TC-CSBP: Compressive Sensing for Time-Correlated Data Based on Belief Propagation

Behzad Shahrasbi, Ali Talari and Nazanin Rahnavard
School of Electrical and Computer Engineering
Oklahoma State University
Stillwater, OK 74078
Emails: {behzad.shahrasbi, ali.talari, nazanin.rahnavard}@okstate.edu

***Abstract*—Existing compressive sensing techniques mostly consider the sparsity of signals in one dimension. However, a very important case that has rarely been studied is when the signal of interest is time varying and signal coefficients have correlation in time. Our proposed algorithm in this paper is a structure-aware version of the compressive sensing reconstruction via belief propagation proposed by Baron et al. that exploits the time correlation between the signal components and provides the belief propagation algorithm with more accurate initial priors. Numerical simulations show that the belief propagation-based compressive sensing algorithm is able to utilize the side information about signal's time correlation and results in enhanced reconstruction performances.**

***Index Terms*—Compressive sensing, belief propagation reconstruction, Markov model parameter estimation, time correlation**

I. Introduction

In many signal processing applications, such as image processing and wideband signal processing, an intelligent inspection of signal components reveals that these components are sparse or can be sparsely expressed in a proper basis, e.g. wavelet domain. This observation motivates us to exploit the sparsity of such *compressible signals* to save sampling, communication, processing, and memory resources. These efforts have opened a new area of research known as *compressive (compressed) sensing* (CS) [1–3]. The underdetermined problem of reconstructing a sparse signal with length n, from its compressed sensed measurement vector with length m, where $m < n$, is possible by using existing CS algorithms [4–10].

In addition to sparsity, many of the real-world signals have signal components which vary slowly in time. An interesting example is sensor networks in which the signal of interest represents data from temperature sensors that are collected during a time interval T with unit time steps. Such readings have both *spatial correlation* due to closeness of sensors and *time correlation* due to the smooth variations in the temperature. Such a time correlation can further help us to reconstruct the signal at each time step using the estimated signal from the previous time step.

In recent years, a considerable amount of research has been conducted to take advantage of the prior knowledge in the reconstruction algorithms [11–20]. The reconstruction algorithms are mostly the modifications of ℓ_1 reconstruction, except [11, 18, 20] which modify Orthogonal Matching Pursuit [7]. In a recent study [17], CS for the time-correlated signals is considered. However, the proposed scheme requires the collection of all measurements over interval T to reconstruct the coefficients of signals in T. In [12, 16] a weighting strategy is applied to include the prior information into the support of the sparse signal. In other words, the extra information is modulated in terms of different weights for the different parts of the signal support. In [19, 20] the reconstruction of jointly sparse signals in both spatial and temporal domains is considered, and the correlation in spatial domain is employed. In [19], it is assumed that the supports of all correlated signals share an equal common part plus a unique sparse innovation part. In [20], the support is fixed and there is spatial and temporal correlation between signals of the different sources. In this paper, we assume a different model for the signal than [19, 20].

In [13–15], the authors have developed a novel algorithm that calculates the least-square residuals of the signal supports in two consecutive time steps instead of directly estimating the signal support. In [15] a Kalman filter-based algorithm is introduced to dynamically estimate the residuals.

The contribution of this paper is two-fold. First, in contrast to the other previous work that are based on linear programming or greedy algorithms, in our proposed algorithm the reconstruction is based on Bayesian inference. Second, the model that is considered in this work can be easily adapted to many real-world applications.

The remainder of the paper is organized as follows. In Section II, we introduce our time-correlated signal model and we describe the proposed reconstruction algorithm based on the proposed model. A discussion on the model parameters mismatch and the model parameters estimation is provided in Section II-C. Section III reports the simulation results. Finally, Section IV concludes the paper.

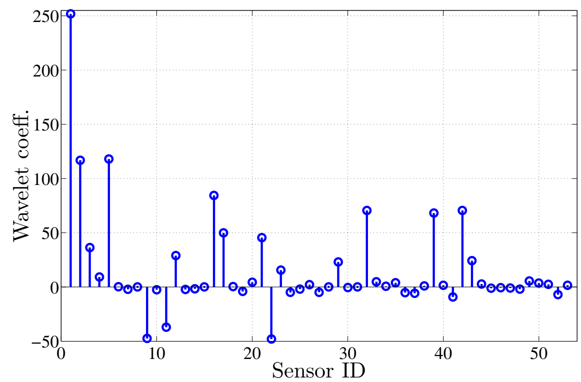
II. Time-correlated data reconstruction

In this section, we first introduce our time correlation model of signal $\underline{x}_t$. Next, we briefly review the CSBP algorithm [10]. Then, we elaborate our approach on incorporating the time correlation in CSBP algorithm.

