Gridless Sparse Reconstruction for the Cyclic Autocorrelation Estimation

Xushan Chen, Xiongwei Zhang, Jibin Yang*, Lin Qiao

Lab of Intelligence Information Processing, PLA University of Science and Technology, Nanjing China
*Email: yjbice@sina.com

Abstract—In order to exploit the inherent cyclostationary properties which vary periodically in most man-made signals, one prerequisite is the knowledge of the signal's cyclic autocorrelation (CA) which can be estimated from finite time-domain samples. In this paper we concern about the sparse, periodic CA estimation and focus on recovering the CA using compressive sampling, i.e. a small amount of time-domain samples. Inspired by atomic norm based technology we model the CA estimation as a denoising problem with atomic norm, or equivalently an atomic norm soft thresholding (AST) problem, and propose a gridless version CA reconstruction which can locate the nonzero cyclic frequencies on an infinitely dense grid. The consequent convex optimization problem can be solved using semidefinite programming (SDP) via Alternating Direction Method of Multipliers (ADMM) in polynomial time. Numerical results demonstrate that the proposed method outperforms the traditional methods as well as the dictionary based CA estimator in terms of the mean square error (MSE) over a wide range of signal to noise ratios (SNR) case.

Keywords—Cyclic autocorrelation; Atomic norm; Atomic soft thresholding; Compressive sampling; ADMM

I. INTRODUCTION

The inherent physical mechanism that generates the random signal makes the parameters of the signal would not vary with time, and thus many traditional signal processing schemes deal with random signals just like they were statistically stationary. Cyclostationary [1] which generally existed in most man-made signals describes some statistical characteristics that vary periodically with time in communication system. The inherent periodicity can be caused by sinusoidal carries, periodic keying, spreading sequences, and so on. Recognizing and exploiting the underlying periodicity can bring to performance improvements of the signal processors.

Given a cyclostationay random process $x(t)$, its autocorrelation $R_x(t, \tau)$ can be expanded in a Fourier series

$$R_x(t, \tau) = \sum_\alpha R_x^\alpha(\tau) e^{j2\pi\alpha t}$$

where $\alpha = k/T_0$, $k \in \mathbb{Z}$ and the Fourier coefficients, referred to as cyclic autocorrelation functions [2], are given by

$$R_x^\alpha(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t+\tau/2) x^*(t-\tau/2) e^{-j2\pi\alpha t} dt$$

The nonzero CA coefficients with respect to α are considered as cycle frequencies. The CA of purely stationary signals exhibits spectral peaks only at the cycle frequency $\alpha = 0$, while for cyclostationay signals $R_x^\alpha(\tau) \neq 0$ at specific frequencies, $\alpha = k/T_0$, $k \in \mathbb{Z}$ [3]. Since the support of CA in a cyclostationary signal is zero except $\alpha = 0$ and the set of cycle frequency as well as its harmonics, the CA can be regarded as sparse. Therefore the sparsity of the CA can be taken fully into account for the goal of recovering it with compressive time-domain samples.

Recently compressive sensing (CS) [4-6] has attracted considerable attention by suggesting that it may be possible to overcome the traditional limits of sampling theory. CS depends on the fundamental assumption of sparsity that many signals can be represented by only a few nonzero entries in a suitable basis or dictionary. Then nonlinear optimization is able to reconstruct such signals from compressive measurements regardless of the locations of the nonzero entries.

For a BPSK signal with symbol length T_b, the CA's theoretical expression is shown in [7]. Enframed by a rectangular transmission window the value of the CA reduces to

$$\tilde{R}_x^\alpha(\tau) = \begin{cases} \dfrac{|T_b - \tau|}{T_b} \dfrac{\sin(\pi\alpha(T_b - \tau))}{\pi\alpha(T_b - \tau)} & \alpha = \dfrac{k}{T_b}, k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Compressive sensing based reconstruction which applied to estimate the CA in the cycle frequency domain is investigated in [8]. This paper proposes a new blind estimation method that only take advantage of a short observation duration to estimate the CA. Compared to the traditional estimator in [1], the proposed method tends to outperform in term of estimation error evaluated by the mean square error (MSE).

