Computing One-bit Compressive Sensing via Double-Sparsity Constrained Optimization

Shenglong Zhou, Ziyan Luo, and Naihua Xiu

***Abstract*—One-bit compressive sensing is popular in signal processing and communications due to the advantage of its low storage costs and hardware complexity. However, it has been a challenging task all along since only the one-bit (the sign) information is available to recover the signal. In this paper, we appropriately formulate the one-bit compressed sensing by a double-sparsity constrained optimization problem. The first-order optimality conditions via the newly introduced τ-stationarity for this nonconvex and discontinuous problem are established, based on which, a gradient projection subspace pursuit (GPSP) approach with global convergence and fast convergence rate is proposed. Numerical experiments against other leading solvers illustrate the high efficiency of our proposed algorithm in terms of the computation time and the quality of the signal recovery as well.**

***Index Terms*—One-bit compressive sensing, double-sparsity constrained optimization, optimality conditions, gradient projection subspace pursuit, global convergence**

I. Introduction

COMPRESSIVE sensing (CS) has seen evolutionary advances in theory and algorithms in the past few decades since it was introduced in the ground-breaking papers [1], [2], [3]. It aims to reconstruct a sparse signal $\mathbf{x}$ from an underdetermined linear systems $\Phi\mathbf{x} = \mathbf{b}$, where $\Phi \in \mathbb{R}^{m\times n}$ is the measurement matrix and $\mathbf{b} \in \mathbb{R}^m$ is the measurement observation. To reduce storage costs and hardware complexity, Boufounos and Baraniuk [4] benefited from the one-bit quantization case of CS, where only the sign information of measurements are preserved, that is

$$\mathbf{c} = \mathrm{sgn}(\Phi\mathbf{x}).$$

Here, $\mathrm{sgn}(t)$ returns one if t is positive and negative one otherwise, and thus $c_i \in \{1, -1\}, i \in [m] := \{1, 2, \cdots, m\}$ is the one-bit measurement. This gives rise to the one-bit CS. It was then extensively applied into applications including communications [5], [6], [7], wireless sensor network [8], [9], [10], [11], cognitive radio [12], [13], and to name a few [14], [15], [16].

The task of one-bit CS constructs the sparse signal from the one-bit measurements. The ideal optimization model is the following ℓ_0-norm minimization,

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0, \quad \text{s.t. } \mathbf{c} = \mathrm{sgn}(\Phi\mathbf{x}), \tag{1}$$

This work is supported in part by the National Natural Science Foundation of China (11971052 and 11771038) and Beijing Natural Science Foundation (Z190002).

Shenglong Zhou is with the Department of EEE, Imperial College London, London SW72AZ, United Kingdom, e-mail: shenglong.zhou@ic.ac.uk.

Ziyan Luo and Naihua Xiu are with the Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, People's Republic of China, e-mail: zyluo@bjtu.edu.cn, nhxiu@bjtu.edu.cn

where $\|\mathbf{x}\|_0$ is the ℓ_0-norm of $\mathbf{x}$, counting the number of its non-zero entries.

A. Related work

An impressive body of work has developed numerical algorithms for solving the problem (1), but most of them placed the interest in its approximations due to the NP-hardness. The earliest work can be traced back to [4] where the ideal model (1) was relaxed by

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1, \quad \text{s.t. } A\mathbf{x} \geq 0, \ \|\mathbf{x}\| = 1. \tag{2}$$

Here, $A := \mathrm{Diag}(\mathbf{c})\Phi$, $\mathrm{Diag}(\mathbf{c})$ represents the diagonal matrix with diagonal entries from $\mathbf{c}$, $\|\cdot\|_1$ is the ℓ_1-norm, and $\|\cdot\|$ is the Euclidean norm. Such a model is valid for the noiseless case. When the measurement $\Phi\mathbf{x}$ is contaminated by noise $\boldsymbol{\varepsilon}$, i.e., $\mathbf{c} = \mathrm{sgn}(\Phi\mathbf{x} + \boldsymbol{\varepsilon})$, a popular approach to reconstruct the signal benefits from the following optimization

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1 + \lambda\varphi(\mathbf{x}), \quad \text{s.t. } \|\mathbf{x}\| = 1, \tag{3}$$

where $\lambda > 0$ and $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a loss function. In [4], they adopted the one-sided ℓ_2 function $\varphi(\mathbf{x}) := \|(-A\mathbf{x})_+\|^2$ with $\mathbf{y}_+ := (\max\{y_1, 0\}, \cdots, \{y_m, 0\})^\top$ and employed a renormalized fixed point iteration algorithm. Since the targeted problem is a nonconvex optimization, the convergence result has not been provided. The same problem was also addressed by a restricted step shrinkage algorithm [17], where the generated sequence was proved to converge to a stationary point of the penalty problem if some slightly strong assumptions on the sequence were satisfied.

Following the work in [4], Boufounos modified CoSaMP [18], one of the most popular greedy methods in CS, to derive the matching sign pursuit method [19]. It turned out to address the sparsity constrained model,

$$\min_{\mathbf{x}\in\mathbb{R}^n} \varphi(\mathbf{x}), \quad \text{s.t. } \|\mathbf{x}\|_0 = s, \ \|\mathbf{x}\| = 1, \tag{4}$$

where $s \ll n$ is a given sparsity level, and φ is the one-sided ℓ_2 function. Based on the framework of the famous iterative hard thresholding algorithm, the modified version BITH was then developed in [20] to solve the problem (4). Apart from the one-sided ℓ_2 function, BITH was also able to process the one-sided ℓ_1 function, namely, $\varphi(\mathbf{x}) = \|(-A\mathbf{x})_+\|_1$. It was claimed that with a high probability, the distance between a reconstructed signal by BIHT and the original one can be bounded by a prefixed accuracy if the former quantizes to the same quantization point as the latter. As a consequence, the method enjoys to a local convergence property. The latest work

on the problem (4) consists of the robust binary iterative hard thresholding [21], the soft consistency reconstructions [22], the binary iterative re-weighted method [23] and the pinball loss iterative hard thresholding [24].

In [25], authors took advantage of the ℓ_1-regularized least squares to deal with the one-bit CS regardless of the sign information of $\Phi\mathbf{x}$, namely,

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1 + \lambda\|\mathbf{c} - \Phi\mathbf{x}\|^2. \tag{5}$$

As stated, with a high probability, the distance between the solution to the model up to a constant and a sparse solution can be bounded by a prefixed accuracy if the sample size m is greater than a threshold. Then a primal dual active set algorithm was proposed to solve the above model and proved to converge within one step under two assumptions: the columns of the matrix Φ indexed on the nonzero components of the sparse solution is full rank and the initial point is sufficiently close to the sparse solution. Therefore, the generated sequence again has a local convergence property. Very recently, the authors in [26] replaced the ℓ_1-norm in (5) by $\|\mathbf{x}\|_p^p := \sum |x_i|^p (0 < p < 1)$ to design a weighted primal dual active set algorithm.

When the number of sign flips $k(\ll m)$ is provided, namely, $\|\mathbf{c} - \mathrm{sgn}(\Phi\mathbf{x})\|_0 = k$, authors in [27] integrated a sparse variable $\mathbf{w}$ to the problem (4). The nonzero components in $\mathbf{w}$ represent the measurements that have sign flips. The resulting optimization problem is

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{w}\in\mathbb{R}^m} \quad & \|(-\mathrm{Diag}(\mathbf{c})(\Phi\mathbf{x}+\mathbf{w}))_+\|_p^p, \\ \text{s.t.} \quad & \|\mathbf{x}\|_0 \le s,\ \|\mathbf{x}\| = 1,\ \|\mathbf{w}\|_0 \le k, \end{aligned} \tag{6}$$

where $p = 1$ or 2. To tackle the above problem, an alternating minimization method (adaptive outliers pursuit, AOP) was cast: solving one variable while fixing the other. However, AOP has been tested to heavily rely on the choice of k and the convergence result remains to be seen. Other work relating to (6) includes the noise-adaptive renormalized fixed point iteration approach [28] and the noise-adaptive restricted step shrinkage [29].

When the number of the sign flips is unavailable, a remedy pursues a solution with sign flips as few as possible, which can be fulfilled by the following one-sided ℓ_0 function minimization [30],

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|(\epsilon\mathbf{1} - A\mathbf{x})_+\|_0 + \eta\|\mathbf{x}\|^2, \quad \text{s.t. } \|\mathbf{x}\|_0 \le s, \tag{7}$$

where η and ϵ are given positive parameters, and the latter is used to majorize the objective function. The first term in the objective function arises from maximizing a posterior estimation from the perspective of statistics. It returns the number of positive components of $(\epsilon\mathbf{1} - A\mathbf{x})$ and can be regarded as the number of the sign flips when ϵ is quite small. Instead of solving the one-sided ℓ_0 model directly, a fixed-point algorithm [30] was created for its approximation,

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{w}\in\mathbb{R}^m} \quad & \|(\epsilon\mathbf{1} - \mathbf{w})_+\|_0 + \mu\|\mathbf{w} - A\mathbf{x}\|^2 + \eta\|\mathbf{x}\|^2, \\ \text{s.t.} \quad & \|\mathbf{x}\|_0 \le s, \end{aligned} \tag{8}$$

where $\mu > 0$. It has shown that the generated sequence converges to a local minimizer of the approximation problem if the maximum singular value of the matrix A is bounded by some chosen parameters. However, the relationship between the solution obtained by the method and the original problem (7) has not been well explored.

Some other numerical algorithms developed to solve the one-bit CS can be seen the convex relaxation [31], the Bayesian approach [32], the passive algorithm [33], the quantized compressed sampling matching pursuit and adaptive outlier pursuit for quantized iterative hard thresholding [34], the superset technique approximation [35], the sparse consistent coding algorithm [36] and ones [37], [38], [39], [40], [41].

