DYNAMIC ZERO-POINT ATTRACTING PROJECTION FOR TIME-VARYING SPARSE SIGNAL RECOVERY

*Jiawei Zhou, Laming Chen, and Yuantao Gu**

State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, CHINA

ABSTRACT

Sparse signal recovery in the static case has been well studied under the framework of Compressive Sensing (CS), while in recent years more attention has also been paid to the dynamic case. In this paper, enlightened by the idea of modified-CS with partially known support, and based on a non-convex optimization approach, we propose the dynamic zero-point attracting projection (DZAP) algorithm to efficiently recover the slowly time-varying sparse signals. Benefiting from the temporal correlation within signal structures, plus an effective prediction method of the future signal based on previous recoveries incorporated, DZAP achieves high-precision recovery with less measurements or larger sparsity level, which is demonstrated by simulations on both synthetic and real data, accompanied by the comparison with other state-of-the-art reference algorithms.

Index Terms— Time-varying, sparse signal recovery, non-convex approach, exponential smoothing, dynamic zero-point attracting projection (DZAP).

1. INTRODUCTION

Compressive Sensing (CS) [1–3] utilizes the inherent sparsity feature of real-world signals and observes them with only a few random projections, making it possible to acquire signals below Nyquist rate while putting more efforts to the signal processing end. Previous works mainly focus on sparse signal recovery in the static setting, while in some applications time-dependent signals are more natural to be modeled as slowly time-varying sparse ones, e.g. real-time magnetic resonance imaging [4, 5] and channel equalization in communications [6].

In this paper, we study the problem of recovering a slowly time-varying sparse signal $\{\mathbf{s}_t \in \mathbb{R}^N, t = 1, 2, \dots\}$, with at most K ($K \ll N$) nonzero entries at each time instance t, from the linear measuring process

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

where $\{\mathbf{y}_t \in \mathbb{R}^{M_t}, t = 1, 2, \dots\}$ is the measurement series, $\mathbf{A}_t \in \mathbb{R}^{M_t \times N}$ is the sensing matrix, and $\mathbf{v}_t$ is the additive noise. Usually $M_t \ll N$, leading to an under-determined system of equations. we are interested in designing an efficient algorithm to recover the sparse signal series $\{\mathbf{s}_t\}$ from the measurement series $\{\mathbf{y}_t\}$. Here by saying "slowly time-varying", we mean that the amplitudes of nonzero entries do not change abruptly with time, and their locations may only change slightly at adjacent time instances. In one case, [7] shows that for a typical sequence of real-world MRI images in wavelet domain, the maximum support changes are less than 2% of the support size, and almost all signal value changes are less than 4% of the signal energy. Therefore, strong temporal correlations are inherent in the signal structure which could be potentially helpful in designing efficient processing schemes.

Much progress has already been made. To name a few, L1 Homotopy Dynamic Updating [8] moves the solution along the piecewise continuous homotopy path, using previous signal estimate as a starting point. DCS-AMP [9] is a representative of complex probabilistic modeling. Other MMV algorithms, such as Group Lasso and Fused Lasso [10], are also raised for the almost invariant support case. And in [11], the dynamic CS problems are instead tackled by reconstructing the difference signals. Namrata Vaswani et al. have shown that sparse recovery with partially known support yields better performance than traditional CS [12, 13], resulting in modified-CS, which could be easily fitting into the dynamic setting. This idea is also extended to modify some other algorithms in [14], but not specially aiming at slowly time-varying signals.

In this paper, we present a framework of dynamic sparse signal recovery using a non-convex optimization approach under the assumption of the very slowly changing pattern. More specifically, we extend one of this approach, called zero-point attracting projection (ZAP) [15], which is later generalized to the approximate projected generalized gradient (APGG) method [16], to the dynamic setting. Since ZAP adopts non-convex non-smooth function as the penalty, it uses the generalized gradient of the penalty function in the update procedure. When the penalty is convex, generalized gradient is commonly known as subgradient. By updating the sparsity penalty based on support approximation and incorporating signal prediction techniques combined with the algorithm setting, the dynamic ZAP (DZAP) can greatly lower the requirements for sampling rate and sparsity level for high recovery accuracy, meanwhile costing much less running time. In the following sections of this paper, we describe the DZAP algorithm for time-varying sparse signal recovery in detail, and use experimental results to demonstrate its validity and good performance in comparison with other state-of-the-art algorithms.