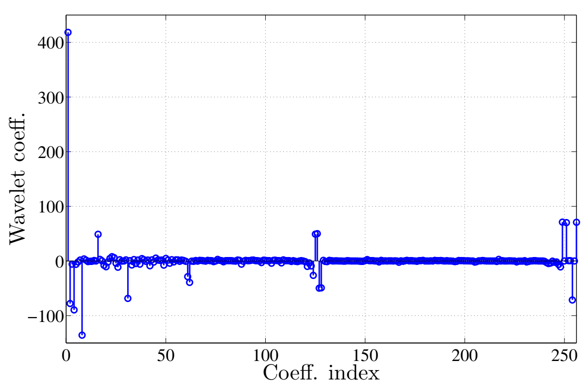
978-1-4244-9848-2/11$26.00©2011 IEEE

A. Time-varying signal model

As mentioned in Section I, sensor readings in sensor networks represent correlation in both space and time. For example, using the data provided in [21], we depict Fig. 1, which shows the correlation of temperature readings of many thermal sensors in space and time. In Fig. 1(a), the readings of 54 sensors at $t = 1$ shows a sparse behavior in wavelet domain (correlation in space). On the other hand, the generated signal from each sensor in time is sparse in wavelet domain (Fig. 1(b)) due to time correlation among the sensor readings. We define $\underline{x}_t = [x_{1,t}, x_{2,t}, \ldots, x_{n,t}]$ and $\underline{y}_t = \phi_t \underline{x}_t$ as the signal and CS measurement vectors in time t, respectively. We have $x_{i,t}$ is the value of i^th signal coefficient at time t and ϕ_t is the CS projection matrix at time t. Let $\hat{\underline{x}}_t$ denote the estimate of $\underline{x}_t$ employing a CS recovery scheme. Our goal is to find $\hat{\underline{x}}_t$ exploiting both $\hat{\underline{x}}_{t-1}$ and $\underline{y}_t$. The information $\hat{\underline{x}}_{t-1}$ is used as *a priori* information in a proper CS decoder to enable the recovery of $\hat{\underline{x}}_t$ with fewer number of measurements as it is shown in Fig 2. Only a few of the existing CS algorithms are capable of taking advantage of a priori information about the signal. Among them, we are choosing the state-of-the-art CSBP [9, 10] to implement our proposed ideas.



(a) Wavelet coefficients of sensor readings signals ($t = 1$)



(b) Wavelet coefficients of sensor readings in time (Sensor ID= 1)

Fig. 1. Real temperature sensor readings from UC-Berkeley Intel lab. The readings of 54 temperature sensors are considered for 256 consecutive time steps (i.e., every 30 secs). The correlation of the sensor readings in both spatial domain and the time domain can be seen.

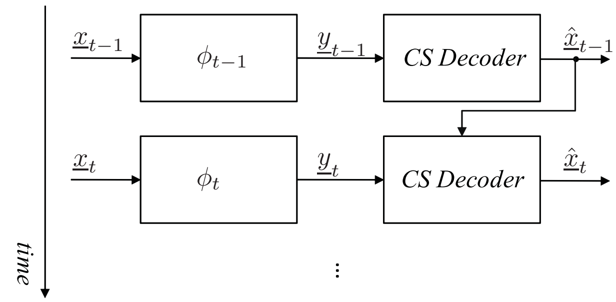


Fig. 2. Reconstructed signal at the previous step is fed as *a priori* knowledge to the decoder.

Let $\underline{X}_t = [X_{1,t}, X_{2,t}, \ldots, X_{n,t}]$ be a random vector and consider $\underline{x}_t$ as an outcome of $\underline{X}_t$. We know $\underline{X}_t$ and $\underline{X}_{t-1}$ are not independent. Hence, in order to build a mathematical model, we can exploit the correlation between their coefficients (i.e. $X_{i,t-1}$ and $X_{i,t}$ for $i = 1, \ldots, n$) in the time domain. We assume $\underline{x}_t$ has only k distinguishable coefficients from the noise level, where $k \ll n$ (k-*sparse* signal). $\frac{k}{n}$ is defined as the sparsity rate and the value of k can be derived using the history of the signal.

We model variations of each coefficient's value in time by a Markov model (MM) with the state transition matrix $A = \begin{pmatrix} a_{ss} & a_{sl} \\ a_{ls} & a_{ll} \end{pmatrix}$ as shown in Fig. 3. According to this model, each element of vector $\underline{X}$ can be in one of the two states large ($\mathcal{L}$) and small ($\mathcal{S}$) that represent whether or not its magnitude is distinguishable from the noise level.

