2016 IEEE International Conference on Data Science and Advanced Analytics

The Uniqueness and Greedy Method for Quadratic Compressive Sensing

Jun Fan and Lingchen Kong
Department of Applied Mathematics
Beijing Jiaotong University
Beijing, China
Email: fanjunmath@hotmail.com
konglchen@126.com

Liqun Wang
Department of Statistics
University of Manitoba
Winnipeg, Canada
Email: liqun.wang@umanitoba.ca

Naihua Xiu
Department of Applied Mathematics
Beijing Jiaotong University
Beijing, China
Email: nhxiu@bjtu.edu.cn

***Abstract*—Quadratic compressive sensing, as a nonlinear extension of compressive sensing, has attracted considerable attention in optical image, X-ray crystallography, transmission electron microscopy, etc. We introduce the concept of uniform s-regularity to study the uniqueness in quadratic compressive sensing and propose a greedy algorithm for the corresponding numerical optimization. Moreover, we prove the convergence of the proposed algorithm under the uniform s-regularity condition. Finally, we present numerical results to demonstrate the efficiency of the proposed method.**

***Index Terms*—quadratic compressive sensing; sparsity; uniform s-regularity; uniqueness; greedy algorithm**

I. Introduction and Motivation

Compressive sensing (CS) has been intensively studied and widely used in the last decade. The main goal is to reconstruct sparse signals from the sampled measurements. Recently the theory has been extended to nonlinear compressive sensing, see, e.g., [3], [4], [13], [14] and [15] for more details. Particularly, the so-called quadratic compressive sensing (QCS) aims to find the sparse signal x to the problem

$$\min_{x\in\mathbb{R}^n} \|x\|_0 \;\; s.t. \;\; y_i = x^T A_i x + b_i^T x + c_i, \;\; i = 1,\cdots,m, \quad (1)$$

where $\|x\|_0$ is the number of nonzero elements of x and $y_i, c_i \in \mathbb{R}, b_i \in \mathbb{R}^n$ and $A_i \in \mathbb{R}^{n\times n}, i = 1,\cdots,m$ are the given real vectors and matrices respectively.

Similar to the linear compressive sensing, QCS has been widely used in scientific discoveries. For example, [15] and [16] point out that the relation between the measurements and image for partially-spatially-incoherent light is quadratic and demonstrate that the sub-wavelength optical images borne on partially-spatially-incoherent light can be recovered from their far-field or the blurred image, given the prior knowledge that the image is sparse. Consequently, they introduce the QCS to recover sub-wavelength information through bandwidth extrapolation algorithms. Mathematically, their problems can be described as follows: given an positive integer s and m symmetric matrices $A_1,\cdots,A_m \in \mathbb{R}^{n\times n}$, find a vector x satisfying

$$\begin{array}{l} x^T A_i x \approx y_i, \qquad i = 1,\cdots,m, \\ \|x\|_0 \le s. \end{array} .$$

Another example is the phase retrieval problem which plays an important role in, e.g., X-ray crystallography, transmission electron microscopy and coherent diffractive imaging. Generally speaking, the problem is to recover the lost phase information through the observed magnitudes. Mathematically, the phase retrieval problem is to find $x \in \mathbb{C}^n$ or $\mathbb{R}^n$ in the following model

$$y_i = |\langle a_i, x\rangle|^2, \qquad i = 1,\cdots,m,$$

where $a_i \in \mathbb{C}^n$ or $\mathbb{R}^n$ are given and $y_i \in \mathbb{R}$ are observed variables ([5], [6], [7], [10]). Let $\mathcal{R}(x)$ and $\mathcal{I}(x)$ denote the real and imaginary part for a complex number x respectively. Then the above relationship can be rewritten as

$$y_i = u^T A_i u, \qquad i = 1,\cdots,m,$$

where $u = (\mathcal{R}(x)^T, \mathrm{Im}(x)^T)^T$ and A_i is

$$\begin{pmatrix} \mathcal{R}(a_i)\mathcal{R}(a_i)^T & \mathcal{I}(a_i)\mathcal{R}(a_i)^T - \mathcal{R}(a_i)\mathcal{I}(a_i)^T \\ \mathcal{R}(a_i)\mathcal{I}(a_i)^T - \mathcal{I}(a_i)\mathcal{R}(a_i)^T & \mathcal{I}(a_i)\mathcal{I}(a_i)^T \end{pmatrix}.$$

In particular, in the real phase retrieval problem the goal is to find $x \in \mathbb{R}^n$ such that

$$y_i = x^T (a_i a_i^T) x, \qquad i = 1,\cdots,m.$$

In the above two examples, the relationship between the input and output signals has the form

$$y_i = x^T A_i x, \quad i = 1,\cdots,m. \quad (2)$$

For the sake of discussion, (2) is called a purely quadratic measurements model. A general quadratic measurements model can be defined as

$$y_i = x^T A_i x + x^T b_i + c_i, \quad i = 1,\cdots,m. \quad (3)$$

As is well-known in the literature, model (2) suffers the problem of identifiability because x and $-x$ are not distinguishable from the observed data. Usually this problem is dealt with by adopting a modular mapping that identifies $\pm x$ for any $x \in \mathbb{R}^n$, so that a unique solution for model (2) is obtained up to a change of sign. This method is widely used in the phase retrieval literature such as [1] and [12].

Therefore by including a linear term, model (3) is not only more general than model (2) mathematically, but it also helps

978-1-5090-5206-6/16 $31.00 © 2016 IEEE
DOI 10.1109/DSAA.2016.94

to ensure unique solution. For this reason sometimes both models are discussed separately below.

To solve problem (1) that includs sparse phase retrieval problem, a general approach is to use a lifting technique and recast it as a semi-definite programming (SDP) problem. See, e.g., [14], [15] and [16].

Another approach to solving problem (1) is to use greedy methods. For example, [4] shows that the iterative hard thresholding (IHT) algorithm, a popular greedy method for CS, can be used to accurately recover sparse or structured signals from a few nonlinear observations. The main ideas of this method is to first linearize the nonlinear mapping between the input and output signals and then use the projected gradient method to solve a sequence of linear least square subproblems with sparse constraints. Further, [3] generalizes the iterative hard thresholding method to solve a more general problem of minimizing a continuously differentiable objective function subject to a sparse constraint, which includes the ℓ_0-constrained least squares methods for the sparse recovery problems from quadratic measurements (2) as a special case. To guarantee the convergence of the proposed algorithms, the gradient of the nonlinear mapping in [4] or the objective function in [3] is assumed to be Lipschitz. For recovering a sparse signal in the phase retrieval problem, [17] employs a ℓ_0-constrained nonlinear least square method and proposed the GrEedy Sparse PhAse Retrieval (GESPAR) algorithm for the corresponding numerical optimization. Generally speaking, the GESPAR is an iterative local-search based algorithm for solving the ℓ_0-constrained nonlinear least square problem, where the support of the sought signal is updated iteratively according to a set of selections and the damped Gauss-Newton method is invoked to solve a subproblem for the given support.

