ROBUST OFF-GRID RECOVERY FROM COMPRESSED MEASUREMENTS

*Xinyue Shen** *Justin Romberg*† *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
†School of Electrical and Computer Engineering
Georgia Institute of Technology Atlanta, Georgia 30332, USA

ABSTRACT

In this paper, the robust off-grid recovery of the compressed signals with atomic norm-regularized least-squares problem is studied. The aim of the recovery is to reconstruct the original signal and to detect its off-grid support set. The general optimality conditions for the solution to this problem and its dual problem are proposed and discussed. A method based on dual certification to detect the support set is introduced and proved to be effective. As a specific case, the target signal is further assumed to have unknown line spectrum. Then the problem is also an estimation of a low dimensional subspace which is indexed by continuous parameters, yet the dimension itself is unknown. Under these presumptions, the squared-error of the reconstruction is derived. Finally, numerical experiments are demonstrated in such case to validate the effectiveness of the method and the plausibility of the theory.

Index Terms— Sparse recovery, robust signal reconstruction, off-grid support detection, atomic norm, line spectra detection.

1. INTRODUCTION

The reconstruction of a sparse signal and the detection of its support set from compressed measurements is a fundamental issue in the thriving realm of compressed sensing [1]. The sparse recovery problem has lately been unified in a general form by the model of atomic norm and atom set [2, 3]. The atomic norm itself, which is heuristically the tightest convex description of the cardinality of the support set, is not a newly invented idea [4]. The best part of it lies in that the sparse domain could be parameterized by continuous factors, hence it facilitates the off-grid support detection which is preferred in some high-resolution-required situations. Among all the important works in this general framework [5, 6], the work [7] has drawn some essential conclusions on the abstract denoising problem with atomic norm, despite that the compression is not in their main consideration.

Without compression, significant results have also been reached in the line spectra estimation and the recovery of the corresponding time-domain signal [8–12]. A similar problem has been named as super resolution [13, 14], in which the higher end of the spectrum is estimated from the samples at the lower end of it, and the estimation could be infinitely accurate given that the lines in the spectrum are separated far enough. In the work [7], the problem of robust line spectra detection is dealt as a special case of their general atomic norm minimization problem.

In the framework of compressive subspace parametric estimation [15] [16, Chap 4], the target signal is assumed to be in an unknown low dimensional subspace indexed by continuous parameters, and the purpose is to find such embedding subspace and reconstruct the target, yet the dimension of the subspace is often pre-known.

1.1. Problem Formulation

To begin with, suppose that the target signal $\mathbf{x}^* \in \mathbb{C}^N$ admits a linear combination of a few atoms from a given atom set.

$$\mathbf{x}^* = \sum_{k=1}^{s} c_k \mathbf{a}_k, \tag{1}$$

where $c_k \in \mathbb{R}_+$, $\mathbf{a}_k \in \mathcal{A}, k = 1, \cdots, s$, and $\mathcal{A}$ is a compact set in $\mathbb{C}^N$. What we observed is a non-adaptively compressed and contaminated measurement

$$\mathbf{y} = \Phi\mathbf{x}^* + \mathbf{w}, \tag{2}$$

where $\mathbf{w} \in \mathbb{C}^N$ is an additive noise, and $\Phi \in \mathbb{R}^{MN}, M < N$ is a known sensing matrix. In order to reconstruct $\mathbf{x}^*$ from $\mathbf{y}$, an atomic norm-regularized least-squares optimization problem is formulated as

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \Phi\mathbf{x}\|_2^2 + \tau\|\mathbf{x}\|_{\mathcal{A}}, \tag{3}$$

where τ is a chosen regularization parameter. For the definition of $\|\cdot\|_{\mathcal{A}}$, please refer to [2, Section 2.1].