B. Our contributions

Inspired by the previous work, we formulate the one-bit CS as the following double-sparsity constrained optimization:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{y}\in\mathbb{R}^m} \quad & \|A\mathbf{x} + \mathbf{y} - \epsilon\mathbf{1}\|^2 + \eta\|\mathbf{x}\|^2, \\ \text{s.t.} \quad & \|\mathbf{x}\|_0 \le s, \quad \|\mathbf{y}_+\|_0 \le k \end{aligned} \tag{9}$$

where $\eta > 0$ is a penalty parameter, $s \ll n$ and $k \ll m$ are two integers representing the prior information on the upper bounds of the signal sparsity and the number of sign flips, respectively. When penalizing the sign flip constraint in our model, it turns to (8) with the auxiliary $\mathbf{y} = \epsilon\mathbf{1} - \mathbf{w}$. It is worth mentioning that the selection of k is very flexible in our approach as will be shown in the numerical experiments, which reveals that our approach has no heavy burden on the pre-knowledge of the true number of sign flips. Our contributions in this paper are threefold:

(1) **The new optimization model.** The double-sparsity constrained optimization problem (9) is formulated to handle the one-bit CS. It is well-known that these two discrete and nonconvex constraints lead to the NP-hardness in general. Nevertheless, a necessary and sufficient optimality condition as stated in (18) for a local minimizer is established, see Lemma 3.1. Moreover, the necessary and sufficient optimality condition for a global minimizer is also studied in terms of the newly introduced τ-stationary point, see Theorem 3.1.

(2) **The efficient `GPSP` algorithm.** As the established optimality conditions indicate a τ-stationary point is instructive to pursue an optimal solution to (9), we thus design a gradient project method with a subspace pursuit scheme interpolated (`GPSP`). The proposed method is proved to be globally convergent to a unique local minimizer without any assumptions, see Theorem 4.1. Furthermore, the produced sequence either enjoys a Q-linear convergence rate or is identical to the limiting point after finite iterations. Particularly, `GPSP` will stop within finite steps once the limiting point reaches the upper bounds of the corresponding double-sparsity, see Theorem 4.2.

(3) **High numerical performance.** `GPSP` is demonstrated to be relatively robust to the parameters k, ϵ, η in (9) in the numerical experiments, which indicates that we do not need an exact upper bound k of the sign flips. Most importantly, `GPSP` outperforms some leading solvers for synthetic data, in both time efficiency and recovery accuracy.

C. Organization

The remainder of the paper is organized as follows. In Section II, some necessary mathematical backgrounds are provided, including the notation and the projection onto the feasible set of the problem (9). Section III is devoted to the optimality conditions of the problem, associated with the τ-stationary points, followed by its relationship to the global minimizers. In Section IV, the gradient projection subspace pursuit (GPSP) method is designed, and the global convergence and the Q-linear convergence rate are established. Numerical experiments are given in Section V, including the involved parameters (s, k, ϵ, η) tuning and comparisons with other state-of-the-art solvers. Concluding remarks are made in Section VI.

II. Preliminaries

We first define some notation employed throughout this paper. To differ from $\mathrm{sgn(t)}$, the sign function is written as $\mathrm{sign(t)}$ that returns 0 if $t = 0$ and $\mathrm{sgn(t)}$ otherwise. Given a subset $T \subseteq [n] := \{1, 2, \cdots, n\}$, its cardinality and complementary set are $|T|$ and $\overline{T} := [n] \setminus T$. For a vector $\mathbf{x} \in \mathbb{R}^n$, the support set $\mathrm{supp}(\mathbf{x})$ represents the indices of nonzero elements of $\mathbf{x}$ and the neighbourhood with a radius $\delta > 0$ is denoted by $N(\mathbf{x}, \delta) := \{\mathbf{w} \in \mathbb{R}^n : \|\mathbf{w} - \mathbf{x}\| < \delta\}$. Let $\|\mathbf{x}\|_{[i]}$ be the ith largest (in absolute) element of $\mathbf{x}$. In addition, $\mathbf{x}_T$ stands for the sub-vector contains elements of $\mathbf{x}$ indexed on T. Similarly, for a matrix $A \in \mathbb{R}^{m\times n}$, $A_{\Gamma T}$ is the sub-matrix containing rows indexed on Γ and columns indexed on T, particularly, $A_{:T} = A_{[m]T}$. Moreover, we merge two vectors $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{z} := (\mathbf{x}; \mathbf{y}) := (\mathbf{x}^\top \ \mathbf{y}^\top)^\top$. For a positive definite matrix H, the H-weighted norm is written $\|\mathbf{z}\|_H^2 = \langle \mathbf{z}, H\mathbf{z}\rangle$, where $\langle \mathbf{z}, \mathbf{z}'\rangle := \sum z_i z_i'$ is the inner product of two vectors. Given a scalar $a \in \mathbb{R}$, $\lceil a \rceil$ returns the smallest integer that is no less than a. For simplicity, denote

$$\begin{array}{rcl} S & := & \{\mathbf{x} \in \mathbb{R}^n : \ \|\mathbf{x}\|_0 \le s\}, \\ K & := & \{\mathbf{y} \in \mathbb{R}^m : \ \|\mathbf{y}_+\|_0 \le k\}. \end{array} \tag{10}$$

The feasible region of (9) is then denoted by $\mathcal{F} := S \times K$, with its interior

$$\mathrm{int}\mathcal{F} := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \ \|\mathbf{x}\|_0 < s, \ \|\mathbf{y}_+\|_0 < k\}.$$

A. Projections

For a nonempty and closed set $\Omega \subseteq \mathbb{R}^n$, the projection $\Pi_\Omega(\mathbf{x})$ of $\mathbf{x} \in \mathbb{R}^n$ onto Ω is given by

$$\Pi_\Omega(\mathbf{x}) = \mathrm{argmin}\ \{\|\mathbf{x} - \mathbf{w}\| : \mathbf{w} \in \Omega\}.$$

By introducing

$$\Sigma(\mathbf{x}; s) := \left\{ T \subseteq [n] : \ \begin{array}{l} |T| = s, \\ |x_i| \ge |x_j|, \forall i \in T, j \notin T \end{array} \right\}, \tag{11}$$

one can easily verify that

$$\Pi_S(\mathbf{x}) = \Big\{ (\mathbf{x}_T; 0) : \ T \in \Sigma(\mathbf{x}; s) \Big\}. \tag{12}$$

To derive the projection of a point $\mathbf{y} \in \mathbb{R}^m$ onto K, denote

$$\begin{array}{rcl} \Gamma_+ & := & \{i \in [m] : \ y_i > 0\}, \\ \Gamma_0 & := & \{i \in [m] : \ y_i = 0\}, \\ \Gamma_- & := & \{i \in [m] : \ y_i < 0\}. \end{array} \tag{13}$$

Note that Γ_+, Γ_0 and Γ_- should depend on $\mathbf{y}$. We drop their dependence if no extra explanations are provided for the sake of notational convenience. Based on the above notation, for a point $\mathbf{y} \in \mathbb{R}^m$ and an integer $k \in [m]$, we define a set by

$$\begin{array}{rcl} & & \Theta(\mathbf{y}; k) \\ & := & \left\{ \Gamma_k \cup \Gamma_- : \ \begin{array}{l} \Gamma_k \subseteq \Gamma_+, |\Gamma_k| = \min\{k, |\Gamma_+|\} \\ y_i \ge y_j \ge 0, \ \forall\, i \in \Gamma_k, \ \forall\, j \notin \Gamma \end{array} \right\}, \end{array} \tag{14}$$

where Γ_+, Γ_0 and Γ_- be given by (13). One can observe that $\Gamma \in \Theta(\mathbf{y}; k)$ consists of the indices of all negative elements and the first $\min\{k, |\Gamma_+|\}$ largest positive elements of $\mathbf{y}$. These notation allow us to derive the projection $\Pi_K(\mathbf{y})$ by

$$\Pi_K(\mathbf{y}) = \Big\{ (\mathbf{y}_\Gamma; 0) : \ \Gamma \in \Theta(\mathbf{y}; k) \Big\}. \tag{15}$$

A simple example is presented for illustration. Given $\mathbf{y} = (3, 2, 2, 0, -2)^\top$, we have

$$\begin{array}{rcl} \Theta(\mathbf{y}; 3) & = & \Big\{\{1, 2, 3, 5\}\Big\}, \ \ \Pi_K(\mathbf{y}) = \{\mathbf{y}\}, \\ \Theta(\mathbf{y}; 2) & = & \Big\{\{1, 2, 5\}, \{1, 3, 5\}\Big\}, \\ \Pi_K(\mathbf{y}) & = & \Big\{(3, 2, 0, 0, -2)^\top, (3, 0, 2, 0, -2)^\top\Big\}. \end{array}$$

B. Properties of the objective function

To end this section, we present some properties of the objective function in (9) which can be written as follows

$$\begin{array}{rcl} f(\mathbf{x}, \mathbf{y}) & := & \|A\mathbf{x} + \mathbf{y} - \epsilon\mathbf{1}\|^2 + \eta\|\mathbf{x}\|^2 \\ & = & \|\mathbf{z}\|_H^2 - 2\langle \mathbf{b}, \mathbf{z}\rangle + m\epsilon^2 \\ & =: & f(\mathbf{z}), \end{array} \tag{16}$$

where the matrix H and the vector $\mathbf{b}$ are given by

$$H := \frac{1}{2}\nabla^2 f(\mathbf{z}) = \begin{bmatrix} A^\top A + \eta I & A^\top \\ A & I \end{bmatrix}, \ \mathbf{b} = \epsilon \begin{bmatrix} A^\top \mathbf{1} \\ \mathbf{1} \end{bmatrix}.$$

It is easy to verify that H is symmetric positive definite and hence has all eigenvalues positive. Denote the smallest and the largest eigenvalues by $\lambda_{\min}$ and $\lambda_{\max}$, respectively. The quadratic objective function f is then strongly convex and strongly smooth since for any $\mathbf{z}$ and $\mathbf{z}'$ in $\mathbb{R}^{n+m}$,

$$\begin{array}{rcl} & & f(\mathbf{z}) - f(\mathbf{z}') - \langle \nabla f(\mathbf{z}'), \mathbf{z} - \mathbf{z}'\rangle \\ & = & \|\mathbf{z} - \mathbf{z}'\|_H^2 \ \in \ \left[\lambda_{\min}\|\mathbf{z} - \mathbf{z}'\|^2, \ \lambda_{\max}\|\mathbf{z} - \mathbf{z}'\|^2\right]. \end{array} \tag{17}$$

III. Optimality Conditions

The first-order necessary and sufficient optimality conditions for (9) are established in this section.