Given multiple prior structure information of the CA, a more detailed model is developed in [7] which can achieve higher reconstruction accuracy but require less time-domain samples, i.e. a shorter observation time. The desired CA vector is divided into a sum of three solution vectors and a dictionary is constructed to describe the structure of the CA including the fundamental cycle frequency and its harmonics. Simulation results show that the additional knowledge of the CA structure

leads to a better MSE performance compared to the basis CS approach in [9].

However in this kind of methods, it is inevitable that the continuous cycle frequency domain is discretized into a finite set of grid points, since CS has been focused on signals that can be expressed sparsely under a finite dictionary, e.g. DFT basis. Assume that the true cycle frequencies are close to some grid points, it is intuitively reasonable that a dense grid leads to a more accurate cycle frequency estimation due to not only grid mismatch but approximation errors can be reduced with a dense grid [10]. However in the light of the behavior of CS the sampling grid should not be too dense, or the significant coherence between adjacent atoms may degrade the sparse recovery performance. The dimension of the finite dictionary reflects the tradeoff. Therefore, one naturally wonders whether the CA can be estimated with an infinitely dense grid and, if implementable, what performance the gridless version CA estimation can obtain.

In this paper we propose a new approach to estimate the CA from compressive sampling based on an atomic norm denoising theory. The corresponding convex optimization problem can be solved by semidefinite programming (SDP) via Alternating Direction Method of Multipliers (ADMM). Therefore we provide the update formulations of ADMM in closed form which can solve the SDP efficiently. We also compare our algorithm with the traditional CA estimation algorithm in [1] and the CS dictionary based approach presented in [7]. Experimental results indicate that the atomic norm based CA estimator outperforms in terms of MSE over a wide range of SNR case. Moreover, the proposed estimator performs superior in some practical cases, such as non-integer period of cycle frequencies, or some cycle frequencies of multiple signals are separated by few spacing.

II. TRADITIONAL AND CS BASED CYCLIC AUTOCORRELATION ESTIMATION MODEL

A. Traditional CA Estimation Model

Consider that the time-domain signal $x(t)$ is sampled uniformly with a sampling period T_s, so that for the sample vector $\mathbf{x} \in \mathbb{C}^N$ we have

$$\mathbf{x} = \left[x(0), x(T_s), \cdots, x((N-1)T_s) \right]^T$$

The traditional unbiased CA estimation is given by

$$\hat{R}_x^\alpha(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x(nT_s) x^*(nT_s + \tau) e^{-j2\pi\alpha\tau}$$

where $\alpha = i/NT_s$, $i = 0, \cdots, N-1$ and the delay τ is fixed to be an integer multiple of the sampling interval T_s. Since the CA is a generalization of the autocorrelation function, in which a cyclic weighting factor $e^{-j2\pi\alpha\tau}$ is included before the time averaging is carried out. Thus denote $\mathbf{F}$ as the discrete Fourier transform matrix and the delay-product $\mathbf{y}_{\tau_0} = \mathbf{x} \odot \mathbf{x}_{\tau_0}$, where $\odot$ denotes the point-wise multiplication, the CA vector can then be described as

$$\hat{\mathbf{R}}_x(\tau_0) = \frac{1}{N} \mathbf{F} \mathbf{y}_{\tau_0} \tag{2}$$

where $\hat{\mathbf{R}}_x(\tau_0) = \left[\hat{R}_x^0(\tau_0), \cdots, \hat{R}_x^{(N-1)/NT_s}(\tau_0) \right]^T$.

According to [1], the CA's absolute value of a BPSK signal at corresponding cycle frequencies can be obtained from Eq. (1) for a theoretical reference

$$\tilde{\mathbf{R}}_x(\tau_0) = \begin{cases} \left| \tilde{R}_x^{i/NT_s}(\tau_0) \right| & i = 0, \cdots, \dfrac{N}{2} \\[2mm] \left| \tilde{R}_x^{(i-N)/NT_s}(\tau_0) \right| & i = \dfrac{N}{2}+1, \cdots, N-1 \end{cases} \tag{3}$$