Furthermore, the specific case of reconstructing line-spectrum signal could be well fitted into the above framework. In this scenario, the target $\mathbf{x}^*$ admits a decomposition in (1) with $\mathcal{A}$ defined as:

$$\mathcal{A} = \{\mathbf{a}_\omega \in \mathbb{C}^N : a_{\omega n} = \mathrm{e}^{\mathrm{j}\omega n}, \omega \in [0, 2\pi]\}, \tag{4}$$

and $\mathbf{a}_k = \mathbf{a}_{\omega_k}$. Such $\mathbf{x}^*$ could also be regarded as a vector lying in a low dimensional subspace, which is parameterized by an unknown $\omega = [\omega_1, \cdots, \omega_s]^{\mathrm{T}} \in \Omega$, where $\Omega = [0, 2\pi]^s$, and s is also not known in advance.

This work aims to recover $\mathbf{x}^*$ from $\mathbf{y}$ from two aspects: to reconstruct the original signal and to estimate its support in an off-grid manner, i.e. the positions of the elements in the support set are not assumed to be on any fixed grid.

*This work was partially supported by the National Program on Key Basic Research Project (973 Program 2013CB329201), the National Natural Science Foundation of China (NSFC 61371137) and the Innovation Foundation of CAST. The corresponding author is Yuantao Gu (gyt@tsinghua.edu.cn).

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

1.2. Main Contribution

The first contribution in this paper is the general optimality conditions for problem (3). In addition, a general off-grid support detection method based on dual certification is proposed. These results have underlying connections to [7], but it should be emphasized that the basic difference lies in that the compression is utilized in this work. The second contribution is the squared reconstruction error in the problem of line-spectrum signal recovery. We find that the compressing matrix brings in new challenges which are partly dealt with the restricted isometry property with respect to a low dimensional subspace [16, Chap 4] [17]. Further more, the performance of the proposed general support detection method is verified in numerical experiments together with the signal reconstruction method in the line-spectrum case.

2. GENERAL OPTIMALITY CONDITIONS AND SUPPORT DETECTION

In this section, the general problem of robust compressed signal reconstruction and off-grid support detection based on atomic norm-regularized least-squares minimization are studied.

To begin with, a lemma states the optimality conditions for problem (3).

Lemma 1. $\hat{\mathbf{x}}$ *is the solution to problem* (3)*, if and only if the conditions below hold simultaneously*

1. $\|\Phi^{\mathrm{T}}(\mathbf{y}-\Phi\hat{\mathbf{x}})\|_{\mathcal{A}}^{*} \leq \tau$;
2. $\langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi\hat{\mathbf{x}}\rangle = \tau\|\hat{\mathbf{x}}\|_{\mathcal{A}}$.

Proof. Please see 6.1. □

From (3) one may readily accept that τ plays an important role in balancing the two penalties of reconstruction error and sparsity. If τ tends to be 0, then $\Phi\hat{\mathbf{x}}$ will tend to be close to $\mathbf{y}$, which is preferred when the noise level is relatively low. Otherwise, $\Phi\hat{\mathbf{x}}$ is allowed to be far from $\mathbf{y}$, which brings benefit to the heavy-noise case. To be more concrete, Lemma 1 describes the quantitative relation between the regularization parameter τ and the reconstructed $\hat{\mathbf{x}}$.

The next lemma claims the dual problem of problem (3).

Lemma 2. *The dual problem of problem* (3) *is*

$$\max_{\lambda} \frac{1}{2}\|\lambda(\Phi\Phi^{\mathrm{T}})^{-1}\Phi\|_2^2 + \lambda^{\mathrm{T}}\mathbf{z}, \tag{5}$$

$$\text{s.t.} \quad \|\lambda\|_{\mathcal{A}}^{*} \leq \tau, \lambda - \Phi^{\mathrm{T}}\mathbf{y} = \Phi^{\mathrm{T}}\Phi\mathbf{z}. \tag{6}$$

Proof. Please see 6.2. □

Suppose that $\hat{\lambda}$ is the solution to problem (5). From the strong duality which the problem holds, it could be easily verified that if $\hat{\mathbf{x}}$ is the solution to (3), and $\hat{\lambda}$ satisfies that

$$\hat{\lambda} = \Phi^{\mathrm{T}}(\mathbf{y}-\Phi\hat{\mathbf{x}}), \tag{7}$$

then $\hat{\lambda}$ is the solution to problem (5). Equation (7) could be used to get the needed $\hat{\lambda}$ in the next proposition, in which an off-grid support detection method is proposed based on those two lemmas above.