2. TIME-VARYING SIGNAL RECOVERY: DYNAMIC ZAP

2.1. Notation

Throughout this paper, we use boldface lowercase letters to denote vectors, and boldface uppercase letters to denote matrices. The subscript $_t$ is strictly used to indicate time instances in time-dependent quantities, whereas the superscript with parentheses, such as $^{(n)}$, is the signal element index, and the letter p indexes an iteration. There

This work was supported by National 973 Program of China (Grant No. 2013CB329203) and National Natural Science Foundation of China (NSFC 61371137). The corresponding author of this paper is Yuantao Gu (gyt@tsinghua.edu.cn).

are mainly three different notations of a signal: $\mathbf{s}_t$ denotes the true signal at time t, $\hat{\mathbf{s}}_t$ denotes the recovered signal, and $\mathbf{s}_t^*$ denotes predicted signal for time t based on previous recoveries. Finally, we use $supp(\cdot)$ to represent the support of a sparse signal, and for a signal that is not strictly sparse, we use $supp_{\beta\%}(\cdot)$ to denote the smallest set containing $\beta\%$ energy of the signal, called the "shrinkage support".

2.2. Basic Approach

Essentially different from the popular l_1 relaxation and greedy pursuits, our proposed method is based on a non-convex approach. For a time-varying signal, at time instance t, we aim to solve the following optimization problem:

$$\min J_t(\mathbf{s}), \quad \text{s.t.} \quad \mathbf{y}_t = \mathbf{A}_t\mathbf{s} \tag{2}$$

where $J_t(\cdot)$ is a non-convex penalty function serving as an approximation to the l_0-norm, and it is separable down to the sum of some sort of sparsity-inducing penalties on each signal element

$$J_t(\mathbf{s}) = \sum_{n=1}^{N} F_{\alpha_t^{(n)}}(\mathbf{s}^{(n)}). \tag{3}$$

There are many choices for the parameter controlled sparsity-inducing function $F_\alpha(\cdot)$ (see more details in [16]), and here we choose it as

$$F_\alpha(x) = \begin{cases} 2\alpha|x| - \alpha^2x^2 & |x| \le \dfrac{1}{\alpha}; \\ 1 & \text{elsewhere.} \end{cases} \tag{4}$$

To find the optimal solution to problem (2), ZAP does the following two steps in each iteration. First, the solution is updated along the negative gradient direction of the sparsity penalty

$$\tilde{\mathbf{s}}_t(p+1) = \hat{\mathbf{s}}_t(p) - \kappa_p \cdot \nabla J_t(\hat{\mathbf{s}}_t(p)) \tag{5}$$

where $\kappa_p > 0$ is the variant step size (decreasing at convergence) and $\nabla J_t(\cdot)$ is the generalized gradient of the penalty $J_t(\cdot)$. Then the update is projected back to the solution space of $\mathbf{y}_t = \mathbf{A}_t\mathbf{s}$

$$\hat{\mathbf{s}}_t(p+1) = \tilde{\mathbf{s}}_t(p+1) + \mathbf{A}_t^\dagger(\mathbf{y}_t - \mathbf{A}_t\tilde{\mathbf{s}}_t(p+1)) \tag{6}$$

where $\mathbf{A}_t^\dagger = \mathbf{A}_t^T(\mathbf{A}_t\mathbf{A}_t^T)^{-1}$ is the pseudo inverse matrix of $\mathbf{A}_t$. The static algorithm usually starts from the initialization position $\hat{\mathbf{s}}_t(0) = \mathbf{A}_t^\dagger\mathbf{y}_t$. For large scale problems where the accurate pseudo inverse matrix is difficult to derive, [16] shows that approximate calculation of it would also lead to favorable recovery performance with reduced computational complexity.

2.3. J-function update based on support approximation

The optimality of the solution to the problem (2) is heavily dependent on the non-convexity of the J-function, which can be purely measured by the parameter α. The larger α is, the closer the elemental sparsity-inducing penalty $F_\alpha(\cdot)$ is to the l_0 norm, thus attracting the corresponding signal element to zero-point more strongly in every iteration. However, not all the elements are those we want to reduce to zero, especially those inside the signal support.