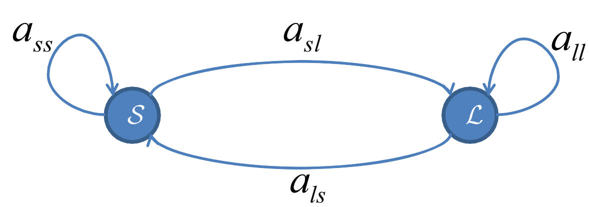


Fig. 3. Markov model for transition from $X_{i,t-1}$ to $X_{i,t}$.

Initially (at $t = 0$) and similar to [10], the signal coefficients are generated according to the following mixture Gaussian distribution

$$f(X_{i,0}) = (\frac{k}{n})\mathcal{N}(0, {\sigma_1}^2) + (1 - \frac{k}{n})\mathcal{N}(0, {\sigma_0}^2), \qquad (1)$$

where $\sigma_1 \gg \sigma_0$ and $f(.)$ is the probability density function (pdf). This assumption implies that a fraction $\frac{k}{n}$ of the coefficients of the vector $\underline{X}_0$ are distributed normally according to $\mathcal{N}(0, {\sigma_1}^2)$ and the remaining coefficients are distributed according to $\mathcal{N}(0, {\sigma_0}^2)$. Given (1), we define a threshold as $th = 3\sigma_0$ (almost all of the small coefficients lie inside this range), which means that if $\hat{x}_{i,t} > th$ then $\hat{x}_{i,t}$ is in the state $\mathcal{L}$, otherwise it is in the state $\mathcal{S}$.

We assume an element in any state at time $t - 1$ is more likely to take its new value in the same state at time step t. This is important knowledge about the variations of the signal in time, which implies that the signal is slowly varying.

978-1-4244-9848-2/11$26.00©2011 IEEE

We model this assumption such that every element in state $\mathcal{S}$ takes a Gaussian distributed value with mean zero and small variance σ_0^2 at each time step, while the coefficients in state $\mathcal{L}$ take a Gaussian distributed value with mean $\hat{x}_{i,t-1}$ and variance σ_0^2.

In order to maintain a fixed sparsity rate during all of the time steps, $a_{sl} = \frac{k}{n-k} a_{ls}$ should hold. In other words, at each time step on average, the total number of coefficients transiting from $\mathcal{S}$ to $\mathcal{L}$ is equal to the number of coefficients transiting from $\mathcal{L}$ to $\mathcal{S}$. Otherwise, the signal sparsity changes in time. In the rest of this paper, we assume that signal sparsity is preserved.

We ignore the effect of noisy measurements on the validity of the model for now. Later in Section II-C2, the effect of noisy measurements with standard deviation σ_z is considered. In case of noisy measurements, $\underline{y} = \phi\underline{x} + \underline{\nu}$ where $\underline{\nu} = [\nu_1, \ldots, \nu_n]$ is an outcome of random vector $\underline{N}$ and $\nu_i \sim \mathcal{N}(0, \sigma_z^2)$.

B. CSBP: Compressive sensing reconstruction using belief propagation [9, 10]

The key concept in belief propagation (BP) algorithms is the exchange of beliefs back-and-forth between factor nodes and variable nodes of a factor graph. A factor graph is a bipartite graph in which any vertex from one side of the graph is only connected to the vertices on the other side of the graph [22]. The variable nodes of a factor graph represent the coefficients of $\underline{x}_t$, and the factor nodes represent randomly generated CS measurements, $\underline{y}_t$. Moreover, the connecting edges represent which coefficients of the signal vector $\underline{x}_t$ contribute in generating different measurements. The problem is finding the best estimate of each variable node's value using the observations of the factor nodes employing the BP algorithm. CSBP [9] considers the conditional pdf of each element of signal vector $\underline{x}$ as a belief [9]. and is especially very interesting since we can employ the prior knowledge of the signal model in terms of a pdf in the reconstruction algorithm.