Proposition 1. *Suppose that $\hat{\lambda}$ is the solution to problem* (5) *. If for each element in $\mathcal{A}$ and a subset $\mathcal{S} \subset \mathcal{A}$ satisfy that*

$$\langle\hat{\lambda}, \mathbf{a}\rangle \begin{cases} = \tau, & \forall \mathbf{a} \in \mathcal{S}; \\ < \tau, & \forall \mathbf{a} \notin \mathcal{S}, \end{cases} \tag{8}$$

then $\mathcal{S}$ is the support set of the solution to problem (3).

Proof. Please see Appendix 6.3. □

The proposition suggests that the support set could be estimated by finding all the maximums of the inner products of $\hat{\lambda}$ and all the potential atoms. If the atoms are indexed by continuous parameters which belong to a compact set, then the support set could be detected by identifying a continuous function's all maximums, which are guaranteed to be reached in that the domain of the function is compact. Therefore, an off-grid support detection is realized.

One may notice that a similar proposition is addressed in [7]. However, the essential difference is that our proposition actually works with the compressed signals.

3. RECOVERY ERRORS OF LINE-SPECTRUM SIGNAL

In this section, the above general framework is applied to the problem of line-spectrum signal recovery. The core result includes a squared reconstruction error. The impact of the compressing matrix can not be neglected, and is partly represented by the restricted isometry property with respect to the underlying subspace.

Theorem 1. *Suppose that the target signal $\mathbf{x}^* \in \mathbb{C}^N$ has line spectrum, i.e., it could be decomposed as* (1) *with $\mathcal{A}$ specifically defined in* (4)*. $\hat{\mathbf{x}}$ is the solution to problem* (3) *in which $\mathbf{y}$ is obtained by* (2)*, $\Phi \in \mathbb{R}^{MN}$, and $\Phi_{ij} \sim \mathcal{N}(0, 1/M)$ are* i.i.d.*. If the regularization parameter τ is not less than $\|\Phi^{\mathrm{T}}\mathbf{w}\|_{\mathcal{A}}^{*}$, then*

$$\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2^2 \leq \frac{2\tau}{1-\delta}\|\mathbf{x}^*\|_{\mathcal{A}}, \tag{9}$$

$$P\left\{\delta > C\left(\sqrt{\frac{st}{M}} + \frac{s\log(2s)t}{M}\right)\right\} \leq s(2\pi)^s \mathrm{e}^{s-t}, \tag{10}$$

where t is a positive factor and C is a constant.

Proof. Please see 6.4. □

There are several points to be discussed in the theorem. Firstly, the reconstruction error is bounded, and it increases linearly with the tightest convex relaxation of the support set's cardinality.

Secondly, compared with Lemma 1, the theorem provides more detailed relation between τ and the reconstruction error. When τ grows from $\|\Phi^{\mathrm{T}}\mathbf{w}\|_{\mathcal{A}}^{*}$, initially the error increases linearly with τ, and then $\hat{\mathbf{x}}$ approaches $\mathbf{0}$ when τ tends infinity. When τ equals to $\|\Phi^{\mathrm{T}}\mathbf{w}\|_{\mathcal{A}}^{*}$, the bound reaches its optimal value. As for the case when τ increases from 0 to $\|\Phi^{\mathrm{T}}\mathbf{w}\|_{\mathcal{A}}^{*}$, the penalty function is changing from the non-regularized one to the optimally regularized one, thus the error diminishes. Such prediction will be verified in the experiments, yet it is not in the theoretical results.