Inspired by these works, we will employ the following ℓ_0-constrained least squares method for solving the QCS problem (1), i.e.,

$$\begin{aligned} \min_{x\in\mathbb{R}^n} f(x) &:= \sum_{i=1}^{m}(x^TA_ix + b_i^Tx + c_i - y_i)^2 \\ s.t. \qquad & \|x\|_0 \le s, \end{aligned} \tag{4}$$

where $s < n$ is a positive integer. Note that the algorithms of [4] and [3] cannot be used here because the gradient of $f(\cdot)$ is not Lipschitz. In the GESPAR algorithm, [17] shows that each limit point of the iterative sequence is a stationary point if the minimum eigenvalues of the Jacobian matrices calculated in the subproblems are uniformly bounded from below. Recall that in [3] the authors show that the iterative sequence generated by their algorithm converges when the mapping matrix is s-regular. In this paper we extend the s-regularity to uniform s-regularity. We will show that the uniform s-regularity plays an important role both in the uniqueness of the solution in the QCS problem and in the convergence of the algorithm we propose.

This paper is organized as follows. In section 2, we introduce the concept of uniform s-regularity and provide some sufficient conditions. We also prove the uniqueness in QCS problem. In section 3 we establish a fixed point equation for the minimization (4) under the assumption of uniform s-regularity. We also construct a projected gradient algorithm and discuss the convergence of the proposed algorithm. In section 4 we calculate some numerical examples to demonstrate our method. Conclusions are given in section 5, while some technical lemmas and mathematical proofs are given in the Appendices.

Throughout the paper we use the following notations. For any d-dimensional vector $v = (v_1, \cdots, v_d)^T$, let $|v| = (|v_1|, \cdots, |v_d|)^T, v^2 = (v_1^2, \cdots, v_d^2)^T, \|v\|_2 = (\sum_{i=1}^d v_i^2)^{\frac{1}{2}}$, $\|v\|_1 = \sum_{i=1}^d |v_i|$ and $\|v\|_\infty = \max\{|v_1|, \cdots, |v_d|\}$. For any set $\Gamma \subseteq \{1, \cdots, d\}$, $|\Gamma|$ denotes its cardinality and $\Gamma^c = \{1, \cdots, d\}/\Gamma$. For any $n \times d$ matrix $A = [a_{ij}]$, denote $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d a_{ij}^2}$ and $|A|_\infty = \max_{1\le i,j\le d} |a_{ij}|$. Further denote by A_Γ the sub-matrix of A consisting of the columns of A with index in $\Gamma \subseteq \{1, \cdots, d\}$, by $A^{\Gamma'}$ the sub-matrix of A consisting of the rows with index in $\Gamma' \subseteq \{1, \cdots, n\}$, and by $A^{\Gamma'\Gamma}$ the sub-matrix of A consisting of the rows and columns indexed by Γ' and Γ respectively. Especially, we use the notation v_Γ to denote the sub-vector for either a column or a row vector v. Further, denote by $A \otimes B$ the Kronecker product of two matrices A and B. Finally, $e_{d,j}$ denotes the jth column of the $d \times d$ identity matrix I_d.

II. Uniqueness

A. Uniform s-regularity

For a sparse linear model, [9] introduced the concept of spark and showed that the uniqueness of the underlying signal x can be characterized by the $\mathrm{spark}(B)$ which is defined as the minimum number of linearly dependent columns of the design matrix B. Another way to express this property is via the s-regularity of B, i.e., any s columns of B are linearly independent. Indeed, B is s-regular if and only if $\mathrm{spark}(B) \ge s + 1$ [3]. Further, in a linear model the residual function $R(x) = y - Bx$ has Jacobian matrix $-B$, where $y = (y_1, \cdots, y_n)^T$. Correspondingly, in model (3) the residual function $R(x) = \big(R_1(x), \cdots, R_n(x)\big)^T$, where $R_i(x) = y_i - x^TA_ix - b_i^Tx - c_i$, has Jacobian $\big(-2A_1x - b_1, \cdots, -2A_nx - b_n\big)^T$. This leads to the following definition.

Definition II.1. The linear transform $\mathbf{A}(x) = (A_1x, \cdots, A_mx)^T$ is said to be *uniformly s-regular* if $\mathbf{A}(x)_\Gamma$ has full column rank for any $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$ and $x \in \mathbb{R}^n/\{0\}$ with $\mathrm{supp}(x) \subseteq \Gamma$.

Definition II.2. The affine transform $\mathcal{A}(x) = (A_1x + b_1, \cdots, A_mx + b_m)^T$ is said to be *uniformly s-regular*, if $\mathcal{A}(x)_\Gamma$ has full column rank for any $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$ and $x \in \mathbb{R}^n$ with $\mathrm{supp}(x) \subseteq \Gamma$.

It is easy to see that when all A_i are zero matrices, the affine transform $\mathcal{A}(x)$ reduces to the constant transform B and therefore the uniform s-regularity coincides with s-regularity. Further, by taking $x = 0$ one can see that the uniform s-regularity of the affine transform $\mathcal{A}(x)$ implies the s-regularity

of $B = (b_1, \cdots, b_m)^T$. However, since the rank of the affine transform $\mathcal{A}(x)$ depends on x, its rank cannot always be determined by the ranks of A_i, $i = 1, \cdots, m$ and B.

For any $u \in \mathbb{R}^s$, denote $\mathbf{A}_\Gamma(u) = (A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u)^T$ and

$$\begin{aligned}\mathcal{A}_\Gamma(u) &= (A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u)^T + (b_{1\Gamma}, \cdots, b_{m\Gamma})^T \\ &= \mathbf{A}_\Gamma(u) + B_\Gamma(u).\end{aligned}$$

Then it is straightforward to verify the following results.

Property II.1. $\mathbf{A}(\cdot)$ is uniformly s-regular if and only if $\mathbf{A}_\Gamma(\cdot)$ has full column rank for any $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$ and $u \in \mathbb{R}^s/\{0\}$.

Property II.2. $\mathcal{A}(\cdot)$ is uniformly s-regular if and only if $\mathcal{A}_\Gamma(\cdot)$ has full column rank for any index set $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$ and $u \in \mathbb{R}^s$.

Moreover, we have the following results, where we denote by $V_{i\Gamma} = \{v \in \mathbb{R}^s : A_i^{\Gamma\Gamma}v = \lambda v, \lambda \neq 0\}/\{0\}$, $i = 1, \cdots, m$, for any index set $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$.

Proposition II.1. (i) Suppose matrices $\{A_i\}$ are symmetric. Then $\mathbf{A}(\cdot)$ is uniformly s-regular, if for any $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$ and $v_i \in V_{i\Gamma}, i = 1, \cdots, m$, either $\{v_i\}_{i\in\mathbb{K}}$ or $\{v_i\}_{i\in\mathbb{K}^c}$ span $\mathbb{R}^s$ for every subset $\mathbb{K} \subseteq \{1, \cdots, m\}$.

(ii) Especially, if $A_i = a_i a_i^T$ for some $\{a_i\} \in \mathbb{R}^n/\{0\}$, $i = 1, \cdots, m$, then the uniform s-regularity of $\mathbf{A}(\cdot)$ is equivalent to the so-called s-complement property of $\{a_i\}$, i.e., either $\{a_i^\Gamma\}_{i\in\mathbb{K}}$ or $\{a_i^\Gamma\}_{i\in\mathbb{K}^c}$ span $\mathbb{R}^s$ for every subset $\mathbb{K} \subseteq \{1, \cdots, m\}$ and $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$.

Remark II.1. In the phase retrieval problem, [1] and [2] introduce the complement property and show that it is a necessary and sufficient condition for the measurement vectors to yield injective and stable intensity measurements. For the sparse case, [12] propose the concept of s-complement property which is less restrictive than the complement property. The above proposition shows that the uniform s-regularity can be applied to more negeral models.