C. TC-CSBP: Time-correlated CS algorithm based on belief propagation

If we assume the coefficients of vector $\underline{x}_t$ do not show any correlation in time, the only knowledge about the signal is its sparsity in space. Thus, the prior belief about the value of each variable node at the decoder is in the form of (1). Now we consider the time-correlation model of Section II-A is added as extra information about the signal to the BP decoding. According to the signal model introduced in Section II-A, we have

$$f\left(X_{i,t}|S\left(\hat{x}_{i,t-1}\right) = \mathcal{L}\right) \\ = a_{ll}\mathcal{N}(\hat{x}_{i,t-1}, {\sigma_0}^2) + a_{ls}\mathcal{N}(0, {\sigma_0}^2), \tag{2a}$$

$$f\left(X_{i,t}|S\left(\hat{x}_{i,t-1}\right) = \mathcal{S}\right) \\ = a_{ss}\mathcal{N}(0, {\sigma_0}^2) + a_{sl}\mathcal{N}(0, {\sigma_1}^2), \tag{2b}$$

for $i \in \{1, 2, \cdots, n\}$ and $S(.)$ represents the state of the signal that can be either large ($\mathcal{L}$) or small ($\mathcal{S}$). The proposed TC-CSBP algorithm is different from the original CSBP algorithm [9, 10] in the following key points. First, unlike the conventional CSBP, every variable node receives a *unique a priori belief* according to its previous value ($\hat{x}_{i,t-1}$). Adding this information to the model, the variability of the random vector $\underline{X}_t$ decreases (the first term in RHS of (2a) has variance σ_0^2 instead of σ_1^2). Second, in TC-CSBP the time correlation modeling precision is in tradeoff with the number of required measurements. In other words, the more accurate our time correlation model is, the fewer number of measurements is required to achieve a specific reconstruction quality. However, CSBP algorithm's performance only depends on the number of measurements.

The performance of TC-CSBP algorithm highly depends on the model parameters. Therefore, it is necessary to update the model parameters in time and to analyze the effects of model mismatch.

1) Model mismatch analysis: Although considering a time correlation model for $\underline{x}_t$ can help the decoder to reconstruct the signal more accurately and with fewer number of measurements, it could also be a source of further errors if the model is not accurate enough. Hence, we need to analyze the robustness of the proposed algorithm to model mismatch and parameter variations. We model these anomalies with random matrix Δ_t that adds up to the state transition probabilities at time step t. Therefore, we face a non-ideal state transition matrix,

$$A_\Delta = A + \Delta_t = \begin{pmatrix} a_{ss} + \Delta_s^t & a_{sl} - \Delta_s^t \\ a_{ls} - \Delta_l^t & a_{ll} + \Delta_l^t \end{pmatrix}, \tag{3}$$

where $\Delta_t = \begin{pmatrix} +\Delta_s^t & -\Delta_s^t \\ -\Delta_l^t & +\Delta_l^t \end{pmatrix}$ is the anomaly matrix at time t, and Δ_s^t and Δ_l^t are deviations in transition probabilities of states $\mathcal{S}$ and $\mathcal{L}$, respectively. We consider a time-invariant anomaly matrix (i.e., $\Delta_l^t = \Delta_l$ and $\Delta_s^t = \Delta_s$); however, extension to the time variant case is straightforward. By substituting A with A_Δ in (2), we obtain

$$\begin{aligned} &f\left(X_{i,t}|S\left(\hat{x}_{i,t-1}\right) = \mathcal{L}\right) \\ &\quad = (a_{ll} + \Delta_l)\mathcal{N}(\hat{x}_{i,t-1}, {\sigma_0}^2) + (a_{ls} - \Delta_l)\mathcal{N}(0, {\sigma_0}^2), \\ &f\left(X_{i,t}|S\left(\hat{x}_{i,t-1}\right) = \mathcal{S}\right) \\ &\quad = (a_{ss} + \Delta_s)\mathcal{N}(0, {\sigma_0}^2) + (a_{sl} - \Delta_s)\mathcal{N}(0, {\sigma_1}^2). \end{aligned} \tag{4}$$

We also define random variable $Z_t = [Z_{1,t}, Z_{2,t}, \ldots, Z_{n,t}]$ which represents the signal without the effects of anomalies and random variable $E_{\Delta_t} = [E_{\Delta_{1,t}}, E_{\Delta_{2,t}}, \ldots, E_{\Delta_{n,t}}]$ which represents the difference between the models with and without the effects of anomalies. Yet, $Z_{i,t}$ follows the distribution in (2).

The pdf of error can be found using $f(E_{\Delta t}) = f(X_{i,t} - Z_{i,t}) = f(-X_{i,t}) \oplus f(Z_{i,t})$ where $\oplus$ represents convolution