B. Dictionary based CA Estimation Model

In order to significant reduce the observation time, the CS based estimation presented in [7] can only use the first M samples of $\mathbf{y}_{\tau_0}$, $M \ll N$, to recover the CA vector. It follows from Eq. (2) that the delay-product $\mathbf{y}_{\tau_0}$ satisfies $\mathbf{y}_{\tau_0} = N\mathbf{F}^{-1}\hat{\mathbf{R}}_x(\tau_0)$, where $\mathbf{F}^{-1}$ is the inverse discrete Fourier transform matrix. The notation $\bar{\mathbf{y}}_{\tau_0}$ is introduced for the first M known elements of $\mathbf{y}_{\tau_0}$ such that the underdetermined system of equations can be expressed as

$$\bar{\mathbf{y}}_{\tau_0} = \mathbf{M}\mathbf{y}_{\tau_0} = N\mathbf{M}\mathbf{F}^{-1}\hat{\mathbf{R}}_x(\tau_0) \tag{4}$$

where $\mathbf{M} \in \mathbb{R}^{M \times N}$ contains the first M rows of the $N \times N$ identity matrix.

In fact we do know something about the CA's feature, so that the CA vector is divided into three parts, that is, $\hat{\mathbf{R}}_x(\tau_0) = \mathbf{e} + \mathbf{f} + \mathbf{g}$ [7], where the vector $\mathbf{e}$ denotes the conventional autocorrelation function that contains the DC component, $\mathbf{f}$ represents the different amplitudes of the nonzero CA coefficients and $\mathbf{g}$ describes the complex-valued characteristics of the CA vector.

The key of the method is to construct a dictionary $\mathbf{D}$ describing the structure of the CA vector. Each column of $\mathbf{D}$ represents one of the possible cycle frequencies. Therefore the overall cycle frequencies can be modeled with $\mathbf{D}$. However the dictionary base model do not take the noise corrupted measurements into account, and the continuous cycle frequency domain is discretized into finite points by the dictionary.

III. GRIDLESS CA RECONSTRUCTION AND OPTIMIZATION PROBLEM

A. Atomic Norm [11]

Atomic norm generalizes the notion of ℓ_1 norm and the nuclear norm. Let $conv(\mathcal{A})$ is the convex hull of points in $\mathcal{A}$ and the atoms in $\mathcal{A}$ is compact, centrally symmetric and contains the origin as an interior point. The atomic norm $\|\cdot\|_{\mathcal{A}}$ is defined as the Minkowski function associate with $conv(\mathcal{A})$

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf\{t > 0 \mid \mathbf{x} \in t \cdot conv(\mathcal{A})\}$$

$$= \inf\left\{\sum_k c_k : \mathbf{x} = \sum_k c_k \mathbf{a}_k, c_k \geq 0, \mathbf{a}_k \in \mathcal{A}\right\}$$

When $\mathcal{A}$ is the set of unit norm 1-sparse elements in $\mathbb{C}^N$, the atomic norm $\|\cdot\|_{\mathcal{A}}$ is ℓ_1 norm. Similarly, when $\mathcal{A}$ is the set of unit norm rank-1 matrices, the atomic norm is the nuclear norm. The atomic norm $\|\cdot\|_{\mathcal{A}}$ can be computed via SDP

$$\|\mathbf{x}\|_{\mathcal{A}} = \min_{t,\mathbf{u}} \frac{1}{2}(t + u_1)$$

$$s.t. \quad \begin{bmatrix} T(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^* & t \end{bmatrix} \geq 0$$

where $\mathbf{u} \in \mathbb{C}^M$ and the map $T : \mathbb{C}^N \to \mathbb{C}^{N \times N}$ which denotes a Hermitian Toeplitz matrix with

$$T(\mathbf{u}) = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \\ u_2^* & u_1 & \cdots & u_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_N^* & u_{N-1}^* & \cdots & u_1 \end{bmatrix}$$

where u_j denotes the jth element of $\mathbf{u}$.

The atomic norm based approach is able to recover the incomplete or missing samples and determine the unknown frequencies precisely in a continuous domain that can be widely applied in matrix completion [12], direction of arrival estimation [13,14], mobile communication [15], radar imaging [16] and so on.