Thirdly, since δ needs to be less than 1 with positive probability, t should be chosen such that $t \sim s$, and M should be larger than $s^2 \log s$. Such requirement is not as ideal as the famous $M > s\log(N/s)$ in the "on-grid" compressed sensing [18], but our method and result are actually dealing with the off-grid situation where the well-known incoherent condition is not available. The term $1-\delta$ indicates the negative effect of the compression. The randomness is caused by the compressing matrix, and the theorem states that δ will have a higher probability to be relatively small, if s is smaller or M is larger. Since δ delineates the restricted isometry property of the compression with respect to the sparse domain, the bound in (9) is tight. Compared with the result on the reconstruction error in [7], our theorem indicates that if the compressing matrix is set properly, then the error of the compressed recovery could have a high probability to be as small as the error of the non-compressed recovery problem.

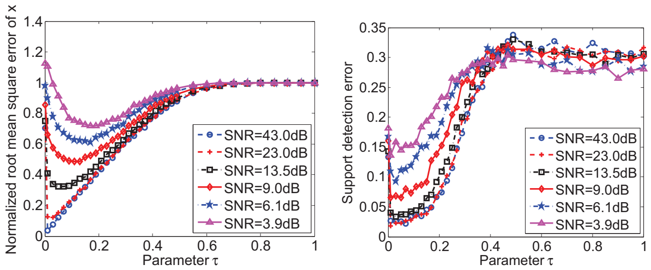

Fig. 1. The reconstruction error and the support detection error deteriorate as the measurement SNR decreases. The length of $\mathbf{x}^*$ is 64, the number of observations is 30, and the cardinality of the spectrum is 2. The SNR varies from 3.93dB to 43.01dB, and τ is tested from 0 to 1. 500 trials are simulated for each point.

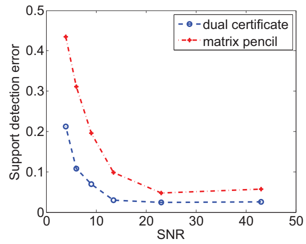

Fig. 2. Support detection error of the proposed method and the matrix pencil method. Parameter settings are the same as the first experiment, and τ is selected according to the lowest point in each curve in the left sub-figure of Fig.1. 500 trials are simulated for each point.

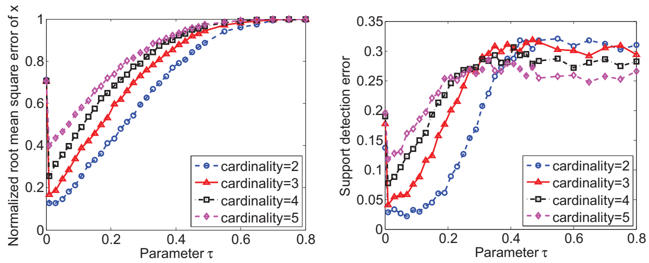

Fig. 3. The reconstruction error and the support detection error deteriorate as the cardinality increases. The length of $\mathbf{x}^*$ is 64, the number of observations is 30, and the cardinality varies from 2 to 5. The SNR is 23dB, and τ is tested from 0 to 1. 500 trials are simulated for each point.

4. NUMERICAL EXPERIMENTS

In this section, line-spectrum signal reconstruction and support detection are implemented to verify the above work. Three sets of experiments are displayed. During each trial of experiments, the compressing matrix and the noise are Gaussian randomly generated, and the positions of the lines in the spectrum are uniformly randomly generated. The signal reconstruction is done by solving problem (3) in which the convex optimization problem is solved by CVX [19], and the support detection is by the method in Proposition 1 except for the parallel experiment in the second set. The simulations are implemented in MATLAB on the Windows 7 operating system[1].

The first experiment is to demonstrate how the reconstruction error and the support detection error deteriorate as the measurement SNR decreases. The length of the unknown signal, the number of observations, and the cardinality of the spectrum are 64, 30, and 2, respectively. The SNR varies from 3.9dB to 43.0dB, and τ is tested from 0 to 1. The experiment is simulated by 500 trials for each point, and the results are shown in Fig.1. Each curve in the left sub-figure shows that when τ is 0, which is equivalent to the case where the least-square term does not appear in problem (3), the error is larger than the optimal error obtained by choosing the optimal τ, i.e. the lowest point in each curve. When τ is increasing from this optimal point, the error ascends to 1. The curves in the right sub-figure show that the support detection error increases in a different manner compared with the reconstruction error in the left when τ grows. In fact, the support detection error does not change too much when τ grows from the optimal value at the beginning, but has a rapid increment after τ grows out of a certain neighbourhood. Besides, the optimal support detection error , i.e. the lowest point in each curve, is almost the same when the SNR is high, but it degenerates more obviously when the SNR continues to descend.