978-1-4244-9848-2/11$26.00©2011 IEEE

operation,

$$
\begin{aligned}
& f\left(E_{\Delta_{i,t}} | S\left(\hat{x}_{i,t-1}\right) = \mathcal{L}\right) \\
& (a_{ll}^2 + a_{ls}^2 + \Delta_l (a_{ll} - a_{ls})) \mathcal{N}(0, 2{\sigma_0}^2) \\
& \quad + (a_{ll} a_{ls} - \Delta_l a_{ll}) \mathcal{N}(\hat{x}_{i,t-1}, 2{\sigma_0}^2) \\
& \quad + (a_{ll} a_{ls} + \Delta_l a_{ls}) \mathcal{N}(-\hat{x}_{i,t-1}, 2{\sigma_0}^2), \\
& f\left(E_{\Delta_{i,t}} | S\left(\hat{x}_{i,t-1}\right) = \mathcal{S}\right) \\
& \quad = (2 a_{ss} a_{sl} + \Delta_s (a_{sl} - a_{ss})) \mathcal{N}(0, 2{\sigma_0}^2 + {\sigma_1}^2) \\
& \quad + (a_{ss}^2 + \Delta_s a_{ss}) \mathcal{N}(0, 2{\sigma_0}^2) \\
& \quad + (a_{sl}^2 + \Delta_s a_{sl}) \mathcal{N}(0, 2{\sigma_1}^2).
\end{aligned} \tag{5}
$$

Therefore, the mean, $E\{E_{\Delta t,i}\}$ of the error imposed by model mismatch is as follows,

$$
\begin{aligned}
E\{E_{\Delta t,i}\} = & \frac{k}{n}((a_{ll}^2 + a_{ls}^2 + \Delta_l (a_{ll} - a_{ls})) \times 0 \\
& + (a_{ll} a_{ls} - \Delta_l a_{ll}) \times \hat{x}_{i,t-1} \\
& + (a_{ll} a_{ls} + \Delta_l a_{ls}) \times (-\hat{x}_{i,t-1})) \\
& + (1 - \frac{k}{n})((2 a_{ss} a_{sl} + \Delta_s (a_{sl} - a_{ss})) \times 0 \\
& + (a_{ss}^2 + \Delta_s a_{ss}) \times 0 + (a_{sl}^2 + \Delta_s a_{sl}) \times 0) \\
= & - \frac{k}{n} \Delta_l \hat{x}_{i,t-1}
\end{aligned} \tag{6}
$$

As it can be seen from (6) the average error due to anomalies is directly proportional to the sparsity rate and Δ_l. We define a parameter ξ, which we call *model variation to mismatch ratio*, as $\xi_j = \frac{\Delta_j}{min(a_{js}, a_{jl})}$ for $j \in \{l, s\}$. We assume, this parameter is equal for all rows of a transition matrix (i.e. different states); therefore, expressing the model anomalies using $\xi = \xi_l = \xi_s$.

It is worth noting that this algorithm is robust to the error propagation. If any coefficient is initialized erroneously, after a few iterations with high probability it converges to its true value. More details on this issue are discussed in Section III.

2) Online model parameters estimation: The Markov model that is used in previous sections may not be priori known. In this case it is required to employ a learning process along with the signal reconstruction. In such cases, a training data set is required to capture the model. Even after training the algorithm, model parameters mismatch happens in the case of noisy measurements. Measurement noise prevents exact model parameters prediction using the decoded coefficients. Moreover, for non-stationary signals, model parameters gradually change in time. Therefore, in order to avoid anomalies described in Section II-C1, the model parameters should be periodically estimated using noisy measurements. The offline methods are not of our interest here because they require all decoded signal values in order to predict the model parameters. Consequently, we apply an online model parameter estimator such as the one introduced in [23], in our simulations. A *sequential expectation maximization (EM)* algorithm is adapted to estimate the parameters of Markov model (i.e. $\lambda = [A, \sigma_0, \sigma_1]$) sequentially. These sequential algorithms are derived based on maximizing the Kullback-Leibler (KL) information measure, $J(.)$. Given the true model is λ^0 the KL measure between the true model and any model, λ, is defined as $J(\lambda) = E_{\lambda^0}\{\log f(y_t|\lambda)\}$, where $E_{\lambda^0}\{.\}$ is expectation with respect to the true model. The EM algorithm can be summarized as follows,

$$
\lambda_t = \arg\max_{\lambda} E_{y_t, \Lambda_t}\{\log f(y_t|\lambda)\}, \tag{7}
$$

where $\Lambda_t = [\lambda_1, ..., \lambda_t]$. Using (7) the sequential algorithm which is presented in [23, Equations 3.24, 3.30 and 3.33-3.35] can be obtained. This process is fully described in [23].

We update the model parameters only at the end of CSBP iterations at every time step. According to this algorithm, the parameters of Markov model are estimated based on the reconstructed signal coefficient at the previous time steps. Therefore, the algorithm needs to store the previous values of signal coefficients to use them in model parameter estimation. The online parameter estimator architecture can be seen in Fig. 4. In Fig. 4, the results of signal reconstruction is used in estimating the model parameters A, σ_1 and σ_0 for the next time step.