Lemma 3.1: Consider a point $\mathbf{z}^* := (\mathbf{x}^*; \mathbf{y}^*) \in \mathcal{F}$ with

$$T_* := \mathrm{supp}(\mathbf{x}^*), \quad \Gamma_* := \mathrm{supp}(\mathbf{y}^*).$$

If $\mathbf{z}^*$ is a local minimizer of (9), then it must satisfy

$$\begin{array}{rl} \nabla_{\mathbf{x}} f(\mathbf{z}^*) = 0, & \text{if } \|\mathbf{x}^*\|_0 < \mathrm{s}, \\ (\nabla_{\mathbf{x}} f(\mathbf{z}^*))_{T_*} = 0, & \text{if } \|\mathbf{x}^*\|_0 = \mathrm{s}, \\ \nabla_{\mathbf{y}} f(\mathbf{z}^*) = 0, & \text{if } \|\mathbf{y}_+^*\|_0 < \mathrm{k}, \\ (\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\Gamma_*} = 0, \ (\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\overline{\Gamma}_*} \le 0, & \text{if } \|\mathbf{y}_+^*\|_0 = \mathrm{k}, \end{array} \tag{18}$$

Conversely, let $\mathbf{z}^*$ satisfy (18). Then it is the unique global minimizer if $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$ and the unique local minimizer

otherwise. Furthermore, there is a $\delta_* > 0$ satisfying the following quadratic growth property

$$f(\mathbf{z}) - f(\mathbf{z}^*) \geq \|\mathbf{z} - \mathbf{z}^*\|_H^2, \quad \forall\, \mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \delta_*). \tag{19}$$

The above lemma shows the optimality conditions of a point being a local minimizer. We further establish the conditions for a global minimizer. To do that, we introduce a τ-stationary point. A point $\mathbf{z}^* := (\mathbf{x}^*; \mathbf{y}^*)$ is called a τ-stationary point of (9) with some $\tau > 0$ if it satisfies

$$\mathbf{z}^* \in \Pi_{\mathcal{F}}\left(\mathbf{z}^* - \tau \nabla f(\mathbf{z}^*)\right). \tag{20}$$

An equivalent characterization is presented as follows.

Lemma 3.2: A point $\mathbf{z}^*$ is a τ-stationary point of the problem (9) with some $\tau > 0$ if and only if it satisfies

$$\begin{array}{l} \|\mathbf{x}^*\|_0 \leq s, \\ \tau(\nabla_{\mathbf{x}} f(\mathbf{z}^*))_i \left\{ \begin{array}{ll} = 0, & i \in T_*, \\ \in [-\|\mathbf{x}^*\|_{[s]}, \|\mathbf{x}^*\|_{[s]}], & i \notin T_*, \end{array} \right. \\ \|\mathbf{y}_+^*\|_0 \leq k, \\ \tau(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i \left\{ \begin{array}{ll} = 0, & i \in \Gamma_*, \\ \in [-\|\mathbf{y}_+^*\|_{[k]}, 0], & i \notin \Gamma_*. \end{array} \right. \end{array} \tag{21}$$

The following theorem reveals the relationships between τ-stationary points and global minimizers of the problem (9).

Theorem 3.1: For (9) and a point $\mathbf{z}^* \in \mathcal{F}$, the following statements hold.

a) For $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$, the point $\mathbf{z}^*$ is a global minimizer if and only if it is a τ-stationary point with $\tau > 0$.
b) For $\mathbf{z}^* \notin \mathrm{int}\mathcal{F}$, a global minimizer $\mathbf{z}^*$ is a τ-stationary point with $0 < \tau \leq 1/(2\lambda_{\max})$, and conversely, a τ-stationary point with $\tau \geq 1/(2\lambda_{\min})$ is also a global minimizer.

IV. Gradient Projection Subspace Pursuit

A gradient projection method with a subspace pursuit strategy is proposed to handle the problem (9) by seeking a τ-stationary point. For notational simplicity, hereafter, for a parameter $\tau \in (0, 1]$, we always let

$$\mathbf{z}^\ell(\tau) = \left[\mathbf{x}^\ell(\tau);\ \mathbf{y}^\ell(\tau)\right] \in \Pi_{\mathcal{F}}(\mathbf{z}^\ell - \tau \nabla f(\mathbf{z}^\ell)), \tag{22}$$

for the ℓth iteration $\mathbf{z}^\ell := (\mathbf{x}^\ell; \mathbf{y}^\ell)$. Analogous to the Γ-related indices defined for $\mathbf{y}$ in (13), we also define

$$\begin{array}{rcl} \widetilde{\Gamma}_+^\ell & := & \{i \in [m] : (\mathbf{y}^\ell(\tau_\ell))_i > 0\}, \\ \Gamma_+^\ell & := & \left\{i \in [m] :\ y_i^\ell > 0\right\}, \\ \Gamma_0^\ell & := & \left\{i \in [m] :\ y_i^\ell = 0\right\}, \\ \Gamma_-^\ell & := & \left\{i \in [m] :\ y_i^\ell < 0\right\}. \end{array} \tag{23}$$

We denote the support sets for $\mathbf{x}^\ell$ and $\mathbf{x}^\ell(\tau_\ell)$ as follows

$$T^\ell := \mathrm{supp}(\mathbf{x}^\ell), \quad \widetilde{T}^\ell := \mathrm{supp}(\mathbf{x}^\ell(\tau_\ell)). \tag{24}$$

Given $\mathbf{z}^\ell = (\mathbf{x}^\ell; \mathbf{y}^\ell) \in \mathcal{F}$, define the following subspace

$$\Omega(\mathbf{z}^\ell) := \left\{ \mathbf{z} = \left[\begin{array}{c} \mathbf{x} \\ \mathbf{y} \end{array} \right] :\ \begin{array}{l} \mathbf{x}_{\overline{T}^\ell} = 0, \\ \mathbf{y}_{\Gamma_0^\ell} = 0,\ \mathbf{y}_{\Gamma_-^\ell} \leq 0 \end{array} \right\}. \tag{25}$$

It is easy to see that $\Omega(\mathbf{z}^\ell) \subseteq \mathcal{F}$. Based on these notation, we summarize the framework of the proposed method in Algorithm 1.

Algorithm 1 GPSP: Gradient projection subspace pursuit

Initialize $\mathbf{z}^0 \in \mathcal{F}, \beta \in (0, 1)$ and $\rho, \varepsilon, \mathtt{tol}_0 > 0$. Set $\ell := 0$.
while $\mathtt{tol}_\ell > \varepsilon$ **do**
  Gradient descent: Find the smallest integer $\sigma = 0, 1, 2, \cdots$ such that

$$f(\mathbf{z}^\ell(\beta^\sigma)) \leq f(\mathbf{z}^\ell) - \rho\|\mathbf{z}^\ell(\beta^\sigma) - \mathbf{z}^\ell\|^2. \tag{26}$$

  Set $\tau_\ell = \beta^\sigma$, $\mathbf{u}^\ell := \mathbf{z}^\ell(\tau_\ell)$ and $\mathbf{z}^{\ell+1} = \mathbf{u}^\ell$.
  Subspace pursuit: **if** $T^\ell = \widetilde{T}^\ell$ *and* $\Gamma_+^\ell = \widetilde{\Gamma}_+^\ell$ **then**

$$\mathbf{v}^\ell = \mathrm{argmin}\ \{f(\mathbf{z}) : \mathbf{z} \in \Omega(\mathbf{z}^\ell)\}. \tag{27}$$

    If $f(\mathbf{v}^\ell) \leq f(\mathbf{u}^\ell) - \rho\|\mathbf{v}^\ell - \mathbf{u}^\ell\|^2$, then set $\mathbf{z}^{\ell+1} = \mathbf{v}^\ell$.
  end
  Compute $\mathtt{tol}_\ell := \|\mathbf{u}^\ell - \mathbf{z}^\ell\|$ and set $\ell := \ell + 1$.
end
Output the solution $\mathbf{x}^* = \mathbf{x}^\ell / \|\mathbf{x}^\ell\|$.

Observing that the initial point $\mathbf{z}^0 \in \mathcal{F}$, and $\Omega(\mathbf{z}^\ell) \subseteq \mathcal{F}$, we can see that all iterations are feasible. Particularly, if the gap $\mathtt{tol}_\ell = \|\mathbf{z}^\ell - \mathbf{u}^\ell\|$ vanishes, then

$$\mathbf{z}^\ell = \mathbf{u}^\ell \in \Pi_{\mathcal{F}}(\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell)),$$

which indicates that $\mathbf{z}^\ell$ is a τ-stationary point with $\tau \leq \tau_\ell$. Additionally, once the conditions $T^\ell = \widetilde{T}^\ell$ and $\Gamma_+^\ell = \widetilde{\Gamma}_+^\ell$ are satisfied, we have $\mathbf{u}^l \in \Omega(\mathbf{z}^\ell)$. The unique minimizer $\mathbf{v}^l$ of f over $\Omega(\mathbf{z}^\ell)$ implies that

$$\langle \nabla f(\mathbf{v}^\ell), \mathbf{u}^\ell - \mathbf{v}^\ell \rangle \geq 0. \tag{28}$$

In virtue of (17), we have

$$f(\mathbf{v}^\ell) \quad \leq \quad f(\mathbf{u}^\ell) - \lambda_{\min}\|\mathbf{u}^\ell - \mathbf{v}^\ell\|^2. \tag{29}$$

Suppose $0 < \rho \leq \lambda_{\min}$. Then the candidate $\mathbf{v}^\ell$ will be taken, namely, $\mathbf{z}^{\ell+1} = \mathbf{v}^\ell$.