Proposition II.2. If $\sum_{i=1}^m b_i \otimes A_i = 0$ and $B = (b_1, \cdots, b_m)^T$ is s-regular, then $\mathcal{A}(\cdot) = \mathbf{A}(\cdot) + B$ is uniformly s-regular. Especially, if $A_i = I$, $i = 1, \cdots, m$, $\sum_{i=1}^m b_i = 0$ and $B = (b_1, \cdots, b_m)^T$ is s-regular, then $\mathcal{A}$ is uniformly s-regular.

B. Uniqueness

In [12] it is pointed out that the unique recovery of an s-sparse real signal is guaranteed by the s-complement property of $\{a_i\}$. As discussed above, the s-complement property can be generalized to the uniform s-regularity. A natural question is whether the later condition implies the uniqueness solution in QCS. To answer this quesiton, we provide the following results.

Theorem II.1. (i) Let $\bar{y}_i = x^{*T}A_i x^*$, $i = 1, \cdots, m$. Then model (2) has unique solution x^* satisfying $\|x^*\|_0 \leq s$ if $\mathbf{A}(\cdot)$ is uniformly $2s$-regular.

(ii) Let $\bar{y}_i = x^{*T}A_i x^* + b_i^T x^* + c_i$, $i = 1, \cdots, m$. Then model (3) has unique solution x^* satisfying $\|x^*\|_0 \leq s$ if $\mathcal{A}(\cdot)$ is uniformly $2s$-regular.

III. Optimization algorithm

In this section we discuss the numerical computation of problem (4). To this end we define $S = \{x \in \mathbb{R}^n : \|x\|_0 \leq s\}$ for a positive integer s and $P_S(\cdot)$ to be the orthogonal projection onto S, which is a vector consisting of the s elements of x with the largest absolute values. For any $x \in \mathbb{R}^n$, let $M_i(x)$ be the ith largest absolute value component in x, where $i = 1, \cdots, n$.

We first establish a fixed point equation for the optimization problem (4), which is used to construct a projected gradient algorithm.

Theorem III.1. Assume that (i) in model (2), $\mathbf{A}(\cdot)$ is uniformly s-regular; or (ii) in model (3), $\sum_{i=1}^m b_i \otimes A_i = 0$ and $B = (b_1, \cdots, b_m)^T$ is s-regular. Then there exists a vector $\hat{x} \in \mathbb{R}^n$ that is a minimizer of problem (4). Further, there exists a positive constant $\hat{L}$ such that for any $\tau \in \big(0, \min\{\hat{L}^{-1}, 1\}\big]$, it holds

$$\hat{x} = \mathcal{P}_S\big(\hat{x} - \tau\nabla f(\hat{x})\big). \tag{5}$$

Based on the fixed point equation (5), we propose the following algorithm for the computation of (4).

Algorithm:

Step 0. Given $\lambda > 0, \epsilon \geq 0, \gamma, \alpha \in (0, 1), \delta > 0$, choose an arbitrary x^0 and set $k = 0$.

Step 1. (a) Compute $\nabla f(x^k)$ from

$$\nabla f(x) = 2\sum_{i=1}^m (x^T A_i x + b_i^T x - y_i)\big((A_i + A_i^T))x + b_i\big);$$

(b) Compute $x^{k+1} = P_S(x^k - \tau_k \nabla f(x^k))$, where $\tau_k = \gamma\alpha^{j_k}$ and j_k is the smallest nonnegative integer such that

$$f(x^k) - f(x^{k+1}) \geq \frac{\delta}{2}\|x^k - x^{k+1}\|_2^2. \tag{6}$$

Step 2. Stop if

$$\frac{\|x^{k+1} - x^k\|_2}{\max\{1, \|x^k\|_2\}} \leq \epsilon.$$

Otherwise, replace k by $k + 1$ and go to Step 1.

Remark III.1. A key point in the above algorithm is to find the smallest nonnegative integer j_k such that (6) holds, which can be done successfully by Lemmas B.3 and B.4 in Appendix B. Another key point is the choice of the sparsity parameter s which may not be known *a priori* in some applications. A popular and efficient method for choosing the penalty parameter in ℓ_1-regularized minimization is cross validation which can be applied to problems such as compressed sensing. In the next

section we will calculate some numerical examples using 5-fold cross validation to determine the sparsity parameter s. The numerical results demonstrate the efficiency of this method.

Remark III.2. In [3] the authors studied the sparsity constrained optimization problem

$$\min_{x\in\mathbb{R}^n} h(x) \quad s.t. \quad \|x\|_0 \le s, \tag{7}$$

where $h(\cdot)$ is a continuously differentiable function, and introduced the so-called L-stationary point x that satisfies

$$x \in P_S(x - \frac{1}{L}\nabla h(x)), \quad L > 0.$$

They proved that the above relation holds if and only if $\|x\|_0 \le s$ and

$$|\nabla h(x)_i| \begin{cases} \le LM_s(x), & \text{if } i \notin \text{supp}(x); \\ = 0, & \text{if } i \in \text{supp}(x) \end{cases} \tag{8}$$

where $\nabla h(x)_i$ is the ith element of the gradient vector ∇h. In articular, it is pointed out that if x is an L-stationary point for some $L > 0$, then (i) $\nabla h(x) = 0$ when $\|x\|_0 < s$; and (ii) $\nabla h(x)_i = 0$ for all $i \in \text{supp}(x)$ when $\|x\|_0 = s$. Furthermore, it is shown in [3] that under the assumption that the gradient of the objective function ∇h is Lipschitz with constant $L(h)$, the minimizer of problem (7) must be an L-stationary point when $L > L(h)$. Unfortunately our function $f(\cdot)$ does not satisfy this assumption. However, the problem can be overcome if the uniform s-regularity is assumed.

Now we consider the convergence of the proposed algorithm.

Theorem III.2. Let $\{x^k\}$ be the sequence generated by the above algorithm. Then,

(i) $\lim_{k\to\infty} \|x^{k+1} - x^k\|_2 = 0$;

(ii) any accumulation point of $\{x^k\}$ is a stationary point of the minimization problem (4).

IV. Numerical Examples

In this section we demonstrate the efficiency of our proposed algorithm by calculating some numerical examples related to the phase retrieval problem with real measurement vectors and real signal [1], [3], [12].

Example. Suppose that $x \in \mathbb{R}^n$ is a discrete signal and we observe the squared modulus of the inner product between the signal and some vectors $a_j \in \mathbb{R}^n$

$$y_j = \langle a_j, x\rangle^2, \qquad j = 1, \cdots, m. \tag{9}$$

Let the true value x^* be generated randomly with s nonzero components from the standard Gaussian distribution. To recover x^*, we use the ℓ_0-constrained least squares method which can be formulated as

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^{m} \left((x^T a_i)^2 - y_i\right)^2 \quad s.t. \quad \|x\|_0 \le s, \tag{10}$$

where s is a positive integer and is treated as unknown here. As mentioned in Remark III.1, we use cross validation method to choose the sparsity s. To evaluate the performance of our algorithm, we carry out 100 Monte Carlo runs in each simulation and report the mean and standard errors (SE) of $\|\hat{x}\|_0$, where $\hat{x}$ is a minimizer of the optimization (10). We also report the successful recovery (SR) rate using the criterion $\hat{\Gamma} = \Gamma^*$ and $\|\hat{x} - x^*\|_2 \le 0.01$, where $\Gamma^* = \{j : x^*_j \ne 0\}$ and $\hat{\Gamma} = \{j : \hat{x}_j \ne 0\}$.