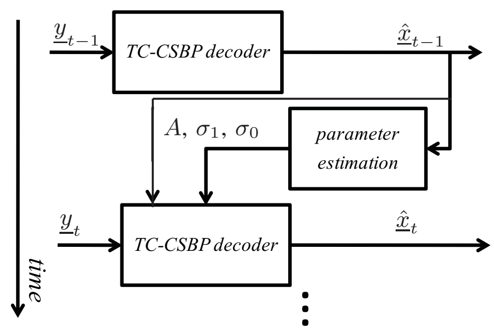


Fig. 4. Block diagram of the learning phase for model parameter estimation.

III. Simulation results

In this section, we evaluate our proposed TC-CSBP method through simulations. The model of Section II-A is fed to the proposed TC-CSBP algorithm with and without model mismatch, and then the results are compared with a model-ignorant CSBP algorithm. We also provide the results for comparing the performance of Modified-CS reconstruction algorithm [14] with TC-CSBP algorithm. The numerical simulation parameters are reported in Table I. The simulations of CSBP [24] and modified-CS [14] algorithms are performed using the MATLAB code that the developers provided online. The CSBP algorithm requires a sparse measurement matrix with fixed number of non-zero coefficients at each row and column of ϕ_t. Therefore, we assume a randomly generated measurement matrix with 20 ones at each row. We also assume the model is known at time step $t = 1$ (i.e. the model learning process is finished).

Fig. 5 depicts the ℓ_2-reconstruction error (LRE), which is defined as $\sqrt{\sum_{i=1}^{n} (x_{i,t} - \hat{x}_{i,t})^2}$, as a function of the number of collected measurements at each time step for CSBP and TC-CSBP with different ξ parameters. It can be seen that TC-CSBP outperforms the conventional CSBP algorithm. For example, for $\xi = 0$ (no model mismatch) and LRE= 30 TC-CSBP requires 400 measurements while CSBP requires

978-1-4244-9848-2/11$26.00©2011 IEEE

TABLE I
SIMULATION PARAMETERS

Parameter name	Value
n: Number of signal samples at each time step	1000
k: Number of coefficients of x_t in state $\mathcal{L}$	100
σ_1: Standard deviation of values in state $\mathcal{L}$	10
σ_0: Standard deviation of values in state $\mathcal{S}$	1
σ_z: Standard deviation of observation noise at decoder	1
A: Markov Model State transition matrix	$\begin{pmatrix} 0.989 & 0.011 \\ 0.1 & 0.9 \end{pmatrix}$

about 590 measurements. This means TC-CSBP results in 32% reduction in the number of required measurements. It can also be seen that the proposed algorithm is robust to model mismatch to a good extent. As seen, for $\xi < 0.5$ model parameter variations does not have any destructive effect on the algorithm's performance. On the other hand, it reveals that the proposed algorithm is misled by the model mismatch for relatively large variations in model parameters (i.e., $\xi > 0.5$). Fig. 6 illustrates the LRE versus the number of measurement

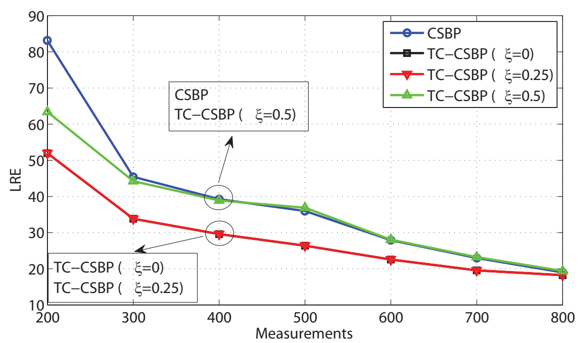


Fig. 5. Reconstruction LRE versus the number of measurements for TC-CSBP and CSBP algorithms (effect of model mismatch illustration).

with different signal time correlation parameters. For a highly variable signal, a considerable number of coefficients change their state (i.e., from $\mathcal{L}$ to $\mathcal{S}$ and vice versa) at each time step. As we can see from Fig. 6, by increasing the signal variability (i.e. decreasing a_{ll} and a_{ss}), the performance of TC-CSBP algorithm degrades. However, even for a highly variable signal that half of its coefficients at state $\mathcal{L}$ change their state at each time step, TC-CSBP is performs better than conventional CSBP.

In Fig. 5 and Fig. 6, the LRE performances of TC-CSBP and CSBP algorithms are measured at the time step $t = 10$.