B. Gridless CA Reconstruction

Consider the problem of recovering a vector $\hat{\mathbf{R}}_x(\tau_0)$ from the noisy measurement $\bar{\mathbf{y}}_n \in \mathbb{C}^M$, for gridless CA estimation, the resulting convex optimization problem is given by

$$\min \|\mathbf{G}\hat{\mathbf{R}}_x(\tau_0)\|_{\mathcal{A}} \quad s.t. \quad \mathbf{G}(\hat{\mathbf{R}}_x(\tau_0) + \mathbf{n}) = \bar{\mathbf{y}}_n \quad (5)$$

where $\mathbf{G} = N\mathbf{M}\mathbf{F}^{-1}$, $\mathbf{n}$ is independently and identically distributed zero-mean Gaussian noise with noise variance σ^2. It is known a prior that $\hat{\mathbf{R}}_x(\tau_0)$ can be expressed as a linear combination of a few atom from $\mathcal{A}$. Therefore the way to estimate $\hat{\mathbf{R}}_x(\tau_0)$ from the observation $\bar{\mathbf{y}}_n$ can be characterized as

$$\min_{\hat{\mathbf{R}}_x(\tau_0)} \frac{1}{2}\|\bar{\mathbf{y}}_n - \mathbf{G}\hat{\mathbf{R}}_x(\tau_0)\|_2^2 + \tau\|\hat{\mathbf{R}}_x(\tau_0)\|_1$$

where τ is a regularization parameter.

The sparse recovery performance depends on the finite grid size of dictionary, i.e. $\mathbf{G}$, while gridless reconstruction considers $\mathbf{G}\hat{\mathbf{R}}_x(\tau_0)$ as a whole and then can eliminate the gridded influence. Therefore, the AST method for estimating the noiseless CA vector is listed below

$$\min_{\mathbf{p}} \frac{1}{2}\|\bar{\mathbf{y}}_n - \mathbf{p}\|_2^2 + \tau\|\mathbf{p}\|_{\mathcal{A}} \quad (6)$$

where $\mathbf{p} = \mathbf{G}\hat{\mathbf{R}}_x(\tau_0)$. It can be formulated by

$$\min_{t,\mathbf{u},\mathbf{p},\mathbf{Z}} \frac{1}{2}\|\bar{\mathbf{y}}_n - \mathbf{p}\|_2^2 + \frac{\tau}{2}(t + u_1)$$

$$s.t. \quad \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{p} \\ \mathbf{p}^* & t \end{bmatrix} \geq 0 \quad (7)$$

C. Solving via Alternating Direction Method of Multipliers

The dual atomic norm is given by

$$\|\mathbf{z}\|_{\mathcal{A}}^* = \sup_{\|\mathbf{x}\|_{\mathcal{A}} \leq 1} \langle \mathbf{x}, \mathbf{z} \rangle$$

There exists a $\mathbf{z}$ that achieve the above supremum for any $\mathbf{x}$. Since $\mathcal{A}$ contains all external points of $\{\mathbf{x} : \|\mathbf{x}\|_{\mathcal{A}} \leq 1\}$, it is guaranteed that the optimal solution will actually lie in the set $\mathcal{A}$ [17]

$$\|\mathbf{z}\|_{\mathcal{A}}^* = \sup_{\mathbf{a} \in \mathcal{A}} \langle \mathbf{a}, \mathbf{z} \rangle$$

Assume $\mathcal{S} \subset \mathcal{A}$, $\hat{\mathbf{z}}$ is a solution of the dual problem satisfying

$$\begin{cases} |\langle \hat{\mathbf{z}}, \mathbf{a} \rangle| = \tau & \mathbf{a} \in \mathcal{S} \\ |\langle \hat{\mathbf{z}}, \mathbf{a} \rangle| < \tau & \mathbf{a} \notin \mathcal{S} \end{cases}$$

The dual norm will play an important role that the dual solution $\hat{\mathbf{z}}$ can determine the support of $\hat{\mathbf{x}}$ by deciding whether the absolute value of the inner product reaches τ.

According to Eq. (7), [10] provides a reasonably fast method for solving the SDP via ADMM. Therefore, we dualize the equality constraint via an Augmented Lagrangian

$$L_\rho(t, \mathbf{u}, \mathbf{p}, \mathbf{Z}, \mathbf{\Lambda}) = \frac{1}{2}\|\bar{\mathbf{y}}_n - \mathbf{p}\|_2^2 + \frac{\tau}{2}(t + u_1)$$

$$+ \left\langle \mathbf{\Lambda}, \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{p} \\ \mathbf{p}^* & t \end{bmatrix} \right\rangle + \frac{\rho}{2}\left\| \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{p} \\ \mathbf{p}^* & t \end{bmatrix} \right\|_F^2 \quad (8)$$