A. Computational complexity analysis

To update $\mathbf{u}^\ell$, we need to select one point from $\Pi_{\mathcal{F}}(\mathbf{z}^\ell - \tau\nabla f(\mathbf{z}^\ell))$. Namely, three quantities are computed: $\bar{\mathbf{z}}^\ell := (\bar{\mathbf{x}}^\ell; \bar{\mathbf{y}}^\ell) := \mathbf{z}^\ell - \tau\nabla f(\mathbf{z}^\ell)$, $\Pi_S(\bar{\mathbf{x}}^\ell)$ and $\Pi_K(\bar{\mathbf{y}}^\ell)$. For the former, the computational complexity is about $O(mn)$. To select one point from $\Pi_S(\bar{\mathbf{x}}^\ell)$, we only pick the first s largest (in absolute) elements of $\bar{\mathbf{x}}^\ell$. This allows us to use a MATLAB built-in function `maxk` whose computational complexity is $O(n + s\log s)$. Similarly, for $\Pi_K(\bar{\mathbf{y}}^\ell)$, the computational complexity is $O(m + k\log k)$. Thus, updating $\mathbf{u}^\ell$ takes a computational complexity of order $O(\sigma mn)$, where σ is the smallest integer satisfying (26).

To update $\mathbf{v}^\ell$, we solve a quadratic programming,

$$\begin{array}{rll} \mathbf{v}^\ell = \underset{(\mathbf{x};\mathbf{y})}{\mathrm{argmin}} & \|A\mathbf{x} + \mathbf{y} - \epsilon\mathbf{1}\|^2 + \eta\|\mathbf{x}\|^2 \\ \mathrm{s.t.} & \mathbf{x}_{\overline{T}^\ell} = 0,\ \mathbf{y}_{\Gamma_0^\ell} = 0,\ \mathbf{y}_{\Gamma_-^\ell} \leq 0, \end{array} \tag{30}$$

for the fixed T^ℓ, Γ_0^ℓ and Γ_-^ℓ. Any solvers for solving the quadratic programming can be used to solve (30) to pursue a solution in good quality. To further reduce the computation

cost, we drop the constraint $\mathbf{y}_{\Gamma_-^\ell} \le 0$ from (30), and simply solve the system of equations:

$$\mathbf{x}_{\overline{T}^\ell} = 0, \quad \mathbf{y}_{\Gamma_0^\ell} = 0,$$
$$\begin{bmatrix} A_{:T^\ell}^\top A_{:T^\ell} + \eta I & A_{\overline{\Gamma_0^\ell} T^\ell}^\top \\ A_{\overline{\Gamma_0^\ell} T^\ell} & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T^\ell} \\ \mathbf{y}_{\overline{\Gamma_0^\ell}} \end{bmatrix} = \begin{bmatrix} A_{:T^\ell}^\top \epsilon \mathbf{1} \\ \epsilon \mathbf{1} \end{bmatrix}.$$

The solution $(\widetilde{\mathbf{x}}^\ell; \widetilde{\mathbf{y}}^\ell)$ can be derived by

$$\begin{aligned} \widetilde{\mathbf{x}}_{T^\ell}^\ell &= \left[A_{\Gamma_0^\ell T^\ell}^\top A_{\Gamma_0^\ell T^\ell} + \eta I\right]^{-1} \left[A_{\Gamma_0^\ell T^\ell}^\top \epsilon \mathbf{1}\right], \\ \widetilde{\mathbf{x}}_{\overline{T}^\ell}^\ell &= 0, \\ \widetilde{\mathbf{y}}_{\Gamma_0^\ell}^\ell &= 0, \\ \widetilde{\mathbf{y}}_{\overline{\Gamma_0^\ell}}^\ell &= \epsilon \mathbf{1} - A_{\overline{\Gamma_0^\ell} T^\ell} \widetilde{\mathbf{x}}_{T^\ell}^\ell. \end{aligned}$$

If $\widetilde{\mathbf{y}}_{\Gamma_-^\ell}^\ell \le 0$, namely, $(\widetilde{\mathbf{x}}^\ell; \widetilde{\mathbf{y}}^\ell)$ is the solution to (30), then we set $\mathbf{v}^\ell = (\widetilde{\mathbf{x}}^\ell; \widetilde{\mathbf{y}}^\ell)$. Otherwise, this point will not be taken into consideration, and we set $\mathbf{z}^{\ell+1} = \mathbf{u}^\ell$. The computational complexity of addressing the above equations is about $O(ms^2 + s^3)$.

Overall, the computational complexity of each iteration is

$$O(\sigma mn + ms^2 + s^3).$$

B. Convergence analysis

The first result shows that the Armijo-type step size $\{\tau_\ell\}$ is well defined.

Lemma 4.1: For any $0 < \tau \le \frac{1}{2\rho + 2\lambda_{\max}}$, it holds that

$$f(\mathbf{z}^\ell(\tau)) \le f(\mathbf{z}^\ell) - \rho \|\mathbf{z}^\ell(\tau) - \mathbf{z}^\ell\|^2, \tag{31}$$

and thus $\inf_{\ell \ge 0}\{\tau_\ell\} \ge \underline{\tau} > 0$, where

$$\underline{\tau} := \min\left\{1, \ \frac{\beta}{2\rho + 2\lambda_{\max}}\right\}. \tag{32}$$

Lemma 4.2: Let $\{\mathbf{z}^\ell\}$ be the sequence generated by GPSP and $\underline{\tau}$ be given by (32). Then the following results hold.

a) The sequence $\{\mathbf{z}^\ell\}$ is bounded and $\lim_{\ell\to\infty} \|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\| = \lim_{\ell\to\infty} \|\mathbf{u}^\ell - \mathbf{z}^\ell\| = 0$.

b) Any accumulating point of the sequence $\{\mathbf{z}^\ell\}$ is a τ-stationary point with $0 < \tau \le \underline{\tau}$ of the problem (9).

The above lemma allows us to conclude that the whole sequence converges.

Theorem 4.1: Let $\{\mathbf{z}^\ell\}$ be the sequence generated by GPSP. Then the whole sequence converges to $\mathbf{z}^*$, which is necessarily the unique global minimizer of (9) if $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$ and the unique local minimizer otherwise.

The following theorem establishes the convergence rate that GPSP either enjoys a Q-linear convergence rate or terminates at the limit of the sequence after a certain point.

Theorem 4.2: Let $\{\mathbf{z}^\ell\}$ be the sequence generated by GPSP and $\mathbf{z}^*$ be the limit.

i) For sufficiently large ℓ, it follows

$$\|\mathbf{u}^\ell - \mathbf{z}^*\|^2 \le \frac{1 - 2\underline{\tau}\lambda_{\min}}{1 + 2\underline{\tau}\lambda_{\min}} \|\mathbf{z}^\ell - \mathbf{z}^*\|^2. \tag{33}$$

ii) The sequence either has infinitely many and sufficiently large ℓ satisfying

$$\|\mathbf{z}^{\ell+1} - \mathbf{z}^*\|^2 \le \frac{1 - 2\underline{\tau}\lambda_{\min}}{1 + 2\underline{\tau}\lambda_{\min}} \|\mathbf{z}^\ell - \mathbf{z}^*\|^2, \tag{34}$$

or remains identical to $\mathbf{z}^*$ after a certain point, say $\hat{\ell} \ge 1$, namely,

$$\mathbf{z}^\ell = \mathbf{v}^{\hat{\ell}} = \mathbf{z}^*, \quad \forall\ \ell > \hat{\ell}. \tag{35}$$

iii) Let ρ be chosen as $0 < \rho \le \lambda_{\min}$. If the limit $\mathbf{z}^*$ satisfies $\|\mathbf{x}^*\|_0 = s$ and $\|\mathbf{y}_+^*\|_0 = k$, then GPSP will terminate at the limit $\mathbf{z}^*$ within finite steps.

V. Numerical Experiments

In this section, we will conduct extensive numerical experiments to showcase the performance of our proposed GPSP (available at https://github.com/ShenglongZhou/GPSP), by using MATLAB (R2019a) on a laptop of 32GB memory and Inter(R) Core(TM) i9-9880H 2.3Ghz CPU.

A. Testing examples

Examples with the data generated from the Gaussian distributions are taken into account.

Example 5.1 (Independent covariance [27], [30]): Entries of $\Phi := [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_m]^\top \in \mathbb{R}^{m\times n}$ and the nonzero entries of the ground-truth s_*-sparse vector $\mathbf{x}^* \in \mathbb{R}^n$ (i.e., $\|\mathbf{x}^*\|_0 \le s_*$) are generated from the independent and identically distributed (i.i.d.) samples of the standard Gaussian distribution $\mathcal{N}(0,1)$. To avoid tiny nonzero entries of $\mathbf{x}^*$, let $x_i^* = x_i^* + \mathrm{sign}(x_i^*)$ for nonzero x_i^*, followed by normalizing $\mathbf{x}^*$ to be a unit vector. Let $\mathbf{c}^* = \mathrm{sgn}(\Phi\mathbf{x}^*)$ and $\tilde{\mathbf{c}} = \mathrm{sgn}(\Phi\mathbf{x}^* + \varepsilon)$, where entries of the noise ε are the i.i.d. samples of $\mathcal{N}(0, 0.1^2)$. Finally, we randomly select $\lceil rm \rceil$ entries in $\tilde{\mathbf{c}}$ and flip their signs, and the flipped vector is denoted by $\mathbf{c}$, where r is the flipping ratio.

Example 5.2 (Correlated covariance [25]): Rows of Φ are generated from the i.i.d. samples of $\mathcal{N}(0, \Sigma)$ with $\Sigma_{ij} = v^{|i-j|}, i, j \in [n]$, where $v \in (0,1)$. Then $\mathbf{x}^*, \mathbf{c}^*$ and $\mathbf{c}$ are generated the same as those in Example 5.1.