The second experiment shows that by using the method based on dual certification in Proposition 1, the line spectra detection is much more accurate than the matrix pencil method [20] in the noisy circumstances. The result is in Fig.2, where the experimental scenario is the same as that of the first experiment. Before the support detection, the signal is reconstructed by solving problem (3). For each SNR level, τ is optimally chosen as suggested by the lowest point for each curve of the reconstruction error in the first experiment.

The third experiment demonstrates how the reconstruction error and the support position detection error deteriorate as the cardinality of the spectrum increases. The length of the unknown signal and the number of observations are 64 and 30, respectively. The SNR is chosen as 23dB. The parameter τ is tested from 0 to 1. The experiment is simulated by 500 trials for each point. The corresponding results are in Fig.3. Each curve in the left sub-figure indicates that when τ increases from its optimal point, the errors ascend almost linearly before τ is over large. Besides, the square of the optimal error also increases linearly with s. Such is also in the results of Theorem 1, where the convex hull of s appears instead of s. The curves in the right sub-figure show that the optimal support detection error degenerates when the cardinality continues to ascend.

5. CONCLUSION

In this work, the general sparse recovery problem is tackled by solving an atomic norm-regularized least-squares problem, in which the off-grid sparsity is delineated by atomic norm. The optimality conditions are provided and an off-grid support detection method is proposed. As a specific case, the problem of the line-spectrum signal recovery is formulated under the above general framework. The quantitative effect of the regularization parameter on the reconstruction error is discussed, and the deterioration brought in by the compression is bounded with probability dependent on the sparsity and the number of compressed measurements. Furthermore, numerical experiments of line-spectrum signal recovery focused on the recovery errors and the regularization parameter are simulated under different measurement SNR and different sparsity. When compared with the matrix pencil method, our support detection method has better performance with respect to the support position error. This paper is not

[1]The code for these simulations is available at
http://gu.ee.tsinghua.edu.cn/publications#sx1

exhaustive. More issues could be discussed in some related topics, such as to formulate the concrete expression on the optimal regularization parameter, and to theoretically analyze the support position detection error of the proposed method.

6. PROOF

6.1. The Proof of Lemma 1

Proof. Let's denote the objective function of a minimization problem as $f(\mathbf{x})$. Then $\hat{\mathbf{x}}$ being its solution is equivalent to

$$f(\hat{\mathbf{x}}+\alpha(\mathbf{x}-\hat{\mathbf{x}})) \geq f(\hat{\mathbf{x}}), \quad \forall \alpha \in (0,1),\ \mathbf{x} \in \mathbb{R}^N. \tag{11}$$

By substituting $f(\mathbf{x})$ in (11) with the objective function in problem (3) and using the convexity of atomic norm, one has

$$\tau(\|\mathbf{x}\|_{\mathcal{A}} - \|\hat{\mathbf{x}}\|_{\mathcal{A}}) \geq \langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi(\mathbf{x}-\hat{\mathbf{x}})\rangle, \tag{12}$$

which is the necessary and sufficient condition for $\hat{\mathbf{x}}$ being the solution. Equation (12) may be rewritten as

$$\tau\|\hat{\mathbf{x}}\|_{\mathcal{A}} - \langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi\hat{\mathbf{x}}\rangle \leq \inf_{\mathbf{x}}\{\tau\|\mathbf{x}\|_{\mathcal{A}} - \langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi\mathbf{x}\rangle\}. \tag{13}$$

From the definition of the dual norm of atomic norm, it yields that

$$\inf_{\mathbf{x}}\{\tau\|\mathbf{x}\|_{\mathcal{A}} - \langle \Phi^{\mathrm{T}}(\mathbf{y}-\Phi\hat{\mathbf{x}}), \mathbf{x}\rangle\} = \begin{cases} 0, & \|\Phi^{\mathrm{T}}(\mathbf{y}-\Phi\hat{\mathbf{x}})\|_{\mathcal{A}}^* \leq \tau; \\ -\infty, & \text{otherwise.}\end{cases}$$