In the slowly time-varying signal model, we suppose that there is, if existing, only a very small portion of support that changes every time, thus the following approximation has very high probability of success (given a recovery with high accuracy at time $t-1$):

$$supp(\mathbf{s}_t) \approxeq supp(\mathbf{s}_{t-1}) \approxeq supp_{\beta\%}(\hat{\mathbf{s}}_{t-1}) \tag{7}$$

Table 1. Outline of the DZAP Algorithm

Set parameters: α_{in}, α_{out}, κ_0, β, stopping criterion;
Input: $\{\mathbf{y}_t\}$, $\{\mathbf{A}_t\}$, $t = 1, 2, 3, \ldots$;
Output: $\{\hat{\mathbf{s}}_t\}$.
Initialization: $\hat{\mathbf{s}}_t(0) = \mathbf{A}_t^\dagger\mathbf{y}_t$ and $J_t(\cdot)$ for $t = 1$.
Repeat: (at time instance t; starting from $t = 1$)
Recover: Solve the problem (2):
Repeat: (at p^{th} iteration; initialize $p = 0$)
Zero-point attracting by (5);
Solution space projection by (6);
$p \to p + 1$;
Until: stopping criterion for iteration satisfied;
Output: recovered signal $\hat{\mathbf{s}}_t$;
Sparsity penalty function update: Adjust $J_t(\cdot)$ by (8);
Signal prediction: Compute $\mathbf{s}_t^*$ by (9), set $\hat{\mathbf{s}}_{t+1}(0) = \mathbf{s}_t^*$;
Move forward: $t \to t + 1$;
Until: t is the last time instance.

where $supp_{\beta\%}(\cdot)$ is adopted since $\hat{\mathbf{s}}_{t-1}$ may not be strictly sparse. According to the support approximation, we update the J-function at every time instance by setting different values for $\alpha_t^{(n)}$ for signal elements depending on where they locate

$$\alpha_t^{(n)} = \begin{cases} \alpha_{in} & n \subseteq supp_{\beta\%}(\hat{\mathbf{s}}_{t-1}) \\ \alpha_{out} & n \nsubseteq supp_{\beta\%}(\hat{\mathbf{s}}_{t-1}) \end{cases} \tag{8}$$

where α_{in} and α_{out} are constants satisfying $0 \le \alpha_{in} < \alpha_{out}$, which means more emphasis on the pursuit of sparsity outside of support. The algorithm can tolerate small errors in support approximation because of the projection procedure, and the parameters α_{in}, α_{out}, and β are chosen empirically by considering the approximation quality, iterative initialization, and signal energy comprehensively.

2.4. Signal Prediction for Initialization Setting

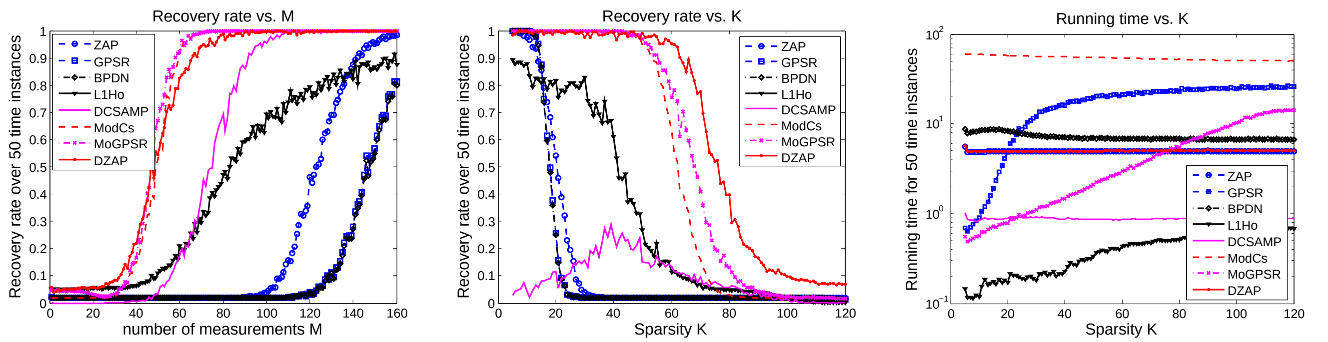
The initialization setting $\hat{\mathbf{s}}_t(0)$ for the ZAP iteration is very important in avoiding a local minima and guaranteeing the algorithm convergence for sparsity penalties with different non-convexities [16], especially when with not adequate measurements. For every recovery, a closer initialization position to the true signal than $\mathbf{A}_t^\dagger\mathbf{y}_t$ is desirable for better performance, and that's where we incorporate the signal prediction strategy in the dynamic case.