In [12] it is pointed out that a set of $4s - 1$ independent samples from an n dimensional standard Gaussian distribution satisfies the $2s$ complement property with probability 1. Therefore in our simulations vectors $a_j \in \mathbb{R}^n$ are generated from the standard Gaussian distribution and the relation between the dimension n and the sparsity s satisfies $n \ge 4s$. Similar to [3] we consider the cases $n = 120$ and $m = 80$ with $s = 3, 4, \cdots, 10$ respectively.

The numerical results are given in Table I. These results show that the averages of $\|\hat{x}\|_0$ are fairly close to the corresponding true values $\|x^*\|_0$ and the SE are very small overall. They also confirm that the cross validation method works well in choosing the right sparsity parameter. The rates of successful recovery are over 50% and are stable when the sparsity of true value increases. These rates are comparable with similar studies in the literature. For example, compared to the results in [3], our SR rates are lower in the cases where $s = 3, 4$ but significantly higher in the cases where $s = 5, 6, 7, 8, 9, 10$. Moreover, our SR rates can be significantly higher if more relaxed numerical accuracy criterion is used. For example, if an estimated value $|\hat{x}(i)| \le 5 \times 10^{-4}$ can be regarded as being estimated as zero, then the SR can be higher than 90%.

TABLE I
THE AVERAGE RESULTS OF 100 SIMULATIONS WITH $n = 120$ AND $m = 80$.

$\|x^*\|_0$	($\|\hat{x}\|_0$, SE)	SR
3	(3.7, 0.078)	0.52
4	(4.9, 0.127)	0.55
5	(5.8, 0.099)	0.50
6	(6.9, 0.115)	0.51
7	(8.0, 0.124)	0.51
8	(8.9, 0.135)	0.58
9	(10.2, 0.154)	0.52
10	(10.7, 0.087)	0.50

To assess the efficiency of our method in the situation of high-dimensional signal recovery with low sample size, we also run the simulations with $m = 3n/4$, $s = 0.05n$ and $n = 100,\ 200,\ 300, 400,\ 500$, respectively. The numerical results are given in Table II, which demonstrate further that cross validation is an appropriate method for the choice of sparsity used in the projection operator. While the last column in II show that our method can recover the unknown signal with higher success rates even in relatively high dimensional cases.

TABLE II
THE AVERAGE RESULTS OF SUCCESSFUL RECOVERIES.

n	$\|x^*\|_0$	($\|\hat{x}\|_0$, SE)	SR
100	5	(5.9, 0.120)	0.51
200	10	(10.6, 0.103)	0.67
300	15	(15.67, 0.111)	0.67
400	20	(20.6, 0.133)	0.71
500	25	(26.2, 0.233)	0.64

V. CONCLUSION

Quadratic compressive sensing (QCS) has been widely used in many fields such as optical images, X-ray crystallography, transmission electron microscopy. We have introduced the concept of uniform s-regularity to study the uniqueness of the solutions in quadratic measurements models. In the framework of sparsity-constrained nonlinear optimization, we have derived a fixed point equation and proposed a greedy algorithm for the numerical computaiton. The convergence of this algorithm is proved under fairly mild assumptions. Finally, we have presented a numerical example to illustrate the proposed method. The numerical results show that the proposed method performs well in cases with various dimensions and sparsity.

ACKNOWLEDGMENT

The research of Jun Fan, Lingchen Kong and Naihua Xiu was supported by the National Natural Science Foundation of China (11431002,11671029). The research of Liqun Wang was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] R. Balan, P. Casazza and D. Edidin, *On signal reconstruction without phase*, Appl. Comput. Harmon. Anal., vol. 20, pp. 345-356, 2006.

[2] A. Bandeira, J. Cahill, D. Mixon and A. Nelson, "Saving phase: Injectivity and stability for phase retrieval", *Appl. Comput. Harmon. Anal.* vol.37, pp. 106-125, 2014.

[3] A. Beck and Y. Eldar, "Sparsity constrained nonlinear optimization: Optimality conditions and algorithms", *SIAM J. Optim.*, vol. 23, pp. 1480-1509, 2013.

[4] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems", *IEEE Trans. Inform. Theory*, vol. 59, pp. 3466-3474, 2013.

[5] E.J. Candès, T. Strohmery and V. Voroninski, "PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming", *Comm. Pure Appl. Math.*, vol. 66, pp. 1241-1274, 2013.

[6] E.J. Candès, Y.C. Eldar, T. Strohmery and V. Voroninski, "Phase retrieval via matrix completion", *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 199-225, 2013.

[7] E.J. Candès, X., Li and M., Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms", Preprint. Available at arXiv: 1407.1065, 2014.

[8] D. Donoho, "High-dimensional data analysis: The curses and blessings of dimensionality", *AMS Math Challenges Lecture*, pp. 1-32, 2000.

[9] D. Donoho & M. Elad, M, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization", *Proc. Nat. Acad. Sci.*, vol. 100, pp. 2197-2202, 2003.

[10] Y. Eldar & S. Mendelson, "Phase retrieval: Stability and recovery guarantees", *Appl. Comput. Harmon. Anal.*, vol. 36, pp. 473-494, 2014.

[11] Y. Eldar, P. Sidorenko, D. Mixon, S. Barel, O. Cohen, "Sparse phase retrieval from short-time Fourier measurements", *Signal Processing Letters, IEEE,* vol. 22, no. 5, pp. 638-642, 2015.

[12] H. Ohlsson & Y. Eldar, "On conditions for uniqueness in sparse phase retrieval", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1841-1845, 2014.

[13] H. Ohlsson, A. Yang, R. Dong and S. Sastry, "CPRL–An Extension of Compressive Sensing to the Phase Retrieval Problem", In *Advances in Neural Information Processing Systems*, pp. 1367-1375, 2012.

[14] H. Ohlsson, A. Yang, R. Dong, M. Verhaegen and S. Sastry, "Quadratic basis pursuit", Preprint. Available at arXiv:1301.7002, 2013.

[15] Y. Shechtman, Y. Eldar, A. Szameit, and M. Segev, "Sparsity-based sub-wavelength imaging with partially spatially incoherent light via quadratic compressed sensing", *Optics Express*, vol. 19, pp. 14807-14822, 2011.

[16] Y. Shechtman, A. Szameit, E. Bullkich, et al. (2012). Sparsity-based single-shot sub-wavelength coherent diffractive imaging. *Nature Materials*, vol. 11, pp. 455-459, 2012.

[17] Y. Shechtman, A. Beck, and Y. Eldar, "GESPAR: Efficient phase retrieval of sparse signals. Signal Processing", *IEEE Transactions on*, vol. 62, no. 4, pp.928-938, 2014.