To demonstrate the performance of one method, apart from the CPU TIME, we will also report the signal-to-noise ratio (SNR), the hamming error (HE) and the hamming distance (HD). They are defined by

$$\begin{aligned} \texttt{SNR} &:= -20\log_{10}\|\mathbf{x} - \mathbf{x}^*\|, \\ \texttt{HD} &:= (1/m)\|\mathrm{sgn}(\Phi\mathbf{x}) - \mathbf{c}\|_0, \\ \texttt{HE} &:= (1/m)\|\mathrm{sgn}(\Phi\mathbf{x}) - \mathbf{c}^*\|_0, \end{aligned}$$

where $\mathbf{x}$ is the solution obtained by one method. The larger SNR (or the smaller HE or HD) means the better recovery.

B. Implementation and parameters selection

Set the accuracy parameter $\varepsilon = 10^{-4}$ in Algorithm 1. GPSP will terminate if $\|\mathbf{z}^\ell - \mathbf{u}^\ell\| \le \varepsilon$ or $\ell > 2000$. In all experiments, we set $\beta = 0.5$, $\rho = 10^{-6}$ and $\mathbf{z}^0 = 0$. The parameters ϵ, η, s and k in (9) are tuned as follows.

(I) Selection of η. For Example 5.1, fix $n = 500$, $\epsilon = 0.01, s = 5$ and $k = \lceil 0.01m \rceil$ but vary $m \in$

$\{0.25, 0.5, 0.75, 1\}n$ and $\eta \in \{10^{-8}, 10^{-7}, \cdots, 10^{4}\}$. Average results over 200 trials are reported in Figure 1. It can be evidently seen that results are stabilized when $\eta \leq 10$ while getting worse when $\eta > 10$ is rising. Similar trends are also observed for GPSP solving Example 5.2. Therefore, any value in $(0, 10]$ can be used to set η. In fact, GPSP also performs the similar results when $\eta = 0$. For simplicity, we fix $\eta = 10^{-4}$.

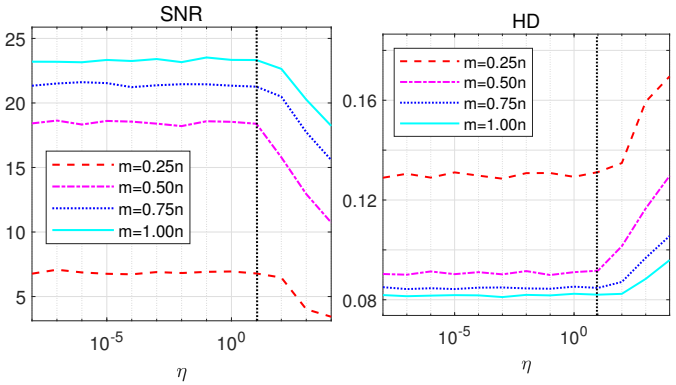


Fig. 1: Effect of η for Example 5.1.

(II) Selection of ϵ. We now fix $n = 500$, $\eta = 10^{-4}, s = 5$ and $k = \lceil 0.01m \rceil$ but alter $m \in \{0.25, 0.5, 0.75, 1\}n$ and $\epsilon \in [10^{-4}, 10]$ for GPSP solving Example 5.1. As shown in Figure 2, the average results over 200 trials do not fluctuate significantly among the changing of ϵ, which indicates GPSP is quite robust to the choices of ϵ in the range $[10^{-4}, 10]$. Similar performance can be also observed for GPSP solving Example 5.2. For simplicity, we fix $\epsilon = 0.01$ in the subsequent numerical experiments.

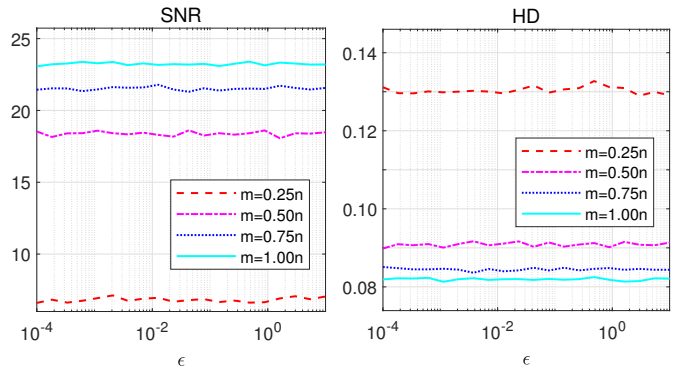


Fig. 2: Effect of ϵ for Example 5.1.

(III) Selection of s. For the sparsity level s, it clearly has a heavy influence of the recovery quality. As shown in Figure 3 where $(n, m, s_*, r) = (500, 250, 5, 0.05)$, the ground-truth signal $\mathbf{x}^*$ has $s_* = 5$ nonzero components with their indices denoted by $T_* = \mathrm{supp}(\mathbf{x}^*)$. Apparently, GPSP gets the most accurate signal if we set $s = s_*$ because it almost exactly recovers those nonzero components. For $s = s_* - 2 = 3$, the recovered signal has 3 nonzero components, however, their indices belong to the true support set T_*. While for $s = s_* + 2$ or $s = s_* + 4$, GPSP generates a solution $\mathbf{x}$ whose support set $\mathrm{supp}(\mathbf{x})$ covers T_* with extra incorrect indices. However, compared with the magnitude $|x_i|, i \in T_*$, those redundant nonzero components $|x_i|, i \in \mathrm{supp}(\mathbf{x}) \setminus T_*$ are pretty small. If we remove those small parts and normalize the signal to have a unit length, then the new signal is much closer to $\mathbf{x}^*$. For simplicity, we set $s = s_*$ in the sequel.

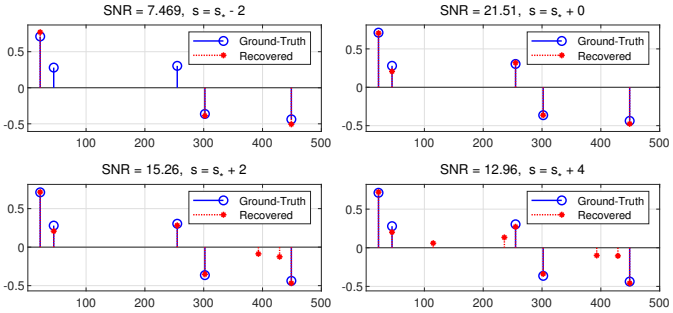


Fig. 3: Effect of s for Example 5.1.

(IV) Selection of k. Note that k is the upper bound of the number of sign flips of $\Phi\mathbf{x}$, and is usually unknown beforehand. However, the model (9) does not require an exact k. One could either fix it by a small integer (e.g. $k = \lceil 0.01m \rceil$) or start with a slightly bigger value and reduce it iteratively. We tested GPSP for solving Example 5.1 and Example 5.2 under both schemes and corresponding numerical performance does not have a big difference. For instance, as indicated in Figure 4, where $(n, m, s_*, v) = (500, 250, 5, 0.5)$, we select $k/m \in \{0.01, 0.03, 0.05, 0.07, 0.09\}$ and then fix it for GPSP. Evidently, for each case of the flipping ratio r, the results SNR and HD do not vary significantly along with k altering. Hence, in our numerical experiments, we pick $k = \lceil 0.01m \rceil$ if no additional information is provided.

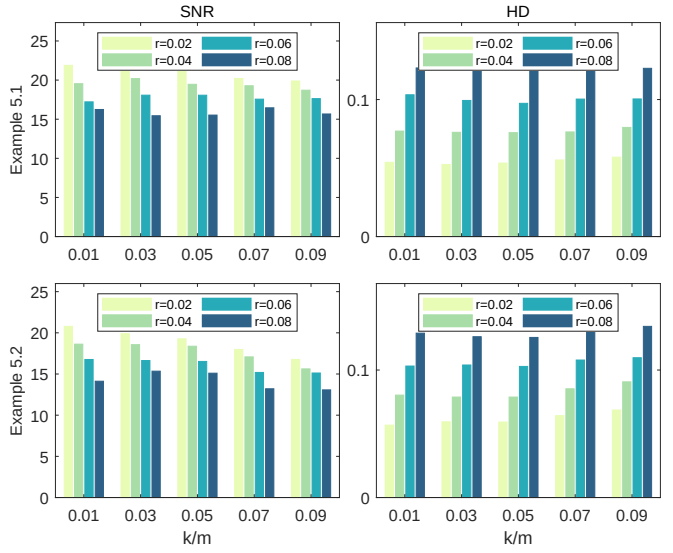


Fig. 4: Effect of k for Example 5.1 and Example 5.2.

C. Benchmark methods

Six state-of-the-art solvers are selected for comparisons. They are BIHT [20], AOPF (BIHT-AOP-flip, [27]), PAOPF (PIHT-AOP-flip, [24]), PIHT [24], PDASC [25] and WPDASC [26]. Like our method, the first four methods need to specify the sparsity level s which is set by $s = s_*$. For AOPF and PAOPF, we adopt one of the suggested options in [27] to set the required number of the sign flips as

$L = \|\mathrm{sgn}(\Phi\mathbf{x}) - \mathbf{c}\|_0$, where $\mathbf{x}$ is the solution generated by BIHT. Moreover, to accelerate the termination of PIHT, we stop it when the number of wrong recovery signs is smaller than L and the number of iterations is over 50. The other parameters for each method are chosen to be their default values. All methods are initialized by $\mathbf{x}^0 = 0$, and their final solutions are normalized to have a unit length. Finally, to make fair comparisons, for PDASC and WPDASC, if one obtains a solution with more than s nonzero entries, we keep the first s largest absolute values. Overall, all methods start with the same initial points and generate solutions with at most s nonzero entries.

D. Numerical comparisons

We now apply the seven methods into solving two examples under different scenarios. For each scenario, we report average results over 200 instances if $n \leq 1000$ and 20 instances otherwise. Note that for each example, there are five factors (n, m, s_*, r, v), where v only makes sense for Example 5.2. In the following numerical comparisons, we shall see the effect of these factors by altering one factor while fixing the others.