For further comparisons, we have considered a scenario in which the support is only allowed to change at certain time steps. We observe the algorithm's performance over a 20 time step period and we assume a support change at time step $t = 10$. This change in support consists of adding new coefficients to and removing some existing coefficients from the support, totally 10% of support set changes in $t = 10$. In this setup, the signal coefficients in the state $\mathcal{S}$ are equal to zero and we assume noisy measurements. Using these

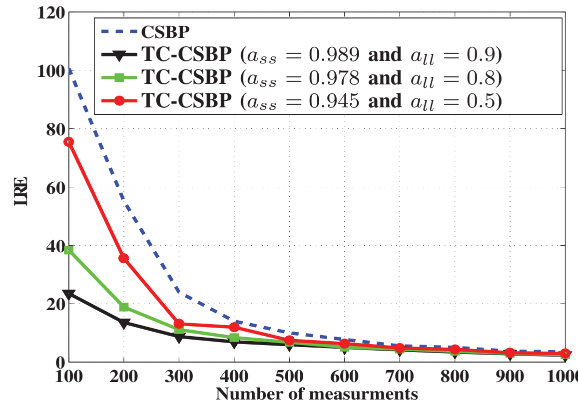


Fig. 6. Reconstruction LRE versus the number of measurements for TC-CSBP and CSBP algorithms at t=10, for different time correlation parameters.

assumptions, we can compare our algorithm with other studies in the literature [13–15]. Fig. 7 shows the reconstruction performances of TC-CSBP and Modified-CS [14] in time with two different values for the number of measurements. At $t = 0$, the reconstruction is performed without using the previously estimated model. The performances of the mentioned algorithms are compared when $m = 75$ and $m = 100$. As we can see, when the number of measurements is small ($m = 75$), TC-CSBP shows a considerable performance improvement, and when an abundance of measurements are available ($m = 100$), Modified-CS (which is based on ℓ_1-minimization) performs slightly better than TC-CSBP. This is explained by the fact that ℓ_1 reconstruction techniques perform better than CSBP when provided with too many measurements, which is not the case in many CS applications. Modified-CS is the latest version in the series of reconstruction algorithms (e.g. KF-CS [15] and LS-CS [13]) with partially known support.

The robustness of TC-CSBP to error propagation can also be seen from Fig. 7. At time step $t = 10$ the support of the signal changes. Consequently, a priori signal values from the previous time step are not valid for the changes in support. As we can see, these types of errors are compensated in a few time steps and do not propagate into the whole algorithm. The learning phase of the model parameters using the online estimator in Section II-C2 is shown in Fig. 8. We have depicted the absolute error in the estimated transition probabilities versus the number of time steps. It can be inferred from Fig. 8 that using the online estimator in Section II-C2 can compensate the effect of model parameter mismatch after a few time steps.

An interesting trade-off exists between the computational requirements for model parameter estimation and the number of required CS measurements for reaching a given performance. Different designs for this problem can be considered by a system engineer according to the limitations on computation or communication resources. For example, in a multi-hop sensor network, transmitting a symbol over the network is energy-wise more costly than extra computations in one node. On the other hand, implementing a complex decoding algorithm

978-1-4244-9848-2/11$26.00©2011 IEEE

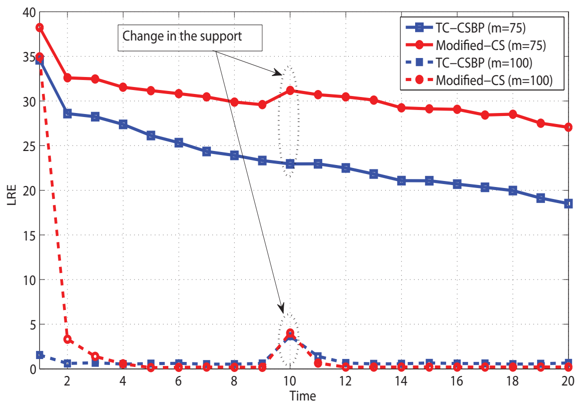


Fig. 7. Reconstruction LRE in time for the CSBP and Modified-CS algorithms. The signal support only changes at $t = 10$ and it is fixed in all other time slots. The results for two different measurement values (i.e. $m = 75$ and $m = 100$) are given. The simulation parameters that are different than those of Table I are $n = 250$, $k = 25$.

inside a sensor node requires employing a more expensive and more energy consuming CPU.