We introduce the partitions

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_0 & \mathbf{z}_1 \\ \mathbf{z}_1^* & Z_{n+1,n+1} \end{bmatrix} \quad \mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_0 & \mathbf{\lambda}_1 \\ \mathbf{\lambda}_1^* & \Lambda_{n+1,n+1} \end{bmatrix}$$

Let T^* denote the adjoint of the map T and $\mathbf{e}_1$ is constructed by an $N \times 1$ vector with only its first element being 1 and all other elements being 0. Then ADMM can be solved by the following update steps

$$\mathbf{p}^{l+1} = \frac{1}{2\rho+1}\left(\bar{\mathbf{y}} + 2\rho\mathbf{z}_1^l + 2\rho\boldsymbol{\lambda}_1^l\right)$$

$$\mathbf{u}^{l+1} = \mathbf{W}\left(T^*\left(\mathbf{Z}_0^l + \boldsymbol{\Lambda}_0^l/\rho\right) - \frac{\tau}{2\rho}\mathbf{e}_1\right)$$

$$t^{l+1} = Z_{n+1,n+1}^l + \left(\Lambda_{n+1,n+1}^l - \frac{\tau}{2}\right)\Big/\rho$$

$$\mathbf{Z}^{l+1} = \arg\min_{\mathbf{Z}\geq 0}\left\|\mathbf{Z} - \begin{bmatrix} T(\mathbf{u}^{l+1}) & \mathbf{p}^{l+1} \\ \mathbf{p}^{l+1*} & t^{l+1} \end{bmatrix} + \boldsymbol{\Lambda}^l/\rho\right\|_F^2$$

$$\boldsymbol{\Lambda}^{l+1} = \boldsymbol{\Lambda}^l + \rho\left(\mathbf{Z}^{l+1} - \begin{bmatrix} T(\mathbf{u}^{l+1}) & \mathbf{p}^{l+1} \\ \mathbf{p}^{l+1*} & t^{l+1} \end{bmatrix}\right)$$

where $\mathbf{W}$ is an $N \times N$ diagonal matrix with its (i,i) th diagonal element being

$$\mathbf{W}_{ii} = \begin{cases} 1/n & i = 1 \\ 1/2(n-i+1) & i > 1 \end{cases}$$

After we obtain the estimate $\hat{\mathbf{p}}$, we also obtain the solution $\hat{\mathbf{q}}$ to the dual problem as $\hat{\mathbf{q}} = \bar{\mathbf{y}} - \hat{\mathbf{p}}$. The dual solution $\hat{\mathbf{q}}$ can determine the support if and only if

$$\left|\sum_{l=0}^{M-1}\hat{\mathbf{q}}_l e^{-j2\pi l f_l}\right| = \tau \qquad (9)$$

In other words, f_l is considered as the support when the maximum absolute value of the dual polynomial achieves magnitude τ. Fig. 1 shows the polynomial value obtained by Eq. (9) with $M = 160$ samples and $k = 6$ randomly chosen frequencies where the SNR is 10dB.
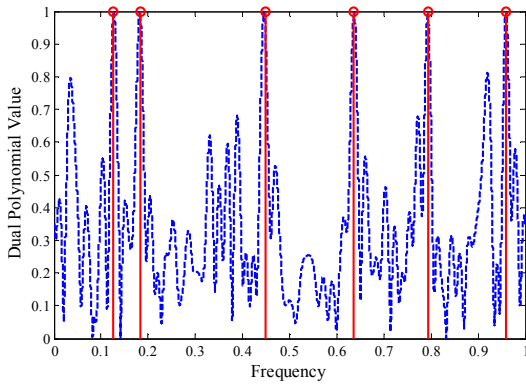

Figure 1. The dashed curve is the dual polynomial value which can be used to determine the frequency localization, and the actual location of the frequencies is shown in solid curve.

As soon as $\hat{\mathbf{p}}$ is obtained, we acquire an optimal solution of the cyclic autocorrelation estimation $\mathbf{R}_{opt}$ by solving

$$\min_{\hat{\mathbf{R}}_x(\tau_0)}\left\|\hat{\mathbf{p}} - \mathbf{G}\hat{\mathbf{R}}_x(\tau_0)\right\|_2^2 .$$

IV. NUMERICAL EXPERIMENTS AND ANALYSIS

The goal of this section is to evaluate the reconstruction properties of proposed method obtained by performing recovery in different cases of various SNR and length of samples, respectively.