(a) Effect of s_*. We first employ seven method to solve Example 5.1 and increase s_* from 2 to 10 with fixing $(n, m, r, v) = (500, 250, 0.05, 0.5)$. As shown in Figure 5a, it can be clearly seen that GPSP gets the highest SNR, the smallest HD and HE for each s_*, followed by AOP, PAOP and PIHT. The lines of SNR display declining trends, which means the signal is getting harder to recover when it has more nonzero components, namely, s_* is getting bigger.

(b) Effect of m. To see the effect of the sample size m, we select it from the range $\{0.1, 0.3, \cdots, 1.5\}n$ and fix $(n, s_*, r, v) = (500, 5, 0.05, 0.5)$. As shown in Figure 5b, again, GPSP outperforms the others for solving Example 5.2 since it delivers much higher SNR and lower HD and HE. It is evidently seen that all methods are behaving better along with the rising of the sample size m, this is because the signal is getting easier to recover when more samples are available.

(c) Effect of v. We note that in Example 5.2, the larger v is, the more correlated each pair of samples (i.e., rows in Φ) are, leading to more difficult recovery. To see this, we alter v from $\{0.1, 0.2, \cdots, 0.9\}$ but fix $(n, m, s_*, r) = (500, 250, 5, 0.05)$, and report the average results in Figure 5c. As expected, the larger v is, the more difficult the recovery is. It is observed that GPSP is quite robust to v between 0.1 and 0.6 since the produced results stay steadily when $v \in [0.1, 0.6]$. No matter how v changes, GPSP always performs the best results among those methods.

(d) Effect of r. To see the effect of the flipping ratio r, we alter it from $\{0, 0.05, 0.1\}$ but fix $(n, m, s_*, v) = (500, 250, 5, 0.5)$. The box-plots of each method for solving Example 5.2 are presented in Figure 6. In each box, the central mark (red line) indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The outliers are plotted individually using the '+' symbols.

As expected, the larger r is, the worse performance of each method, because more correct signs are flipped. This can be testified by SNR (resp. HD and HE) whose median obtained by each method is declining (resp. rising) when r ascends. Once again, GPSP behaves the best because it delivers the highest median of SNR and the lowest median of HD and HE in each box. Similar results can be observed for Example 5.1 and are omitted here.

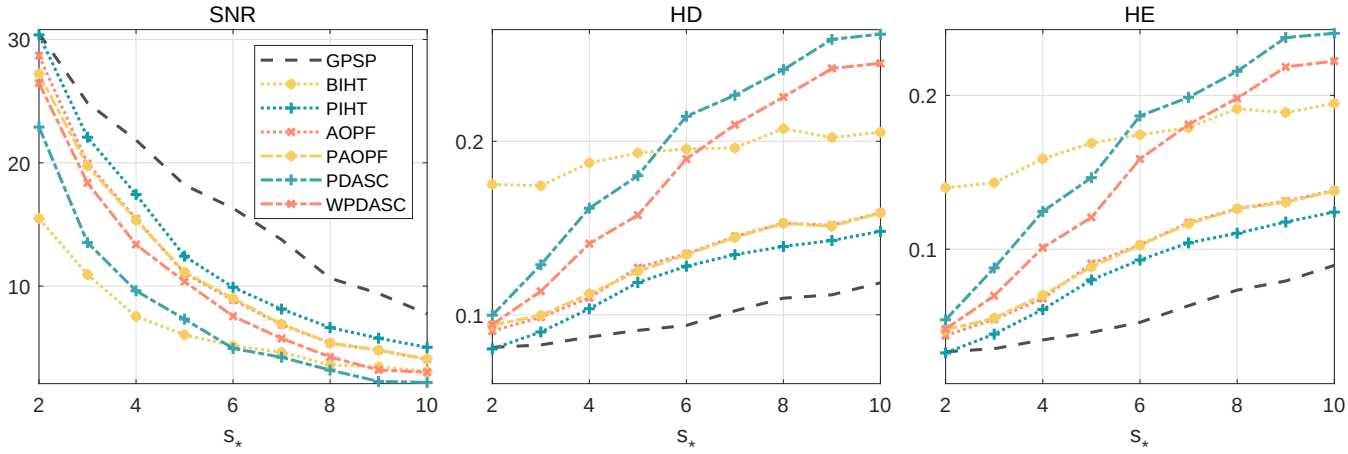
(e) Effect of n. To see the computational speed of each method, we consider some bigger values of n from $\{5000, 10000, 15000, 20000\}$ with fixing $(m, s_*, r, v) = (n/2, n/100, 0.05, 0.5)$. The average results are recorded in Table I. Obviously, GPSP achieves the highest recovery accuracy in terms of the highest SNR, the lowest HD and HE, against the other methods. For the computational speed, PAOPF runs the fastest for Example 5.1 while GPSP is the winner for Example 5.2.

TABLE I: Effect of the bigger values of n.

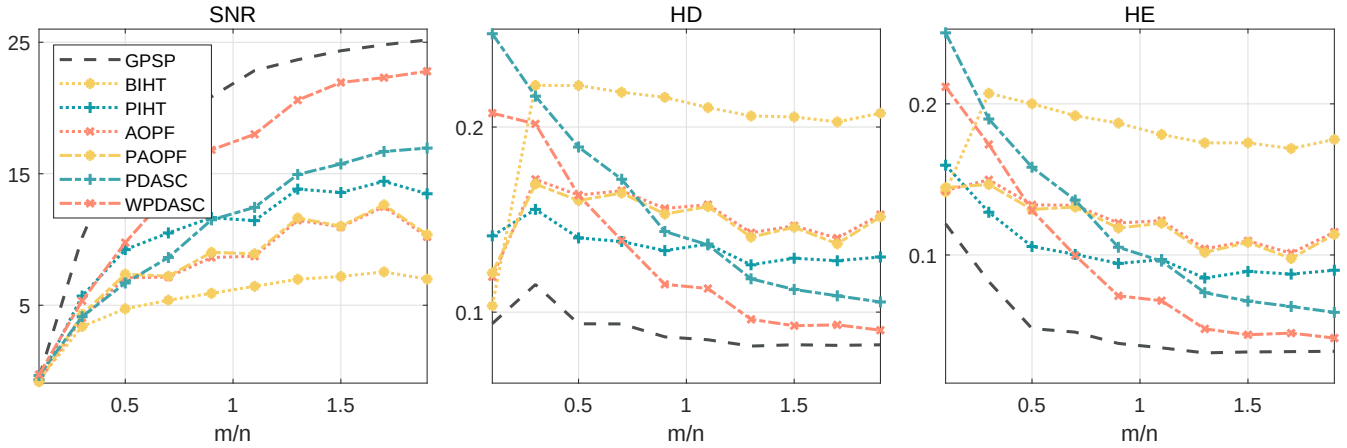
	n	GPSP	BIHT	PIHT	AOPF	PAOPF	PDASC	WPDASC
		Example 5.1						
SNR	5000	15.51	4.612	9.584	7.948	7.948	-0.228	0.786
	10000	12.17	4.431	9.230	7.464	7.464	-1.755	-1.141
	15000	12.47	4.720	8.976	7.256	7.256	-2.023	-1.477
	20000	12.89	4.693	9.095	7.348	7.348	-2.135	-2.111
HD	5000	0.092	0.201	0.122	0.138	0.138	0.353	0.332
	10000	0.106	0.207	0.125	0.141	0.141	0.425	0.394
	15000	0.103	0.202	0.127	0.142	0.142	0.438	0.412
	20000	0.102	0.201	0.126	0.143	0.143	0.444	0.443
HE	5000	0.051	0.180	0.087	0.106	0.106	0.339	0.317
	10000	0.067	0.185	0.089	0.109	0.109	0.417	0.383
	15000	0.065	0.180	0.092	0.111	0.111	0.433	0.404
	20000	0.062	0.178	0.091	0.111	0.111	0.437	0.437
TIME	5000	0.147	0.528	1.187	0.072	0.014	0.378	0.473
	10000	0.580	4.575	6.124	0.258	0.056	1.451	1.698
	15000	1.054	10.09	14.39	0.558	0.139	2.973	3.791
	20000	1.807	17.78	24.82	0.978	0.243	4.788	6.459
		Example 5.2						
SNR	5000	13.35	6.865	7.368	4.794	5.351	-0.948	-0.860
	10000	11.55	7.699	7.180	4.980	5.630	-1.857	-1.340
	15000	11.22	8.183	6.982	4.668	5.386	-1.802	-1.896
	20000	11.67	8.185	6.942	4.774	5.549	-2.224	-2.324
HD	5000	0.099	0.200	0.148	0.171	0.168	0.379	0.375
	10000	0.106	0.193	0.148	0.164	0.163	0.429	0.405
	15000	0.109	0.195	0.149	0.165	0.163	0.426	0.431
	20000	0.106	0.191	0.150	0.162	0.162	0.446	0.452
HE	5000	0.058	0.170	0.113	0.143	0.138	0.366	0.362
	10000	0.070	0.164	0.115	0.137	0.133	0.422	0.396
	15000	0.072	0.164	0.117	0.139	0.135	0.417	0.423
	20000	0.069	0.162	0.117	0.136	0.133	0.441	0.448
TIME	5000	0.155	0.542	1.274	2.201	0.086	0.386	0.474
	10000	0.566	4.523	6.448	13.07	1.600	1.468	1.736
	15000	1.139	10.55	14.71	31.91	3.575	2.947	3.726
	20000	1.996	18.39	25.44	58.05	10.60	5.103	7.150