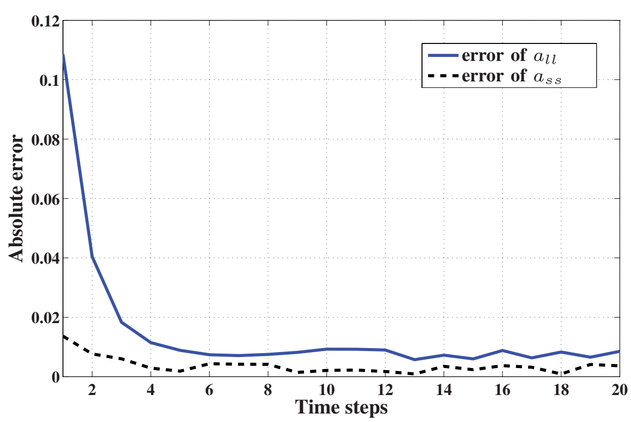


Fig. 8. Transition probabilities error versus number of times that TC-CSBP algorithm runs ($\xi = 1$).

IV. Conclusion

In this paper, we have proposed TC-CSBP, which is a compressive sensing reconstruction algorithm for sparse signals that are also time-correlated. TC-CSBP builds upon previous work on compressive sensing via belief propagation (CSBP) by Baron et al. We considered CSBP as our underlying reconstruction scheme due to its flexibility to consider signal model as a priori knowledge. Other CS reconstruction schemes mostly do not have this flexibility. In TC-CSBP, we consider the time-correlation model of signal as a priori knowledge and our results show a considerable improvement over conventional CSBP and other related work. Moreover, our results show that TC-CSBP is robust to the error in time-correlation model parameters to a great extent and it can maintain its supremacy in the presence of model mismatch. Further, an integration of online model estimation into TC-CSBP was studied for more accurate model estimation.

V. Acknowledgement

This material is based upon work partially supported by the National Science Foundation under Grant No. ECCS-1056065.

References

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

[2] R. Baraniuk, "Compressive sensing," *Lecture notes in IEEE Signal Processing magazine*, vol. 24, no. 4, pp. 118–120, 2007.

[3] E. Candès, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 1, Citeseer, 2006.

[4] E. Candes, M. Rudelson, T. Tao, and R. Vershynin, "Error correction via linear programming," in *Annual Symposiom on Foundations of Computer Science*, vol. 46, p. 295, IEEE Computer Society Press, 2005.

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

[6] E. Candes and T. Tao, "The Dantzig selector: statistical estimation when p is much larger than n," *Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

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

[8] M. Duarte, M. Wakin, and R. Baraniuk, "Fast reconstruction of piecewise smooth signals from incoherent projections," *SPARS05*.

[9] S. Sarvotham, D. Baron, and R. Baraniuk, "Compressed sensing reconstruction via belief propagation," *preprint*, 2006.

[10] D. Baron, S. Sarvotham, and R. Baraniuk, "Bayesian Compressive Sensing via Belief Propagation," *Arxiv preprint arXiv:0812.4627*, 2009.

[11] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982 –2001, 2010.

[12] C. J. Miosso, R. von Borries, M. Argàez, L. Velazquez, C. Quintero, and C. M. Potes, "Compressive sensing reconstruction with prior information by iteratively reweighted least-squares," *Trans. Sig. Proc.*, vol. 57, no. 6, pp. 2424–2431, 2009.

[13] N. Vaswani, "LS-CS-residual (LS-CS): Compressive sensing on least squares residual," *Signal Processing, IEEE Transactions on*, vol. 58, no. 8, pp. 4108 –4120, 2010.

[14] N. Vaswani, "Analyzing least squares and kalman filtered compressed sensing," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 3013 –3016, 2009.

[15] N. Vaswani, "Kalman filtered compressed sensing," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2008.

[16] M. Khajehnejad, W. Xu, S. Avestimehr, and B. Hassibi, "Weighted l1 minimization for sparse recovery with prior information," *Arxiv preprint arXiv:0901.2912*, 2009.

[17] D. Angelosante, G. B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Digital Signal Processing, 2009 16th International Conference on*, pp. 1–8, July 2009.

[18] V. Stankovic, L. Stankovic, and S. Cheng, "Sparse Signal Recovery With Side Information," 2009.

[19] D. Baron, M. Duarte, S. Sarvotham, M. Wakin, and R. Baraniuk, "An information-theoretic approach to distributed compressed sensing," in *Allerton Conf. Comm., Control, Comput*, Citeseer, 2005.

[20] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," 2005.

[21] S. Madden, "UC-Berkeley Intel Lab Data." http://db.csail.mit.edu/labdata/labdata.html, June 2004.

[22] F. Kschischang, B. Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on information theory*, vol. 47, no. 2, pp. 498–519, 2001.

[23] V. Krishnamurthy and J. Moore, "On-line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2557–2573, 1993.

[24] D. Baron and S. Sarvotham, "Compressive sensing via belief propagation software," December 2008. http://www.ece.rice.edu/ drorb/CSBP/.