Therefore, the first condition for the solution is that $\|\Phi^{\mathrm{T}}(\mathbf{y} - \Phi\hat{\mathbf{x}})\|_{\mathcal{A}}^* \leq \tau$. Then it yields from (13) that

$$\tau\|\hat{\mathbf{x}}\|_{\mathcal{A}} - \langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi\hat{\mathbf{x}}\rangle = \inf_{\mathbf{x}}\{\tau\|\mathbf{x}\|_{\mathcal{A}} - \langle \mathbf{y}-\Phi\hat{\mathbf{x}}, \Phi\mathbf{x}\rangle\} = 0, \tag{14}$$

which is the second condition. □

6.2. The Proof of Lemma 2

Proof. The Lagrangian function of problem (3) is

$$L(\mathbf{x},\mathbf{u},\lambda) = \frac{1}{2}\|\mathbf{y}-\Phi\mathbf{x}\|_2^2 + \tau\|\mathbf{u}\|_{\mathcal{A}} + \lambda^{\mathrm{T}}(\mathbf{x}-\mathbf{u}), \tag{15}$$

and the objective function in the dual problem is

$$\begin{aligned} g(\lambda) &= \inf_{\mathbf{x},\mathbf{u}} L(\mathbf{x},\mathbf{u},\lambda) \\ &= \inf_{\mathbf{x}}\left(\frac{1}{2}\|\mathbf{y}-\Phi\mathbf{x}\|_2^2 + \lambda^{\mathrm{T}}\mathbf{x}\right) + \inf_{\mathbf{u}}\left(\tau\|\mathbf{u}\|_{\mathcal{A}} - \lambda^{\mathrm{T}}\mathbf{u}\right) \end{aligned} \tag{16}$$

The two terms in (16) could be solved separately. From the definition of dual atomic norm, the second term is

$$\inf_{\mathbf{u}}\left(\tau\|\mathbf{u}\|_{\mathcal{A}} - \lambda^{\mathrm{T}}\mathbf{u}\right) = \begin{cases} 0, & \|\lambda\|_{\mathcal{A}}^* \leq \tau; \\ -\infty, & \text{otherwise.}\end{cases}$$

The first term in (16) is a semi-definite problem, and its minimum could be solved by derivation. The result is that it has a finite minimum if and only if

$$\lambda - \Phi^{\mathrm{T}}\mathbf{y} \in \mathcal{R}(\Phi^{\mathrm{T}}\Phi), \tag{17}$$

where $\mathcal{R}(\cdot)$ denotes the column space, and the minimum is

$$\frac{1}{2}\|(\Phi\Phi^{\mathrm{T}})^{-1}\Phi\lambda\|_2^2 + \lambda^{\mathrm{T}}\mathbf{z}, \tag{18}$$

where $\mathbf{z}$ satisfies that

$$\lambda - \Phi^{\mathrm{T}}\mathbf{y} = \Phi^{\mathrm{T}}\Phi\mathbf{z}. \tag{19}$$

Hence, considering the dual feasibility, we get that the dual problem $\max_\lambda g(\lambda)$ is formulated as (5). □

6.3. The proof of Proposition 1

Proof. Using (7) to substitute $\hat{\mathbf{x}}$ by $\hat{\lambda}$ into the first term of the inner product in the second condition of Lemma 1, one gets

$$\langle\hat{\lambda}, \hat{\mathbf{x}}\rangle = \tau\|\hat{\mathbf{x}}\|_{\mathcal{A}}. \tag{20}$$