VI. Conclusion

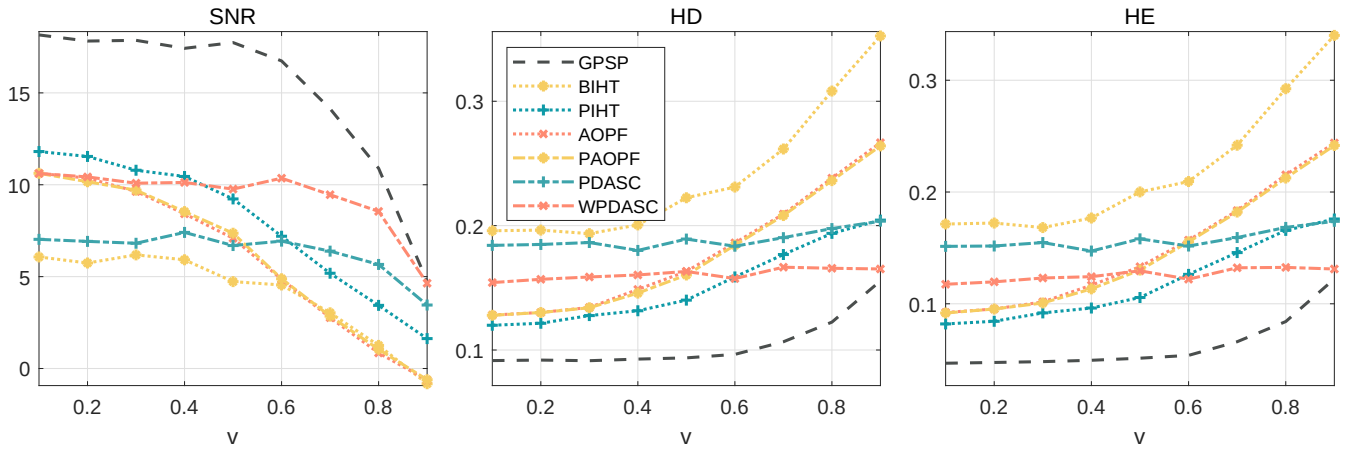
In this paper we have proposed a nonconvex optimization problem (9) to process the one-bit CS, in which the double-sparsity constrains the sparsity of the signal and the number of sign flips. To conquer the hardness resulting from the nonconvex and discrete constraints, we have established necessary and sufficient optimality conditions via the so-called τ-stationarity. These optimality conditions have facilitated the design of a gradient projection subspace pursuit method GPSP, which has been shown to admit global convergence and highly efficient numerical performance both in computation time and the recovery accuracy of signals.



(a) Effect of s_* for Example 5.1.



(b) Effect of m for Example 5.2.



(c) Effect of v for Example 5.2.

Fig. 5: Effect of (s_*, m, r) for seven methods.

APPENDIX A
PROOFS OF ALL THEOREMS

A. *Proof of Lemma 3.1*

Let $\mathbf{z}^*$ be a local minimizer of (9). Then there is a $\delta > 0$ such that, for any $\mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \delta)$,

$$\begin{aligned} 0 &\leq f(\mathbf{z}) - f(\mathbf{z}^*) \\ &\overset{(17)}{\leq} \langle \nabla f(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle + \lambda_{\max} \|\mathbf{z} - \mathbf{z}^*\|^2 =: g(\mathbf{z}). \end{aligned} \tag{36}$$

To verify that $\mathbf{z}^*$ satisfies (18), we consider the four cases.

• $\|\mathbf{y}_+^*\|_0 < k$. If there is an $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i \neq 0$, then for any $t \in \mathbb{R}$, define

$$\begin{aligned} \mathbf{x}_t &:= \mathbf{x}^* + t \cdot 0, \\ \mathbf{y}_t &:= \mathbf{y}^* + t(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i \cdot \mathbf{e}_i, \\ \mathbf{z}_t &:= (\mathbf{x}_t; \mathbf{y}_t). \end{aligned} \tag{37}$$

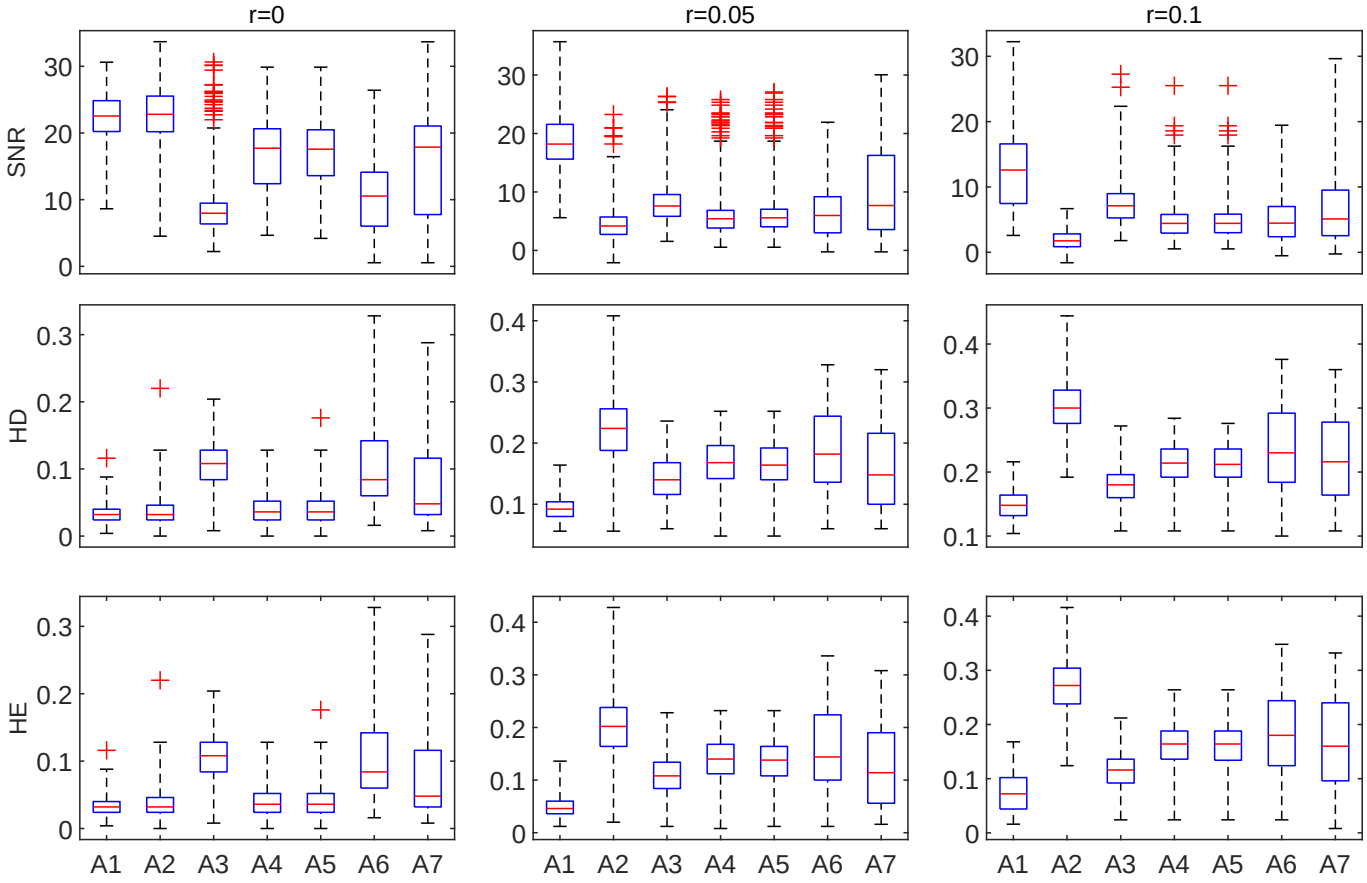


Fig. 6: Effect of r for Example 5.2, where A1-A7 stand for GPSP, BIHT, PIHT, AOPF, PAOPF, PDASC, WPDASC, respectively.

where $\mathbf{e}_i \in \mathbb{R}^m$ is the ith column of the identity matrix. It is easy to see that $\mathbf{z}_t \in \mathcal{F}$ and

$$g(\mathbf{z}_t) = (t + \lambda_{\max} t^2)[(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i]^2.$$

For any $t \in (\max\{-\delta/|(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i|, -1/\lambda_{\max}\}, 0)$, one can verify that $\mathbf{z}_t \in \mathcal{F} \cap N(\mathbf{z}^*, \delta)$ and $g(\mathbf{z}_t) < 0$, contradicting with (36). Thus $\nabla_{\mathbf{y}} f(\mathbf{z}^*) = 0$.

• $\|\mathbf{y}_+^*\|_0 = k$. If there is an $i \in \Gamma_*$ such that $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i \neq 0$, then let $\mathbf{x}_t, \mathbf{y}_t$ and $\mathbf{z}_t$ be as (37). The same reasoning for the case $\|\mathbf{y}_+^*\|_0 < s$ enables to prove $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i = 0$. This displays $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\Gamma_*} = 0$. To show $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\overline{\Gamma}_*} \leq 0$, consider any given $i \in \overline{\Gamma}_*$. For any $t \leq 0$, let

$$\mathbf{x}_t := \mathbf{x}^* + t \cdot 0, \quad \mathbf{y}_t := \mathbf{y}^* + t \cdot \mathbf{e}_i, \quad \mathbf{z}_t := (\mathbf{x}_t; \mathbf{y}_t). \tag{38}$$

It follows from $y_i^* = 0$ and $t \leq 0$ that $(\mathbf{y}_t)_+ = \mathbf{y}_+^*$ and $\|\mathbf{z}_t - \mathbf{z}^*\| = -t$. Thus, $\mathbf{z}_t \in \mathcal{F} \cap N(\mathbf{z}^*, \delta)$ for any $t \in (-\delta, 0)$. Applying (36) yields

$$g(\mathbf{z}_t) = t(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i + \lambda_{\max} t^2 \geq 0,$$

which implies $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i + \lambda_{\max} t \leq 0$. Letting $t \to 0$ derives $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_i \leq 0$. This delivers $(\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\overline{\Gamma}_*} \leq 0$.

• $\|\mathbf{x}^*\|_0 < s$. The same reasoning for the case $\|\mathbf{y}_+^*\|_0 < k$ leads to $\nabla_{\mathbf{x}} f(\mathbf{z}^*) = 0$.

• $\|\mathbf{x}^*\|_0 = s$. The same reasoning for the case $\|\mathbf{y}_+^*\|_0 < k$ yields $(\nabla_{\mathbf{x}} f(\mathbf{z}^*))_{T_*} = 0$.

Conversely, let $\mathbf{z}^*$ satisfy (18). We again consider the following four cases.