APPENDIX A
UNIQUENESS ANALYSIS

***Proof of Proposition II.1**:*

(i) First note that for any index set $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$, it follows from the eigenvalue decomposition that, for each $i = 1, \cdots, m$,

$$A_i^{\Gamma\Gamma} = \sum_{j=1}^{r_i} \lambda_{ij} q_{ij} q_{ij}^T$$

where $\{\lambda_{ij}\}$ are the nonzero eigenvalues and $\{q_{ij}\}$ are orthogonal eigenvectors with $\|q_{ij}\| = 1$ for $j = 1, \cdots, r_i$. Hence it is easy to verify that for any $u \in \mathbb{R}^s/\{0\}$,

$$\begin{aligned}&(A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u)^T\\ =&(\sum_{j=1}^{r_1} \lambda_{1j} u^T q_{1j} q_{1j}, \cdots, \sum_{j=1}^{r_m} \lambda_{mj} u^T q_{mj} q_{mj})^T.\end{aligned}$$

For any $u \in \mathbb{R}^s/\{0\}$, we denote $\mathbb{T} = \{i : u^T q_{ij} = 0, \text{ for each } j = 1, \cdots, r_i\}$ and $t = |\mathbb{T}|$. Then for the case $t \geq 1$ we can assume without loss of generality that $\mathbb{T} = \{1, \cdots, t\}$. It follows that

$$\begin{aligned}&(A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u)^T\\ =&(0, \cdots, 0, \sum_{j=1}^{r_{t+1}} \lambda_{t+1j} u^T q_{t+1j} q_{t+1j}, \cdots, \sum_{j=1}^{r_m} \lambda_{mj} u^T q_{mj} q_{mj})^T\end{aligned}$$

which implies that

$$\begin{aligned}&\text{rank}((A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u))\\ =&\text{rank}(\sum_{j=1}^{r_{t+1}} \lambda_{t+1j} u^T q_{t+1j} q_{t+1j}, \cdots, \sum_{j=1}^{r_m} \lambda_{mj} u^T q_{mj} q_{mj}).\end{aligned}$$

For each $i = t+1, \cdots, m$, since $\{q_{ij}\}$ are linearly independent and $u \neq 0$, we have

$$\sum_{j=1}^{r_i} \lambda_{ij} u^T q_{ij} q_{ij} \in V_i.$$

Therefore by the definition of $\mathbb{T}$ and that $t \geq 1$, there exists nonzero vector u_0 such that $u_0^T q_{ij} = 0$ for $i \in \mathbb{T}$ and $j = 1, \cdots, r_i$. It follows that $\{q_{ij}, j = 1, \cdots, r_i\}_{i \in \mathbb{T}}$ are linearly dependent, which together with the assumption of the Proposition imply

that $\sum_{j=1}^{r_{t+1}} \lambda_{t+1j} u^T q_{t+1j} q_{t+1j}, \cdots, \sum_{j=1}^{r_m} \lambda_{mj} u^T q_{mj} q_{mj}$ span $\mathbb{R}^s$. Therefore we have $\mathrm{rank}((A_1^{\Gamma\Gamma}u, \cdots, A_m^{\Gamma\Gamma}u)) = s$. Finally since u and Γ are arbitrary, property II.1 implies the uniform s-regularity of $\mathbf{A}(\cdot)$.

(ii) Suppose $\{a_i\} \in \mathbb{R}^n/\{0\}$ satisfy the s-complement property. We use the above result to show that $\mathbf{A}(\cdot)$ is uniformly s-regular. For any $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$, it is easy to see that for each $i = 1, \cdots, m$, $\|a_{i\Gamma}\|^{-1} a_{i\Gamma}$ is the unique eigenvector associated with the unique nonzero eigenvalue $\|a_{i\Gamma}\|^2$. That is, $V_{i\Gamma} = \{v \in \mathbb{R}^s : A_i^{\Gamma\Gamma} v = \lambda v, \lambda \neq 0\}/\{0\} = \mathrm{span}(a_{i\Gamma})/\{0\}$. Based on the s-complement property of $\{a_i\}$ and the first result, it follows that $\mathbf{A}(\cdot)$ is uniformly s-regular.

Given that $\mathbf{A}(\cdot)$ is uniformly s-regular, we now show that $\{a_i\}$ satisfy the s-complement property. To this end it suffices to prove that either $\{a_i^\Gamma\}_{i\in\mathbb{K}}$ or $\{a_i^\Gamma\}_{i\in\mathbb{K}^c}$ span $\mathbb{R}^s$ for every subset $\mathbb{K} \subseteq \{1, \cdots, m\}$. Without loss of generality we assume that $\mathbb{K} = \{1, \cdots, k\}$. We prove the result by contradiction. Suppose that both $a_{1\Gamma}, \cdots, a_{k\Gamma}$ and $a_{k+1\Gamma}, \cdots, a_{m\Gamma}$ cannot span $\mathbb{R}^s$ and denote $\tilde{A}_1 = (a_{1\Gamma}, \cdots, a_{k\Gamma})^T$ and $\tilde{A}_2 = (a_{k+1\Gamma}, \cdots, a_{m\Gamma})^T$. Then there exists $u_0 \in \mathbb{R}^s/\{0\}$ such that $\tilde{A}_1 u_0 = 0$ and it follows that

$$\begin{aligned}&\big((a_{1\Gamma}^T u_0)a_{1\Gamma}, \cdots, (a_{m\Gamma}^T u_0)a_{m\Gamma}\big)^T\\ =&\big(0, \tilde{A}_2^T \mathrm{diag}(a_{k+1\Gamma}^T u_0, \cdots, a_{m\Gamma}^T u_0)\big)^T\end{aligned}$$

and therefore

$$\begin{aligned}&\mathrm{rank}\big((a_{1\Gamma}^T u_0)a_{1\Gamma}, \cdots, (a_{m\Gamma}^T u_0)a_{m\Gamma}\big)^T\\ =&\mathrm{rank}\big(0, \tilde{A}_2^T \mathrm{diag}(a_{k+1\Gamma}^T u_0, \cdots, a_{m\Gamma}^T u_0)\big)^T.\end{aligned}$$

Since $a_{k+1\Gamma}, \cdots, a_{m\Gamma}$ cannot span $\mathbb{R}^s$, it follows that $\mathrm{rank}(\tilde{A}_2) < s$, which implies that

$$\begin{aligned}&\mathrm{rank}\big(0, \tilde{A}_2^T \mathrm{diag}(a_{k+1\Gamma}^T u_0, \cdots, a_{m\Gamma}^T u_0)\big)\\ =&\mathrm{rank}\big(\tilde{A}_2^T \mathrm{diag}(a_{k+1\Gamma}^T u_0, \cdots, a_{m\Gamma}^T u_0)\big)\\ \leq&\mathrm{rank}(\tilde{A}_2)\\ <&s.\end{aligned}$$

However, the uniform s-regularity of $\mathbf{A}(\cdot)$ implies that $\mathrm{rank}(\mathbf{A}_\Gamma(u)) = \mathrm{rank}\big((a_{1\Gamma}^T u_0)a_{1\Gamma}, \cdots, (a_{m\Gamma}^T u_0)a_{m\Gamma}\big)^T$ has full column rank, which is a contradiction.