To begin with, we assume that the signal energy stays stable, and the degree of "closeness" (measured in correlation) to the signal $\mathbf{s}_t$ is decreasing in the order of $\mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \ldots, \mathbf{s}_1$, meaning that the more recent a past signal is in time, the more it can tell about the current signal, which is also justified in our slowly time-varying signal model. Therefore, we use the exponential smoothing method to make a prediction of the current signal from previous recoveries:

$$\mathbf{s}_t^* = \begin{cases} 0 & \text{for} \quad t = 1 \\ \gamma\,\hat{\mathbf{s}}_{t-1} + (1-\gamma)\,\mathbf{s}_{t-1}^* & \text{for} \quad t \ge 2 \end{cases} \quad (0 < \gamma \le 1) \tag{9}$$

More clearly, if we expand the recursion and rewrite (9) as

$$\begin{aligned} \mathbf{s}_t^* &= \gamma\,\hat{\mathbf{s}}_{t-1} + (1-\gamma)\,\mathbf{s}_{t-1}^* \\ &= \gamma\,\hat{\mathbf{s}}_{t-1} + (1-\gamma)[\gamma\,\hat{\mathbf{s}}_{t-2} + (1-\gamma)\,\mathbf{s}_{t-2}^*] = \cdots \\ &= \gamma\,\hat{\mathbf{s}}_{t-1} + \gamma(1-\gamma)\,\hat{\mathbf{s}}_{t-2} + \cdots + \gamma(1-\gamma)^{t-2}\,\hat{\mathbf{s}}_1 \end{aligned} \tag{10}$$



(a) (b) (c)

Fig. 1. Results for the simulations on synthetic data. (a) Average rates of successful recovery over 50 time instances vs. the number of measurements. (b) Average rates of successful recovery over 50 time instances vs. signal sparsity. (c) Average running time for recovering a 50 time instance signal vs. signal sparsity.

we can see the prediction equation is actually a weighted average, with the weights for older recoveries decrease at an exponential rate, which fits our signal closeness assumption. Note that by predicting in this way, we do not suppose any trend or properties such as periodicity of the signal are known in advance. The recursion (9) merely makes the computation efficient.

The parameter γ is very crucial since it controls the weights distribution: a larger γ places more emphasis on latest recovery, thus can adjust the prediction to signal changes more quickly, and a smaller γ cares more about the whole period, thus can better dampen out prediction fluctuations introduced by recovery noises. For the case where the signal follows a time-invariant changing pattern, we take a further step to propose an adaptively γ tuning method, based on the prediction error $PE_t = \frac{\|\mathbf{s}_t^* - \hat{\mathbf{s}}_t\|_2}{\|\hat{\mathbf{s}}_t\|_2}$, adjusting γ by

$$\gamma = \gamma_t = \begin{cases} 1 & \text{if} \quad PE_{t-1} > 1 \\ PE_{t-1} & \text{if} \quad PE_{t-1} \in (0, 1] \\ 0.0001 & \text{if} \quad PE_{t-1} = 0 \end{cases} \tag{11}$$

and the initialization is $\gamma_2 = 1$. In other words, we rely on the previous prediction quality to guide on our decision for the next time.

After getting the signal prediction, we initialize the recovery procedure with $\hat{\mathbf{s}}_t(0) = \mathbf{s}_t^*$. Combined with the J-function updating scheme based on support approximation, we develop the dynamic ZAP (DZAP) algorithm, which is summarized in TABLE 1.