A. Experiments on Recovery Performance

We compare the MSE of the corresponding estimated CA vectors with the theoretical curve described by Eq. (3) to demonstrate the reconstruction performance. The parameters are given in Table 1.

TABLE 1.   **SCENARIO PARAMETERS**

Parameter	Symbol	Value
Size of the CA vector	N	1000
# of the measurements	M	{120,160,200,240,280,320}
Time delay	τ_0	$2T_s$
BPSK symbol length	T_b	$10T_s$
Signal to noise ratio	SNR	{-10,-5,0,5,10}dB
Regularization parameters	β_a	0.25
	β_b	1

The overall MSE and the MSE of the support are defined as follows

$$\text{MSE}_{\text{overall}} = \frac{1}{N}\sum_{i=0}^{N-1}\left(\left|\hat{\mathbf{R}}_x(\tau_0)\right| - \tilde{\mathbf{R}}_x(\tau_0)\right)^2$$

$$\text{MSE}_{\text{peak}} = \frac{1}{N_{\text{peaks}}}\sum_{\mathcal{S}}\left(\left|\hat{\mathbf{R}}_x(\tau_0)\right| - \tilde{\mathbf{R}}_x(\tau_0)\right)^2 \qquad (10)$$

Reference [10] indicates that the choice of the regularization parameter depends on the noise model and the optimal choice for white Gaussian noise samples is to set τ equal to an upper bound on the expected dual atomic norm

$$\tau = \sigma\left(1 + \frac{1}{\log n}\right)\sqrt{n\log n + n\log(4\pi\log n)}$$

Fig. 2 shows the overall MSE and the MSE at the peaks arising from applying the three above-mentioned methods, namely the traditional CA estimation (trad.), the dictionary based approach (dict.) and our atomic norm based gridless CA estimator (atom.). The traditional approach requires N samples to compute the delay product vector $\mathbf{y}_{\tau_0}$ for estimating the CA vector $\hat{\mathbf{R}}_x(\tau_0)$ of length N, while the other two approach can recover the CA vector from the first $M \ll N$ elements of $\mathbf{y}_{\tau_0}$. As can be seen in Fig. 2 that the gridless approach can reach even exceed the performance of the traditional estimator with a small amount of time-domain samples. The prior knowledge which exploits sparsity in the CA vector and the atomic norm denoising algorithm lead to a better MSE performance compared to the dictionary based estimator.

Fig. 3 shows the recovery properties of the three estimators over a range of additional white Gaussian noise. It can be seen that the influence of the sparsity assumption improves the reconstruction performance in the noisy case. Although the

traditional takes advantage of much more samples, the gridless estimator perform best regardless of the overall MSE and the MSE at the peaks.
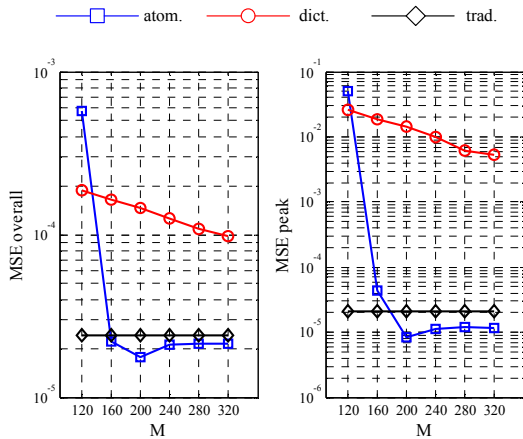

Figure 2. Reconstruction performance evaluated by the overall MSE (left) and peaks MSE (right) with SNR = 10dB. The number of measurements M in traditional estimation are in accordance with size of the CA vector N, while in the other two methods are from 120 to 320.
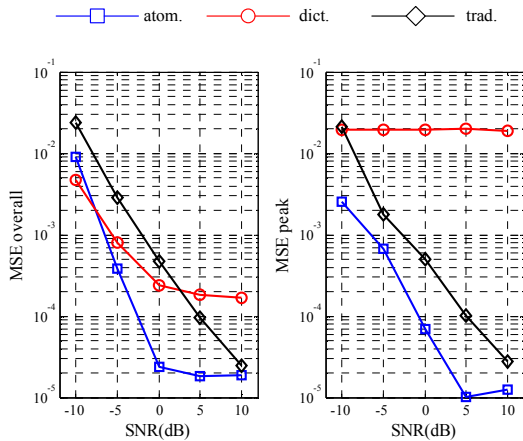

Figure 3. Reconstruction performance evaluated by the overall MSE (left) and peaks MSE (right) with a range of SNR from -10dB to 10dB. The number of measurements M in traditional estimation are in accordance with size of the CA vector N, while in the other two methods are 160.