■

Proof of Proposition II.2: Denote by e_j the jth column of the $n \times n$ identity matrix I_n. For each $k, l = 1, 2, \cdots, n$, it follows that

$$e_k^T (\sum_{i=1}^n A_i x b_i^T) e_l = x^T (\sum_{i=1}^m A_i^T e_k e_l^T b_i) = 0,$$

where the last equality follows from the assumption that $\sum_{i=1}^m b_i \otimes A_i = 0$. Therefore we have

$$\sum_{i=1}^n A_i x b_i^T = 0. \tag{11}$$

Since

$$\begin{aligned}&(\mathbf{A}(x) + B)^T (\mathbf{A}(x) + B)\\ =&\mathbf{A}(x)^T \mathbf{A}(x) + \mathbf{A}(x)^T B + B^T \mathbf{A}(x) + B^T B,\end{aligned}$$

$$\mathbf{A}(x)^T B = \sum_{i=1}^n A_i x b_i^T$$

and $B^T A = (A^T B)^T$, we have

$$(\mathbf{A}(x) + B)^T (\mathbf{A}(x) + B) = \mathbf{A}(x)^T \mathbf{A}(x) + B^T B.$$

For an index set $\Gamma \subseteq \{1, \cdots, n\}$ with $|\Gamma| = s$, we then conclude from the s-regularity of B_Γ that for any $u \in \mathbb{R}^s/\{0\}$ the matrix

$$(\mathbf{A}_\Gamma(u) + B_\Gamma)^T (\mathbf{A}_\Gamma(u) + B_\Gamma) = \mathbf{A}_\Gamma(u)^T \mathbf{A}_\Gamma(u) + B_\Gamma^T B_\Gamma$$

is positive definite, which implies that $\mathrm{rank}(\mathbf{A}_\Gamma(u) + B_\Gamma)^T (\mathbf{A}_\Gamma(u) + B_\Gamma) = s$. Combing this and the fact that $\mathrm{rank}(\mathbf{A}_\Gamma(u) + B_\Gamma) = \mathrm{rank}(\mathbf{A}_\Gamma(u) + B_\Gamma)^T (\mathbf{A}_\Gamma(u) + B_\Gamma)$, we get the desired result.

■

Proof of Theorem II.1: We first prove the result (i). Again we prove it by contradiction. Assume that $\tilde{x} \neq \pm x^*, \|\tilde{x}\|_0 \leq \|x^*\|_0, \mathcal{M}(\tilde{x}) = \mathcal{M}(\pm x^*), \tilde{x} \in \mathbb{R}^n$, where the operator $\mathcal{M}$ is defined as $\big(\mathcal{M}(x)\big)(i) = x^T A_i x$. Denote $\Gamma = \mathrm{supp}(\tilde{x}) \cup \mathrm{supp}(x^*)$. Then, $|\Gamma| \leq 2s$. For any $x \in \mathbb{R}^n$, it follows that $(\mathcal{M}(\tilde{x}))(i) = (x^\Gamma)^T A_i^{\Gamma\Gamma} x^\Gamma$ and therefore

$$\begin{aligned}0 =&\mathcal{M}(\tilde{x}) - \mathcal{M}(x^*)\\ =&\big((\tilde{x}^\Gamma - x^{*\Gamma})^T A_1^{\Gamma\Gamma}(\tilde{x}^\Gamma + x^{*\Gamma}), \cdots,\\ &\qquad (\tilde{x}^\Gamma - x^{*\Gamma})^T A_m^{\Gamma\Gamma}(\tilde{x}^\Gamma + x^{*\Gamma})\big)^T\\ =&A_\Gamma(\tilde{x}^\Gamma + x^{*\Gamma})(\tilde{x}^\Gamma - x^{*\Gamma}).\end{aligned}$$

Since $0 \neq \tilde{x}^\Gamma + x^{*\Gamma}$ and $\mathbf{A}(\cdot)$ is uniformly $2s$-regular, it follows that $\mathbf{A}_\Gamma(\tilde{x}^\Gamma + x^{*\Gamma})$ has full column rank, which implies $\tilde{x}^\Gamma = x^{*\Gamma}$. This is a contradiction.

Next we prove result (ii) by contradiction. Assume that $\tilde{x} \neq x^*, \|\tilde{x}\|_0 \leq \|x^*\|_0, \mathcal{M}(\tilde{x}) = \mathcal{M}(x^*), \tilde{x} \in \mathbb{R}^n$, where the operator $\mathcal{M}$ is defined as $\big(\mathcal{M}(x)\big)(i) = x^T A_i x + x^T b_i$. Denote $\Gamma = \mathrm{supp}(\tilde{x}) \cup \mathrm{supp}(x^*)$. Then, $|\Gamma| \leq 2s$. For any $x \in \mathbb{R}^n$, it follows that $(\mathcal{M}(\tilde{x}))(i) = (x^\Gamma)^T A_i^{\Gamma\Gamma} x^\Gamma + (x^\Gamma)^T b_i^\Gamma$ and therefore

$$\begin{aligned}0 =&\mathcal{M}(\tilde{x}) - \mathcal{M}(x^*)\\ =&\big((\tilde{x}^\Gamma - x^{*\Gamma})^T (A_1^{\Gamma\Gamma} + b_1^\Gamma)(\tilde{x}^\Gamma + x^{*\Gamma}), \cdots,\\ &\qquad (\tilde{x}^\Gamma - x^{*\Gamma})^T (A_m^{\Gamma\Gamma} + b_m^\Gamma)(\tilde{x}^\Gamma + x^{*\Gamma})\big)^T\\ =&\mathcal{A}_\Gamma(\tilde{x}^\Gamma + x^{*\Gamma})(\tilde{x}^\Gamma - x^{*\Gamma}).\end{aligned}$$

Since $\mathcal{A}(\cdot)$ is uniformly $2s$-regular, it follows that $\mathcal{A}_\Gamma(\tilde{x}^\Gamma + x^{*\Gamma})$ has full column rank which leads to $\tilde{x}^\Gamma = x^{*\Gamma}$. This is a contradiciton.

■

APPENDIX B
OPTIMIZATION ALGORITHM ANALYSIS

Lemma B.1. For any $b \in \mathbb{R}^n$, consider the following projection problem

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x - b\|_2^2 \quad s.t. \quad \|x\|_0 \leq s.$$

Then, the minimizer $\tilde{x}$ satisfies $\tilde{x} = P_S(b)$.

Lemma B.2. Let $b_i \in \mathbb{R}^n$ and $A_i \in \mathbb{R}^{n\times n}, i = 1, \cdots, m$ be the given matrices (vectors). Let s be a positive integer satisfying $s < n$ and Γ be any subset of $\{1, 2, \cdots, n\}$ satisfying $|\Gamma| = s$.

(i) If $\mathbf{A}(\cdot)$ is uniformly s-regular, then the function $f_{1\Gamma}(u) = \sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u)^2$ is coercive, i.e., $\lim_{\|u\|\to\infty} f_1(u) = \infty$.

(ii) If $\sum_{i=1}^m b_i \otimes A_i = 0$ and $B = (b_1, \cdots, b_m)^T$ is s-regular, then the function $f_{2\Gamma}(u) = \sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u + u^T b_{i\Gamma})^2$ is coercive, i.e., $\lim_{\|u\|\to\infty} f_{2\Gamma}(u) = \infty$.

Proof: To show result (i), note that there exists a vector $v^* \in \mathbb{R}^s$ satisfying $\|v^*\| = 1$ and

$$f_{1\Gamma}(v^*) = \min f_{1\Gamma}(v) \quad s.t. \quad \|v\| = 1$$

because $f_{1\Gamma}(\cdot)$ is continuous and the set $\|v\| = 1$ is compact. Then the uniform s-regularity implies that the matrix $\sum_{i=1}^m A_i^{\Gamma\Gamma} v^* v^{*T} A_i^{\Gamma\Gamma T}$ is positive definite, and therefore $f_{1\Gamma}(v^*) > 0$. For any nonzero vector $u \in \mathbb{R}^s$, it follows that

$$f_{1\Gamma}(u) = \|u\|^4 \sum_{i=1}^m \Big((\frac{u}{\|u\|})^T A_i^{\Gamma\Gamma} (\frac{u}{\|u\|}) \Big)^2 \geq \|u\|^4 f_{1\Gamma}(v^*)$$

which implies that $f_{1\Gamma}(u) \to \infty$ as $\|u\| \to \infty$. That is, the first result holds.