Suppose that $\hat{\mathbf{x}} = \sum_{\mathbf{a}\in\mathcal{S}} c_{\mathbf{a}}\mathbf{a} + \sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}\mathbf{b}$. Then we have

$$\langle\hat{\lambda}, \hat{\mathbf{x}}\rangle = \sum_{\mathbf{a}\in\mathcal{S}} c_{\mathbf{a}}\langle\mathbf{a}, \hat{\lambda}\rangle + \sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}\langle\mathbf{b}, \hat{\lambda}\rangle \tag{21}$$

$$= \tau\sum_{\mathbf{a}\in\mathcal{S}} c_{\mathbf{a}} + \sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}\langle\mathbf{b}, \hat{\lambda}\rangle \tag{22}$$

For the right hand side of (20),

$$\tau\|\hat{\mathbf{x}}\|_{\mathcal{A}} = \tau\sum_{\mathbf{a}\in\mathcal{S}} c_{\mathbf{a}} + \tau\sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}. \tag{23}$$

If $\exists\mathbf{b} \notin \mathcal{S}$ such that $c_{\mathbf{b}} \neq 0$, then it has to be a contradiction that the last two equations cannot hold, since $\sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}\langle\mathbf{b}, \hat{\lambda}\rangle < \tau\sum_{\mathbf{b}\notin\mathcal{S}} c_{\mathbf{b}}$. Thus, $\forall\mathbf{b} \notin \mathcal{S}$, $c_{\mathbf{b}} = 0$, which means that $\mathcal{S}$ is the support set of $\hat{\mathbf{x}}$. □

6.4. The Proof of Theorem 1

Proof. First, Lemma 1 is used to deduce a bound on $\|\Phi\mathbf{x}^* - \Phi\hat{\mathbf{x}}\|_2^2$.

$$\begin{aligned} &\|\Phi\mathbf{x}^* - \Phi\hat{\mathbf{x}}\|_2^2 \\ =&\langle\Phi\mathbf{x}^*, \mathbf{y}-\Phi\hat{\mathbf{x}}\rangle - \langle\Phi\mathbf{x}^*, \mathbf{w}\rangle + \langle\Phi\hat{\mathbf{x}}, \mathbf{w}\rangle - \langle\Phi\hat{\mathbf{x}}, \mathbf{y}-\Phi\hat{\mathbf{x}}\rangle \\ \leq&\tau\|\mathbf{x}^*\|_{\mathcal{A}} - \langle\Phi\mathbf{x}^*, \mathbf{w}\rangle - \tau\|\hat{\mathbf{x}}\|_{\mathcal{A}} + \|\hat{\mathbf{x}}\|_{\mathcal{A}}\|\Phi^T\mathbf{w}\|_{\mathcal{A}}^* \\ \leq&2\tau\|\mathbf{x}^*\|_{\mathcal{A}}. \end{aligned} \tag{24}$$

Next, it is proved that the subspace that $\mathbf{x}^*$ lies in satisfies the Lipchitz-regularity, so that the results on Φ with respect to such kind of subspace in [16, Theorem 3] can be utilized, and the reconstruction error bound is obtained consequently. The details are as the following.

$\mathbf{x}^*$ is lying in a low dimensional subspace $\mathbf{V}_\omega = [\mathbf{a}_{\omega_1}, \cdots, \mathbf{a}_{\omega_s}]$, and $\hat{\mathbf{x}}$ is lying in $\mathbf{V}_{\bar{\omega}}$, where $\omega, \bar{\omega} \in \Omega = [0, 2\pi]^s$. The projection matrix of $\mathbf{V}_\omega$ is $\mathbf{P}_\omega = \mathbf{V}_\omega\mathbf{V}_\omega^{\mathrm{T}}$, and the norm of the projection operation is defined as the spectrum norm of the matrix. Note that $\mathbf{P}_\omega - \mathbf{P}_{\bar{\omega}}$ is a Hermitian Toeplitz matrix

$$[\mathbf{P}_\omega - \mathbf{P}_{\bar{\omega}}]_{mn} = \sum_{i=1}^{s} \mathrm{e}^{\mathrm{j}\omega_i(m-n)} - \mathrm{e}^{\mathrm{j}\bar{\omega}_i(m-n)}. \tag{25}$$