3. SIMULATION RESULTS

3.1. Synthetic Data

We start by testing our approach on synthetic data. The synthetic time-varying signal has 50 time instances, following the slowly changing pattern: a random number of non-zeros change with time in low-frequency sinusoidal curves (others stay unchanged), and the support changes (only happen when an entry approaches to zero) are limited less than 10% of the support size. The signal length is fixed at $N = 1000$, and the sparsity is kept as $K = 30$. The sensing matrix $\mathbf{A}_t$ has entries following *i.i.d.* $\mathcal{N}(0, 1/M_t)$ distribution, and the white noise $\mathbf{v}_t$ is also Gaussian, fixing MSNR (defined as $10\lg(\frac{\|\mathbf{A}_t\mathbf{s}_t\|^2}{\|\mathbf{v}_t\|^2})$) at 30dB. Because of no prior information for the first time instance, we set $\mathbf{A}_1$ taller and $\mathbf{A}_2$ stays the same for the following time instances.

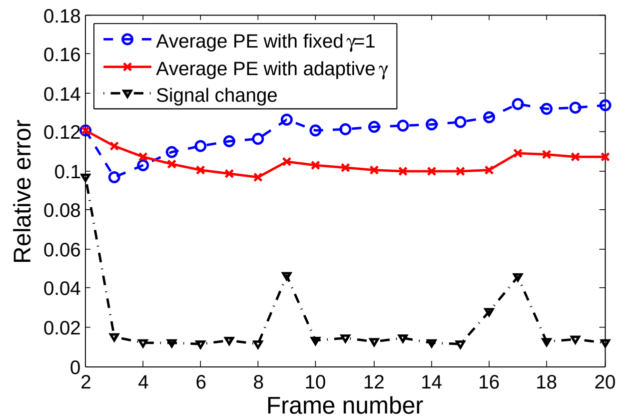


Fig. 2. Average prediction error $PE_t = \|\mathbf{s}_t^* - \mathbf{s}_t\|_2/\|\mathbf{s}_t\|_2$ of DZAP in recovering a cardiac image sequence, and the changing error of the true signal, computed by $\|\mathbf{s}_t - \mathbf{s}_{t-1}\|_2/\|\mathbf{s}_t\|_2$. Note that the red line also illustrates the average time-varying γ in adaptive prediction method, whose value is affected by the signal changing pattern.

Two numerical simulations with the above signal model are implemented. The algorithms used for comparison include static ZAP [15], static BPDN [17,18], static GPSR_BB [19], L1 Homotopy Dynamic Update [8], EM-DCS-AMP [9], ModCs [12], and a modified version (with the idea of modified-CS) of GPSR_BB (coined MoGPSR_BB). The parameter settings of DZAP are: $\alpha_{in} = 0$, $\alpha_{out} = 10$, initialized step size $\kappa_0 = 5 \times 10^{-4}$, and $\beta = 99.9$. In both simulations, we do not apply adaptive γ in the prediction step, instead, we simply set $\gamma = 1$ ($\mathbf{s}_t^* = \hat{\mathbf{s}}_{t-1}$). We evaluate the reconstruction error by mean square deviation defined as $\text{MSD} = \frac{\text{E}\|\hat{\mathbf{s}}_t - \mathbf{s}_t\|^2}{\text{E}\|\mathbf{s}_t\|^2}$.

In the first experiment, we test the algorithm performance versus the number of measurements. We set $M_1 = 200$, M_2 ranging from 1 to 160 to run the recovery algorithms, and for each M_2 we repeat 200 times. Fig. 1(a) shows the average successful recovery rate within 50 time instances (we claim a "successful recovery" at time t if its MSD is below 0.01), from which we can see ModCs, MoGPSR_BB, and DZAP perform the best and are comparable to each other.