To show the result (ii), note that for any vector $u \in \mathbb{R}^s$, we have

$$\begin{aligned} & f_{2\Gamma}(u) \\ =& \sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u + u^T b_{i\Gamma})^2 \\ =& \sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u)^2 + 2\sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u)(u^T b_{i\Gamma}) + \sum_{i=1}^m (u^T b_{i\Gamma})^2. \end{aligned}$$

Analog to the proof of (11), we can show that the assumption $\sum_{i=1}^m b_i \otimes A_i = 0$ implies $2\sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u)(u^T b_{i\Gamma}) = 0$ and therefore

$$f_{2\Gamma}(u) = \sum_{i=1}^m (u^T A_i^{\Gamma\Gamma} u)^2 + u^T (\sum_{i=1}^m b_{i\Gamma} b_{i\Gamma}^T) u.$$

It follows form the s-regularity that $f_{2\Gamma}(u) \to \infty$ as $u \to \infty$. ■

Proof of Theorem III.1: We first prove the existence of $\hat{x}$. By Lemma B.2 it is easy to show that $f(x) \to \infty$ as $\|x\| \to \infty$ and $\|x\|_0 \leq s$. It follows that there exists a positive constant $\hat{r}$ such that the problem (4) is equivalent to

$$\min_{x\in\mathbb{R}^n} f(x) \quad s.t. \quad \|x\|_0 \leq s,\ \|x\| \leq \hat{r}.$$

Since f is continuous and the constrained set is compact, it follows that the problem (12) has a minimizer $\hat{x}$ which is also a solution of (4).

We now prove the fixed point equation (5). For any $\tau > 0$, define $F_\tau(x, \hat{x}) := f(\hat{x}) + \langle \nabla f(\hat{x}), x - \hat{x}\rangle + \frac{1}{2\tau}\|x - \hat{x}\|_2^2$ and consider the following auxiliary problem

$$\begin{aligned} \min \quad & F_\tau(x, \hat{x}) \\ s.t. \quad & \|x\|_0 \leq s, \\ & x \in \mathbb{R}^n. \end{aligned} \tag{12}$$

Denote $B_{\hat{r},s} = \{x \in \mathbb{R}^n : \|x\|_2 \leq \hat{r}, \|x\|_0 \leq s\}$ and $B_{\hat{r},2s} = \{x \in \mathbb{R}^n : \|x\|_2 \leq \hat{r}, \|x\|_0 \leq 2s\}$. It is clear that there exists a positive constant $\hat{L}$ such that $\hat{L} = \sup_{x\in B_{\hat{r},2s}} \|\nabla^2 f(x)\|_2$. Note that for any $x, y \in B_{\hat{r},s}$ the line segment $[x, y] \in B_{\hat{r},2s}$. Therefore for any $\tau \in (0, \hat{L}^{-1}]$ and $x \in B_{\hat{r},s}$, we have

$$\begin{aligned} & f(x) \qquad\qquad (13) \\ =\ & f(\hat{x}) + \langle \nabla f(\hat{x}), x - \hat{x}\rangle + \frac{1}{2}(x - \hat{x})^T \nabla^2 f(\xi)(x - \hat{x}) \\ =\ & F_\tau(x, \hat{x}) + \frac{1}{2}(x - \hat{x})^T \nabla^2 f(\xi)(x - \hat{x}) - \frac{1}{2\tau}\|x - \hat{x}\|_2^2 \\ \leq\ & F_\tau(x, \hat{x}) + \frac{1}{2}\|\nabla^2 f(\xi)\|_2 \|x - \hat{x}\|_2^2 - \frac{1}{2\tau}\|x - \hat{x}\|_2^2 \\ \leq\ & F_\tau(x, \hat{x}) + \frac{\hat{L}}{2}\|x - \hat{x}\|_2^2 - \frac{1}{2\tau}\|x - \hat{x}\|_2^2 \\ \leq\ & F_\tau(x, \hat{x}), \qquad\qquad (14) \end{aligned}$$

where $\xi = \hat{x} + \alpha(x - \hat{x})$ for some $\alpha \in (0, 1)$ and the second inequality follows from the fact that $\xi \in B_{\hat{r},2s}$ and hence $\|\nabla^2 f(\xi)\|_2 \leq \hat{L}$.

Further, let

$$\bar{x} \in \arg\min_{x\in\mathbb{R}^n} F_\tau(x, \hat{x}) \quad s.t. \quad \|x\|_0 \leq s,$$

where $\tau \in (0, \hat{L}^{-1}]$. Since $f(\hat{x}) = F_\tau(\hat{x}, \hat{x})$, one can conclude from the inequality (13) that for any $\tau \in (0, \hat{L}^{-1}]$,

$$F_\tau(\bar{x}, \hat{x}) \leq F_\tau(\hat{x}, \hat{x}) = f(\hat{x}) \leq f(\bar{x}) \leq F_\tau(\bar{x}, \hat{x}),$$

which leads to $F_\tau(\hat{x}, \hat{x}) = F_\tau(\bar{x}, \hat{x})$. Therefore $\hat{x}$ is also a minimizer of the problem (12).

On the other hand, it is easy to check that the problem (12) is equivalent to the following minimization problem

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x - (\hat{x} - \tau\nabla f(\hat{x}))\|_2^2 \quad s.t. \quad \|x\|_0 \leq s$$

which together with Lemma B.1 leads to the desired result. ■

Lemma B.3. Let $g_k = \|\nabla f(x^k)\|_2$ and $G_k = \sup_{\beta\in B_k} \|\nabla^2 f(x)\|_2$, where $B_k = \{x \in \mathbb{R}^n : \|x\|_2 \leq \|x^k\|_2 + g_k\}$. For any $\delta > 0, \gamma, \alpha \in (0, 1)$, define

$$j_k = \begin{cases} 0, & \text{if } \gamma(G_k + \delta) \leq 1; \\ -[\log_\alpha \gamma(G_k + \delta)] + 1, & \text{otherwise.} \end{cases}$$

Then (6) holds.

Proof: From the definition of τ_k and j_k, it is easy to verify that

$$G_k - \frac{1}{\tau_k} \leq -\delta. \tag{15}$$

Indeed, by taking $\tau_k = \gamma$ we have

$$G_k - \frac{1}{\tau_k} = \frac{\gamma G_k - 1}{\gamma} \leq -\delta,$$

when $\gamma(G_k + \delta) \leq 1$. If $\gamma(G_k + \delta) > 1$, then

$$\tau_k = \gamma\alpha^{j_k} \leq \gamma\alpha^{\log_\alpha \gamma(G_k+\delta)} = \frac{1}{G_k + \delta}$$

which also leads to (15).