B. Determining CA vector on an infinitely dense grid

Atomic norm generalizes the notion of ℓ_1 norm and the nuclear norm. Let $conv(\mathcal{A})$ is the convex hull of points in $\mathcal{A}$ and the atoms in $\mathcal{A}$ is compact, centrally symmetric and

The atomic norm based estimator can successfully recover the sparse CA vector from a few time-domain samples without constructing a gridded dictionary. So we concern about the CA support localization which simulated by recovering randomly and uniformly chosen frequencies.

Fig. 4 indicates the estimated frequencies can locate close to the true ones, which provide numerical evidence supporting the feasibility of our estimator, details of the true and estimated frequencies (bold) can be seen in Table 2. Since the estimator does not take the discretization level of the dictionary into account, the support of CA vector can be recovered accurately

even if the period of the cycle frequency is not an integer. The estimator can also be applied to identify multiple sets of cycle frequencies where some pair of them are separated by a small interval.

TABLE 2. COMPARISON BETWEEN TRUE FREQUENCIES AND RECOVERY FREQUENCIES

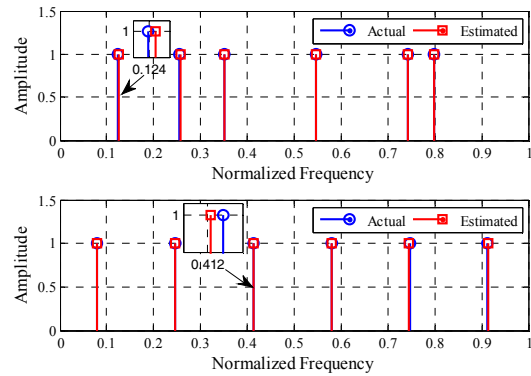| True (random)	Recovery (random)	True (uniform)	Recovery (uniform)
0.1238	**0.1248**	0.0797	**0.0801**
0.2556	**0.2562**	0.2464	**0.2462**
0.3506	**0.3504**	0.4130	**0.4122**
0.5473	**0.5468**	0.5797	**0.5802**
0.7419	**0.7422**	0.7464	**0.7456**
0.7986	**0.7982**	0.9130	**0.9134**

Figure 4. Cycle frequency localization for k randomly (above) and uniformly (below) chosen frequency with $N = 1000$, $M = 200$, $k = 6$, SNR = 10dB. The circles represent the true frequencies, while the squares represent the estimated frequencies.

V. CONCLUSION AND FUTURE WORK

The motivation of exploiting the sparsity of the CA vector is to reduce the amount of the time-domain samples. The main contribution of this paper is to model the CA estimation as a denoising problem inspired by the atomic norm based technology, and propose a gridless CA estimator which can recovery the CA vector with an infinitely dense grid. Interesting aspects of future research include investigating how to jointly estimate d-dimensional $(d \geq 2)$ off-the-grid frequencies by applying the atomic norm minimization [18,19].

REFERENCES

[1] W. Gardner. *Exploitation of spectral redundancy in cyclostationary signal* [J]. Signal Processing Magazine, IEEE, 1991, 8(2): 14-36.

[2] W. Gardner. *Spectral correlation of modulated signals: Part I – analog modulation* [J]. Communications, IEEE Transactions on, 1987, 35(6): 584-594.

[3] W. Gardner, W. Brown and K. Chen. *Spectral correlation of modulated signals: Part II – digital modualtion* [J]. Communications, IEEE Transactions on, 1987, 35(6): 595-601.

[4] R. Baraniuk. *Compressive sensing [lecture notes]* [J]. Signal Processing Magazine, IEEE, 2007, 24(4):118-121.

[5] E. Candes and M. Wakin. *An Introduction to compressive sampling* [J]. Signal Processing Magazine, 2008, 25(3): 21-30.

[6] D. Donoho. *Compressed sensing* [J]. Information Theory, IEEE Transactions on, 2006, 52(4): 1289-1306.

