear measurements.

5. CONCLUSION

In this paper, we have proposed particle filtered modified compressed sensing (PaFiMoCS) for sequential estimation (i.e. tracking) of sparse signals from highly under-sampled measurements. The basic idea is to use recently proposed reg-mod-CS idea by providing it with a 'close enough' prediction of the support. This prediction step is carried out by sequential importance sampling technique used is particle filtering literature. The key difference between PaFiMoCS and other sequential CS based approaches is that it uses dynamical models on the support change as well as signal value change on the known part of the support. It is demonstrated with simulation experiments that PaFiMoCS performs better than other algorithms which do not utilize the importance sampling step for support prediction. As a part of future research, this technique can be used in many practical scenarios where PF-MT [20] is suitable but the residual space has a slowly varying sparsity pattern over time. For example - gradual illumination variations in a scene or slowly moving occlusion in a template based visual tracking framework.

6. REFERENCES

[1] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52(4), pp. 1289–1306, April 2006.

[2] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Info. Theory*, vol. 52(2), pp. 489–509, February 2006.

[3] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Processing*, September 2010.

[4] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," in *ISIT 2009*, pp. 488 –492, June 2009.

[5] N. Vaswani, "LS-CS-residual (LS-CS): Compressive Sensing on Least Squares residual," *IEEE Trans. Signal Processing*, vol. 58(8), pp. 4108–4120, 2010

[6] W. Lu and N. Vaswani, "Modified Basis Pursuit Denoising (Modified-BPDN) For Noisy Compressive Sensing With Partially Known Support," in *ICASSP*, 2010.

[7] W. Lu and N. Vaswani, "Regularized Modified-BPDN for compressive sensing with partially known support," in *Arxiv.org : http://arxiv.org/abs/1002.0019*, 2010.

[8] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53(12), pp. 4655–4666, December 2007.

[9] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. on Information Theory*, (Submitted) 2007.

[10] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," (Preprint) 2008.

[11] S. S. Chen, D. L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.

[12] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. on Information Theory*, vol. 52(12), pp. 5406 – 5425, December 2006.

[13] N. Vaswani, "Kalman filtered compressed sensing," in *IEEE Intl. Conf. on Image Processing*, San Diego, CA, October, 2008.

[14] D. Angelosante, E. Grossi, and G. B. Giannakis, "Compressed sensing of time-varying signals," in *DSP*, 2009.

[15] D. Angelosante,G. B. Giannakis,"RLS-weighted Lasso for adaptive estimation of sparse signals" in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009.

[16] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/nongaussian bayesian state estimation," *IEE Proceedings-F (Radar and Signal Processing)*, pp. 140(2):107–113, 1993.

[17] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, "Weighted l1 minimization for sparse recovery with prior information," in *Proceedings of the 2009 IEEE international conference on Symposium on Information Theory - Volume 1*, ISIT'09, pp. 483–487, 2009.

[18] P. Schniter, L. C. Potter, and J. Ziniel, "Fast bayesian matching pursuit: Model uncertainty and parameter estimation for sparse linear models," 2009.

[19] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," 2007.

[20] N. Vaswani, "Particle Filtering for Large Dimensional State Spaces with Multimodal Observation Likelihoods," *IEEE Trans. Sig. Proc.*, pp. 4583–4597, October, 2008.