Using the bound on the eigenvalues of Hermitian Toeplitz matrix in [21], one has

$$\begin{aligned} &\|\mathbf{P}_\omega - \mathbf{P}_{\bar{\omega}}\| \leq \sup_\beta f(\beta) \\ =&\sup_\beta \sum_{i=1}^{s} \frac{\sin((\omega_i+\beta)(N-\frac{1}{2}))}{\sin(\frac{\omega_i+\beta}{2})} - \frac{\sin((\bar{\omega}_i+\beta)(N-\frac{1}{2}))}{\sin(\frac{\bar{\omega}_i+\beta}{2})} \\ \leq&\sum_{i=1}^{s}(2N-1)^2|\omega_i - \bar{\omega}_i|/\pi \leq \sqrt{s}(2N-1)^2\|\omega-\bar{\omega}\|_2/\pi. \end{aligned}$$

Hence, the Lipchitz-regularity of the subspace $\mathbf{V}_\omega$ is guaranteed. According to the result in [16, Theorem 3] , the restricted isometry constant is bounded by the probability that

$$P\left\{\delta > C\left(\sqrt{\frac{st}{M}} + \frac{s\log(2s)t}{M}\right)\right\} \leq sN_0\mathrm{e}^{d-t}, \tag{26}$$

in which $d = s$, and $N_0 = (2\pi)^s$. Combining the result in (24), we have (9) derived. □

7. REFERENCES

[1] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[2] Venkat Chandrasekaran, Benjamin Recht, Pablo A Parrilo, and Alan S Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.

[3] Nikhil Rao, Parikshit Shah, Stephen Wright, and Robert Nowak. A greedy forward-backward algorithm for atomic norm constrained minimization.

[4] Frank F Bonsall. A general atomic decomposition theorem and banach's closed range theorem. *Quarterly Journal of Mathematics*, 42(1):9–14, 1991.

[5] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *arXiv preprint arXiv:1207.6053*, 2012.

[6] Parikshit Shah, Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Linear system identification via atomic norm regularization. *arXiv preprint arXiv:1204.0590*, 2012.

[7] Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Atomic norm denoising with applications to line spectral estimation. 2012.

[8] Ramdas Kumaresan, DW Tufts, and Loues L Scharf. A prony method for noisy data: Choosing the signal components and selecting the order in exponential signal models. *Proceedings of the IEEE*, 72(2):230–233, 1984.

[9] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on*, 34(3):276–280, 1986.

[10] Zbigniew Leonowicz, Tadeusz Lobos, and Jacek Rezmer. Advanced spectrum estimation methods for signal analysis in power electronics. *Industrial Electronics, IEEE Transactions on*, 50(3):514–519, 2003.

[11] Jean-Marc Azais, Yohann De Castro, and Fabrice Gamboa. Spike detection from inaccurate samplings. *arXiv preprint arXiv:1301.5873*, 2013.

[12] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht. Near minimax line spectral estimation. *arXiv preprint arXiv:1303.4348*, 2013.

[13] Emmanuel Candes and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *arXiv preprint arXiv:1203.5871*, 2012.

[14] Snir Gazit, Alexander Szameit, Yonina C Eldar, and Mordechai Segev. Super-resolution and reconstruction of sparse subwavelength images. *arXiv preprint arXiv:0911.0981*, 2009.

[15] William Maetzel, Justin Romberg, Karim G Sabra, and William Kuperman. Compressive matched field processing. *The Journal of the Acoustical Society of America*, 127:1908, 2010.

[16] William Edward Jr. Mantzel. *Parametric estimation of randomly compressed functions*. PhD thesis, Georgia Institute of Technology, 2013.

[17] Andreas M Tillmann and Marc E Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *arXiv preprint arXiv:1205.2081*, 2012.

[18] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[19] Michael Grant, Stephen Boyd, and Yinyu Ye. cvx users guide. Technical report, Technical Report Build 711, Citeseer. Available at: http://citeseerx. ist. psu. edu/viewdoc/download, 2009.

[20] Yingbo Hua and Tapan K Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 38(5):814–824, 1990.

[21] Robert M Gray. *Toeplitz and circulant matrices: A review*. Now Pub, 2006.