[7] A. Bollig, R. Mathar. *Dictionary-based reconstruction of the cyclic autocorrelation via l_1-minimization for cyclostationary spectrum sensing* [C]// Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, Vancouver, Canada, 4908-4912, 2013.

[8] Z. Tian. *Cyclic feature based wideband spectrum sensing using compressive sampling* [C]// Communications (ICC), 2011 IEEE International Conference on, Kyoto, Japan, 1-5, 2011.

[9] Z. Khalaf and J. Palicot. Exploiting sparse property of the cylic autocorrelation function for cyclostationary process compressed sensing [C]// WSR, Kalsruhe, Germany, 2012.

[10] B. Bhaskar, G. Tang and B. Recht. *Atomic norm denoising with applications to line spectral estimation* [J]. Signal Processing, IEEE Transactions on, 2013, 61(23): 5987-5999.

[11] V. Chandrasekaran, B. Recht, P. Parrilo and A. Willsky. *The convex geometry of linear inverse problems* [J]. Foundation Computational Mathematics, 12: 805-849, 2012.

[12] Z. Yang and L. Xie. *Continuous compressed sensing with a single or multiple measurement vectors* [C]// Statistical Signal Processing (SSP), 2014 IEEE Workshop on, Gold Coast, Australia, 308-311, 2014.

[13] Z. Yang, L. Xie and C. Zhang. *A discretization-free sparse and parametric approach for linear array signal processing* [J]. Signal Processing, IEEE Transactions on, 62(19):4959-4973, 2013.

[14] M. Ibrahim, F. Romer, R. Alieiev, G. Galdo and R. Thoma. *On the estimation of grid offsets in CS-based direction-of-arrival estimation* [C]// Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, Florence, Italy, 6776-6780, 2014.

[15] K. Ranganathan. Enabling off-the-grid telephony: An adaptive probabilistic model for broadcasting in ad-hoc mobile phone networks [J]. Simulation Modelling Practice and Theory, 49: 73-84, 2014.

[16] Z. Lu, R. Ying, S. Jiang, P. Liu and W. Yu. *Distributed compressed sensing off the grid* [J]. Signal Processing Letters, IEEE, 22(1):105-109, 2014.

[17] G. Tang, B. Bhaskar, P. Shah and B. Recht. *Compressed sensing off the grid* [J]. Information Theory, IEEE Transactions on, 59(11): 7465-7490, 2013.

[18] Y. Chi, Y. Chen. *Compressive recovery of 2-D off-gird frequencies* [C]// Signals, Systems and Computers, IEEE Asilomar Conference on, Pacific Grove, CA, USA, 687-691, 2013.

[19] W. Xu, J. Cai, K. Mishra, M. Cho and A. Kruger. Precise semidefinite programming formulation of atomic norm minimization for recovering D-dimensional ($D \geq 2$) off-the-grid frequencies [C]// Information Theory and Applications (ITA), IEEE Workshop on, 1-4, 2014.
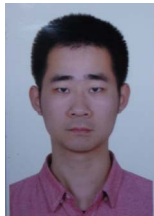
Xushan Chen was born in Jiangsu Province, China, 1987. He received the Bachelor degree in network engineering from PLA University of science and technology, China, in 2009. He received the Master degree in system engineering from PLA University of science and technology, China, in 2012.

He is currently a Ph.D. student in computer science and technology with Lab of Intelligence Information Processing, PLA University of science and technology. His major research interests include compressive sensing, analog to information converter, and sparse signal processing.

Xiongwei Zhang was born in Zhejiang Province, China, 1965. He received the B.S., Master and Ph.D. degrees in 1987, 1990 and 1993, respectively, from PLA University of science and technology, China, both in electrical engineering.

He is currently a full professor with PLA University of science and technology. His research interests include speech signal processing, image processing, and telecommunication systems.

Jibin Yang was born in Anhui Province, China, 1978. He is currently an associate professor with PLA University of science and technology. His research interest include speech signal processing, image processing, and telecommunication systems. His current research interest include DSP algorithms and multimedia signal processing.

Lin Qiao was born in Shanxi Province, China, 1991. He received the Bachelor degree in communication engineering from Xi'an University of posts and telecommunication. He is currently a master student with PLA University of science and technology. His research interests include sparse signal recovery algorithms and DSP application algorithms.