In the second experiment, we test the algorithm performance

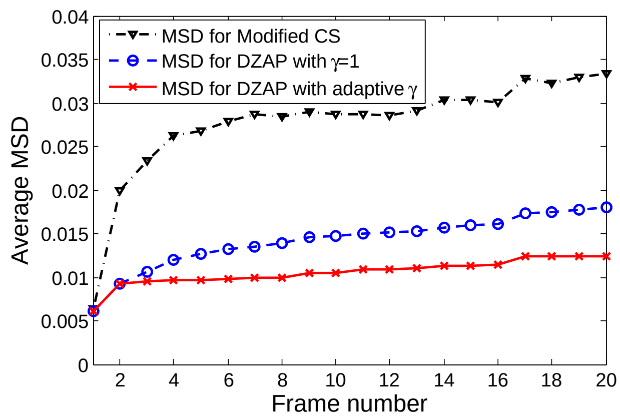


Fig. 3. Average MSD of recovering a 20-frame cardiac image sequence. DZAP with adaptive exponential smoothing prediction works the best.

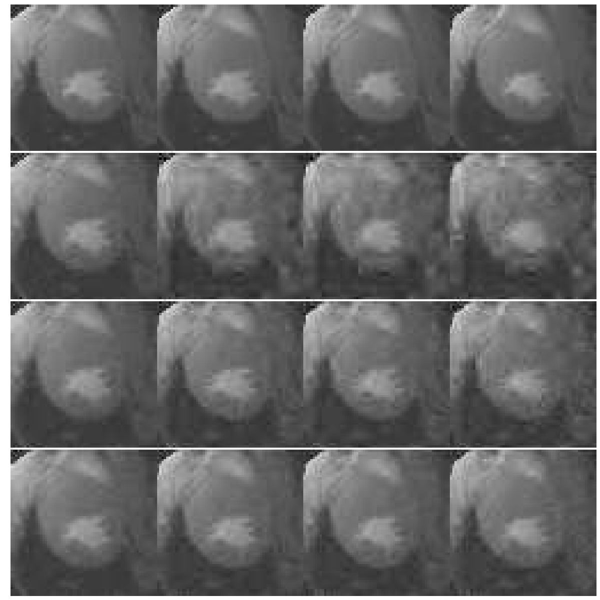


Fig. 4. Frames $1, 4, 10, 20$ of the dynamic 64×64 MRI cardiac image sequence. From top to bottom: original image, recovered image by modified-CS, recovered image by DZAP with $\gamma = 1$, and recovered image by DZAP with adaptive γ.

versus the signal sparsity. We fix $M_1 = 400$ (to avoid the performance downgrades produced by unfit initialization at first time instance), $M_2 = 100$, and change the sparsity level K from 5 to 120, again running 200 times for each K. The recovery rates for different algorithms are shown in Fig. 1(b), in which the proposed DZAP algorithm outperforms all the others.

Fig. 1(c) shows the average running time in the second experiment. Here we should note that although the approximate calculation of the pseudo inverse of sensing matrix $\mathbf{A}_t$ in DZAP algorithm can reduce computational complexity [16], we do not apply it in this simulation to make the speed comparison clear. As can be seen, both EM-DCS-AMP and L1 Homotopy run very fast but are not tolerating more sparsity, and among the three algorithms allowing the largest sparsity levels, ModCs is too slow to implement in practice; DZAP consumes a moderate amount of time that is not affected by sparsity, while the time needed by MoGPSR_BB is exponentially increasing with the scale of problem. Thus, DZAP can be seen as a great balance between performance and complexity.

3.2. Real Data

Finally, we test the proposed algorithm on a dynamic MRI image sequence, as well as validating the effectiveness of the adoption of adaptive exponential smoothing for signal prediction. The 20-frame cardiac images are firstly obtained from Namrata Vaswani's website (*http://www.ece.iastate.edu/ namrata/*) and then down sampled to the size of 64×64, and they are compressible in the two-level Daubechies-4 2D discrete wavelet transform (DWT) domain. The corresponding 1D signal has length $N = 4096$, and we take Gaussian random measurements of the DWT coefficients with down sampling rate 0.5 at the first frame and 0.2 for the following frames. We run DZAP with $\gamma = 1$ in the prediction step, DZAP with adaptive γ, and modified-CS as a benchmark to recover the signals, each for 10 times with different sensing matrices. Here we set $\alpha_{in} = 0$, $\alpha_{out} = 1$, $\kappa_0 = 5 \times 10^{-4}$, and $\beta = 99.9$.