Further, since

$$x^{k+1} \in \arg \min_{x\in\mathbb{R}^n,\ \|x\|_0\le s} F_{\tau_k}(x, x^k) \tag{16}$$

and

$$\|x^{k+1}\|_2 \le \|x^k - \tau_k \nabla f(x^k)\|_2 \le \|x^k\|_2 + g_k,$$

it follows that $x^{k+1} \in B_k$. Similar to (13), we obtain from (15) that

$$\begin{aligned} & f(x^{k+1}) \qquad (17)\\ \le\ & F_{\tau_k}(x^{k+1}, x^k) + \frac{1}{2}\|x^{k+1} - x^k\|_2^2\big(\|\nabla^2 f(\xi_k)\|_2 - \frac{1}{\tau_k}\big)\\ \le\ & F_{\tau_k}(x^{k+1}, x^k) + \frac{1}{2}\|x^{k+1} - x^k\|_2^2(G_k - \frac{1}{\tau_k})\\ \le\ & F_{\tau_k}(x^{k+1}, x^k) - \frac{\delta}{2}\|x^{k+1} - x^k\|_2^2, \end{aligned}$$

where $\xi_k = x^k + \varrho(\tilde{x}^{k,t} - x^k)$ for some $\varrho \in (0, 1)$ and hence the second inequality follows from $\xi_k \in B_k$. Combining this and (16), we have

$$\begin{aligned} f(x^k) - f(x^{k+1}) &= F_{\tau_k}(x^k, x^k) - f(x^{k+1}) \qquad (18)\\ &\ge F_{\tau_k}(x^{k+1}, x^k) - f(x^{k+1})\\ &\ge \frac{\delta}{2}\|x^{k+1} - x^k\|_2^2, \end{aligned}$$

which completes the proof. ∎

Lemma B.4. *Let $\{x^k\}$ and $\{\tau_k\}$ be generated by the algorithm. Assume that $\mathcal{A}$ is uniformly $2s$-regular. Then,*
(i) $\{x^k\}$ is bounded;
(ii) there is a nonnegative integer $\bar{j}$ such that $\tau_k \in [\gamma\alpha^{\bar{j}}, \gamma]$.

Proof: We prove the result (i) by contradiction. Suppose $\{x^k\}$ is bounded, which implies that there exists a subsequence $\{x^{k_j}\}$ tending to infinity as $j \to \infty$. By Lemma B.3, we have $f(x^{k_j}) \to \infty$ as $j \to \infty$. On the other hand, Lemma B.3 implies that $\{f(x^k)\}$ is strictly decreasing which together with $f(\cdot) \ge 0$ implies that $\{f(x^k)\}$ converges to a constant $f^*(\ge 0)$, which is a contradiction.

To show (ii), we note that since $f(\cdot)$ is a twice continuous differentiable function, it follows from the boundedness of $\{x^k\}$ that there exist two positive constants $\bar{g}$ and $\bar{G}$ such that $\sup_{k\ge0}\{g_k\} \le \bar{g}$ and $\sup_{k\ge0}\{G_k\} \le \bar{G}$. Define $\bar{j} = \max(0, [-\log_\alpha \gamma(\bar{G} + \delta)] + 1)$. Therefore $0 \le j_k \le \bar{j}$ and it follows from the definition of τ_k that $\tau_k \in [\gamma\alpha^{\bar{j}}, \gamma]$. ∎

Proof of Theorem III.2: (i) From the definition of x^{k+1} and (6), we have

$$\begin{aligned} \sum_{k=0}^{n} \|x^{k+1} - x^k\|_2^2 &\le \frac{2}{\delta}\sum_{k=0}^{n}[f(x^k) - f(x^{k+1})]\\ &= \frac{2}{\delta}[f(x^0) - f(x^{n+1})]\\ &\le \frac{2}{\delta} f(x^0). \end{aligned}$$

Hence, $\sum_{k=0}^{\infty} \|x^{k+1} - x^k\|_2^2 < \infty$ and therefore $\|x^{k+1} - x^k\|_2 \to 0$ as $k \to \infty$.

(ii) Since $\{x^k\}$ is bounded, it has at least one accumulation point. Let $\tilde{x}$ be an accumulation point and suppose that the subsequence $\{x^{k_j}\}$ tends to $\tilde{x}$. We show that it satisfies (5) for some $\tau > 0$. If $\|\tilde{x}\|_0 < s$, then $\nabla f(\tilde{x}) = 0$ and therefore it follows from $\|\tilde{x}\|_0 < s$ that for any $\tau > 0$,

$$\tilde{x} = P_S(\tilde{x}) = P_S(\tilde{x} - \tau\nabla f(\tilde{x})).$$

Denote $\tilde{\Gamma} = \mathrm{supp}(\tilde{x})$ and let x_i be the ith element of a vector x . If $\|\tilde{x}\|_0 = s$, then $x_i^{k_j} \ne 0$ for large enough j when $i \in \tilde{\Gamma}$. It follows from the property of the projection $P_S(\cdot)$ and the algorithm that

$$x_i^{k_j+1} = x_i^{k_j} - \tau_{k_j}\nabla f(x^{k_j})_i.$$

Since $x_i^{k_j} \to \tilde{x}_i$ and ∇f is continuous, by the first result (i) and Lemma B.4 (ii) we have

$$\nabla f(\tilde{x})_i = 0, \quad \text{for each } i \in \tilde{\Gamma}, \tag{19}$$

which together with $\|\tilde{x}\|_0 = s$ yields that for any $\tau > 0$,

$$P_S(\tilde{x}_i - \tau\nabla f(\tilde{x})_i) = P_S(\tilde{x}_i) = \tilde{x}.$$

Since $\|x_i^{k_j}\|_0 \le s$, $x_i^{k_j} \to \tilde{x}$ and $\|\tilde{x}\|_0 = s$, it follows that $x_i^{k_j} = 0$ for large enough k_j when $i \notin \tilde{\Gamma}$. Combing this and the iterative formula, we obtain that there exists a positive $\underline{\tau}$ such that

$$\begin{aligned} \tau_{k_j}|\nabla f(x^{k_j})_i| &= |x_i^{k_j} - \tau_{k_j}\nabla f(x^{k_j})_i|\\ &< \min_{l\in\tilde{\Gamma}} |x_l^{k_j} - \tau_{k_j}\nabla f(x^{k_j})_l|\\ &= \min_{l\in\tilde{\Gamma}} |x_l^{k_j+1}| \end{aligned}$$

which further implies that

$$\limsup_{j\to\infty} \tau_{k_j}|\nabla f(x^{k_j})_i| \le \min_{l\in\tilde{\Gamma}} \limsup_{j\to\infty} |x_l^{k_j+1}|.$$

By Lemma B.4 (ii) and that $x^{k_j} \to \tilde{x}$, we have

$$\limsup_{j\to\infty} \tau_{k_j}|\nabla f(\tilde{x})_i| \le \min_{l\in\tilde{\Gamma}} |\tilde{x}_l| = M_s(\tilde{x}).$$

On the other hand, (19) implies that for any $\tau > 0$,

$$\min_{l\in\tilde{\Gamma}} |\tilde{x}_l - \tau\nabla f(\tilde{x})_l| = \min_{l\in\tilde{\Gamma}} |\tilde{x}_l|.$$

From the above inequalities, we conclude that for each $i \notin \tilde{\Gamma}$,

$$\begin{aligned} |\tilde{x}_i - \tau\nabla f(\tilde{x})_i| &< |\tilde{x}_i - (\limsup_{j\to\infty} \tau_{k_j})\nabla f(\tilde{x})_i|\\ &\le \min_{l\in\tilde{\Gamma}} |\tilde{x}_l - \tau\nabla f(\tilde{x})_l| \end{aligned}$$

for any $\tau \in (0, \limsup_{j\to\infty} \tau_{k_j})$. For each $i \notin \tilde{\Gamma}$, we then obtain that $P_S(\tilde{x}_i - \tau\nabla f(\tilde{x})_i) = 0$ for any $\tau \in (0, \liminf_{k\to\infty} \tau_k)$. ∎