To demonstrate the performance of our adaptive γ tuning method incorporated in the signal prediction procedure, we plot the average prediction error for adaptive γ (note that this is also the time-varying curve of γ according to (11)) and for $\gamma = 1$ as a comparison in Fig. 2. We also illustrate the signal changes there. It is seen that the signal has quite abrupt changes at frames 2, 9, and 17, where the adaptive γ adjust itself to a higher level to better follow the signal, and at other frames it smoothly lower the prediction error. In contrast, the prediction is becoming worse with frame number increasing for the setting of fixed $\gamma = 1$, which is not so much of a prediction since actually $\mathbf{s}_t^* = \hat{\mathbf{s}}_{t-1}$.

As for the recovery performances, we plot the average MSD for the testing algorithms in Fig. 3. With the close initialization positions provided by prediction, clearly DZAP with adaptive γ tuning method works the best in terms of recovery accuracy. Fig. 4 shows the four frames of the original cardiac images and the recovered ones, from which we can see that the images recovered by modified-CS are contaminated with relatively higher noises at later frames (the second line), while DZAP with adaptive γ performances better in visual quality (the bottom line). In addition, the average running time it takes for DZAP for one sequence recovery is around 3 minutes, whereas modified-CS takes about 2 hours.

4. CONCLUSION

In this paper, we propose an efficient online algorithm DZAP for recovering slowly time-varying sparse signals, which is an extension of the recently proposed non-convex approach from static to dynamic scenarios. As the sparse pattern changes very slowly with time, DZAP utilizes the additional information provided by the previously recovered signals, with the mechanisms of sparsity penalty update based on support approximation, and incorporated with signal prediction techniques to aid the algorithm. In comparison with previously proposed dynamic CS algorithms, experiments on both synthetic and real data attest to the effectiveness of the adaptive prediction method, as well as showing that DZAP needs fewer measurements, allows larger sparsity level, and runs at a moderate speed.

5. REFERENCES

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[2] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[3] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Information Theory*, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.

[4] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 3, pp. 1182-1195, 2007.

[5] W. Lu and N. Vaswani, "Modified compressive sensing for real-time dynamic MR imaging," in *Proc. IEEE Int'l Conf. Image Proc. (ICIP)*, pp. 488-492, 2009.

[6] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Communications*, vol. 50, no. 3, pp. 374-377, 2002.

[7] W. Lu and N. Vaswani, "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *IEEE Trans. Signal Processing*, vol. 60, no. 1, pp. 182-196, Jan. 2012.

[8] M. S. Asif and J. Romberg, "Dynamic updating for sparse time varying signals," in *43rd Annual Conference on Information Sciences and Systems (CISS)*, pp. 3-8, Mar. 2009.

[9] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," arXiv:1205.4080v2 [cs.IT]. Available: http://arxiv.org/abs/1205.4080.

[10] D. Angelosante, G. B. Glannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Int'l Conf. on Digital Signal Processing*, pp. 3-8, Jul. 2009.

[11] V. Cevher, A. Sankaranarayanan, M. Duarte, D. Reddy, R. Baraniuk, and R. Chellappa, "Compressive sensing for background subtraction," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, pp. 155-168, 2008.

[12] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Processing*, vol. 58, pp. 4595-4607, Sept. 2010.

[13] W. Lu and N. Vaswani, "Modified BPDN for noisy compressive sensing with partially known support," in *Proc. IEEE Int'l Conf. Acoustics, Speech, Signal Process (ICASSP)*, pp. 3926-3929, 2010.

[14] R. E. Carillo, L. F. Ploania, and K. E. Barner, "Iterative algorithms for compressed sensing with partially known support," in *Proc. IEEE Int'l Conf. Acoustics, Speech, Signal Process (ICASSP)*, pp. 3654-3657, 2010.

[15] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409-420, Apr. 2010.

[16] L. Chen and Y. Gu, "The convergence guarantees of a non-convex approach for sparse recovery," *IEEE Trans. Signal Processing*, vol. 62, no. 15, pp. 3754-3767, Aug. 2014.

[17] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33-61, Aug. 1998.

[18] E. Candès and J. Romberg, l1-magic: A collection of matlab routines for solving the convex optimization programs central to compressive sensing. Available: www.acm.caltech.edu/l1magic.

[19] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586-597, Dec. 2007.