• $\|\mathbf{y}_+^*\|_0 < k$. This leads to $\langle \nabla_{\mathbf{y}} f(\mathbf{z}^*), \mathbf{y} - \mathbf{y}^* \rangle = 0$ since $\nabla_{\mathbf{y}} f(\mathbf{z}^*) = 0$ by (18).

• $\|\mathbf{y}_+^*\|_0 = k$. Consider a local region $N(\mathbf{z}^*, \delta_1)$ with $\delta_1 := \min\{y_i^* : y_i^* > 0\}$. Thus, for any $\mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \delta_1)$, we have $y_j > 0$ if $y_j^* > 0$ and

$$y_i \leq 0, \quad \forall\, i \in \overline{\Gamma}_*. \tag{39}$$

In fact, if there exists an $i \in \overline{\Gamma}_*$ satisfying $y_i > 0$, then $\|\mathbf{y}_+\|_0 \geq \|\mathbf{y}_+^*\|_0 + 1 = k + 1$, which contradicts with $\mathbf{y} \in K$. Direct calculations yield

$$\begin{aligned} &\langle \nabla_{\mathbf{y}} f(\mathbf{z}^*), \mathbf{y} - \mathbf{y}^* \rangle \\ \overset{(18)}{=}\ &\langle 0, (\mathbf{y} - \mathbf{y}^*)_{\Gamma_*} \rangle + \langle (\nabla_{\mathbf{y}} f(\mathbf{z}^*))_{\overline{\Gamma}_*}, \mathbf{y}_{\overline{\Gamma}_*} \rangle \overset{(18)(39)}{\geq} 0. \end{aligned}$$

• $\|\mathbf{x}^*\|_0 < s$. This yields $\langle \nabla_{\mathbf{x}} f(\mathbf{z}^*), \mathbf{x} - \mathbf{x}^* \rangle = 0$ due to $\nabla_{\mathbf{x}} f(\mathbf{z}^*) = 0$ by (18).

• $\|\mathbf{x}^*\|_0 = s$. Consider a local region $N(\mathbf{z}^*, \delta_2)$ with $\delta_2 := \min\{|x_i^*| : x_i^* \neq 0\}$. For any $\mathbf{z} := (\mathbf{x}; \mathbf{y}) \in \mathcal{F} \cap N(\mathbf{z}^*, \delta_2)$, $x_j \neq 0$ if $x_j^* \neq 0$, which indicates $T_* \subseteq \mathrm{supp}(\mathbf{x})$. This together with $\mathbf{x} \in S$ that $\|\mathbf{x}\|_0 \leq s = |T_*|$ suffices to

$$\mathrm{supp}(\mathbf{x}) = T_*. \tag{40}$$

Now, one can verify that

$$\begin{aligned} & \langle \nabla_{\mathbf{x}} f(\mathbf{z}^*), \mathbf{x}-\mathbf{x}^* \rangle \\ \overset{(18)(40)}{=} \; & \langle 0, (\mathbf{x}-\mathbf{x}^*)_{T_*} \rangle + \langle (\nabla_{\mathbf{x}} f(\mathbf{z}^*))_{\overline{T}_*}, 0 \rangle = 0. \end{aligned}$$

Apparently, if $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$, then $\nabla f(\mathbf{z}^*) = 0$. Combining with (17), we obtain

$$f(\mathbf{z}) - f(\mathbf{z}^*) = \|\mathbf{z}-\mathbf{z}^*\|_H^2 + \langle \nabla f(\mathbf{z}^*), \mathbf{z}-\mathbf{z}^* \rangle > 0,$$

for any $\mathbf{z} \neq \mathbf{z}^*$. Thus, $\mathbf{z}^*$ is the unique global optimal solution to the problem (9).

If $\mathbf{z}^* \notin \mathrm{int}\mathcal{F}$, then for any $\mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \min\{\delta_1, \delta_2\})$, it follows from (17) that

$$\begin{aligned} & f(\mathbf{z}) - f(\mathbf{z}^*) - \|\mathbf{z}-\mathbf{z}^*\|_H^2 \\ = \; & \langle \nabla_{\mathbf{x}} f(\mathbf{z}^*), \mathbf{x}-\mathbf{x}^* \rangle + \langle \nabla_{\mathbf{y}} f(\mathbf{z}^*), \mathbf{y}-\mathbf{y}^* \rangle \geq 0. \end{aligned}$$

Therefore, $\mathbf{z}^*$ is the unique global minimizer of $\min\{f(\mathbf{z}) : \mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \min\{\delta_1, \delta_2\})\}$, which means it is the unique local minimizer of the problem (9). □

B. Proof of Lemma 3.2

A τ-stationary point satisfies (20) which is equivalent to

$$\begin{cases} \mathbf{x}^* \in & \Pi_S\left(\mathbf{x}^* - \tau \nabla_{\mathbf{x}} f(\mathbf{z}^*)\right), \\ \mathbf{y}^* \in & \Pi_K\left(\mathbf{y}^* - \tau \nabla_{\mathbf{y}} f(\mathbf{z}^*)\right). \end{cases} \tag{41}$$

Therefore, we show the equivalence between (21) and (41). For the $\mathbf{x}^*$ part, this can be guaranteed by [42, Lemma 2.2]. For the $\mathbf{y}^*$ part, the projection (15) enables to show that (21) $\Rightarrow$ (41). So we only prove (41) $\Rightarrow$ (21). Let $\lambda^* := \nabla_{\mathbf{y}} f(\mathbf{z}^*)$. It follows from (15) that

$$\begin{aligned} \mathbf{y}^* \; & \in \; \Pi_K(\mathbf{y}^* - \tau\lambda^*) \\ & = \; \left\{ \begin{bmatrix} \mathbf{y}^*_\Gamma - \tau\lambda^*_\Gamma \\ 0 \end{bmatrix} : \; \Gamma \in \Theta(\mathbf{y}^* - \tau\lambda^*; k) \right\}. \end{aligned}$$

This derives $\|\mathbf{y}^*_+\|_0 \leq k$, and for any $\Gamma \in \Theta(\mathbf{y}^* - \tau\lambda^*; k)$,

$$\mathbf{y}^*_{\overline{\Gamma}} = 0, \quad \lambda^*_\Gamma = 0, \quad \mathbf{y}^* - \tau\lambda^* = \left[\mathbf{y}^*_\Gamma; \; -\tau\lambda^*_{\overline{\Gamma}}\right], \tag{42}$$

which together with the definition of $\Theta(\mathbf{y}^* - \tau\lambda^*; k)$ in (14) gives rise to

$$\Gamma = \Gamma_k \cup \Gamma_- = \mathrm{supp}(\mathbf{y}^*), \quad \overline{\Gamma} = (\Gamma_+ \setminus \Gamma_k) \cup \Gamma_0,$$

where Γ_+, Γ_- and Γ_0 are defined as (13) in which $\mathbf{y}$ is replaced by $\mathbf{y}^* - \tau\lambda^*$. On the index set $\overline{\Gamma}$, all elements $y_i^* - \tau\lambda_i^* = -\tau\lambda_i^* \geq 0$, namely, $\lambda_i^* \leq 0, i \in \overline{\Gamma}$.

• For $\|\mathbf{y}^*_+\|_0 < k$, suppose there is an $i \in \overline{\Gamma}$ such that $\lambda_i^* < 0$, then $\mathbf{y}^* - \tau\lambda^*$ has at least $\|\mathbf{y}^*_+\|_0 + 1 \leq k$ positive entries and thus

$$\|\mathbf{y}^*_+\|_0 = \|(\Pi_K(\mathbf{y}^* - \tau\lambda^*))_+\|_0 \geq \|\mathbf{y}^*_+\|_0 + 1.$$

This is a contradiction. So, $\lambda^*_{\overline{\Gamma}} = 0$, leading to $\lambda^* = 0$ by (42), which satisfies (21).

• For $\|\mathbf{y}^*_+\|_0 = k$, (21) is satisfied for any $j \in \mathrm{supp}(\mathbf{y}^*) = \Gamma$ due to $\lambda^*_\Gamma = 0$. For $j \notin \mathrm{supp}(\mathbf{y}^*)$, namely, $j \in \overline{\Gamma}$, the definition of Γ_k in (14) yields

$$0 \leq y_j^* - \tau\lambda_j^* \leq y_i^* - \tau\lambda_i^*, \quad \forall \; i \in \Gamma_k,$$

which together with $\Gamma_k \subseteq \Gamma$ and (42) results in

$$0 \leq -\tau\lambda_j^* \leq y_i^*, \quad \forall \; i \in \Gamma_k.$$

Hence, $-\|\mathbf{y}^*_+\|_{[k]} = -\min_{i \in \Gamma_k} y_i^* \leq \tau\lambda_j^* \leq 0, \forall \; j \in \overline{\Gamma} (j \notin \mathrm{supp}(\mathbf{y}^*))$, showing (21). □

C. Proof of Theorem 3.1

a) If $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$ is a global minimizer, it follows readily from Lemma 3.1 that $\nabla f(\mathbf{z}^*) = 0$. Thus, by definition, $\mathbf{z}^*$ is a τ-stationary point for any $\tau > 0$.

Conversely, if $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$ is a τ-stationary point for some $\tau > 0$, then $\|\mathbf{x}^*\|_{[s]} = \|\mathbf{y}^*_+\|_{[k]} = 0$, which further implies $\nabla_{\mathbf{x}} f(\mathbf{z}^*) = 0$ and $\nabla_{\mathbf{y}} f(\mathbf{z}^*) = 0$ by (21). Applying Lemma 3.1 again, one can conclude that $\mathbf{z}^*$ is a global minimizer.

b) Let $\mathbf{z}^*$ be a global minimizer. If it is not a τ-stationary point with $0 < \tau \leq 1/(2\lambda_{\max})$, then we have the condition $\overline{\mathbf{z}} (\neq \mathbf{z}^*) \in \Pi_{\mathcal{F}}\left(\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*)\right)$. Thus,

$$\|\overline{\mathbf{z}} - (\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*))\|^2 < \|\mathbf{z}^* - (\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*))\|^2,$$

which suffices to

$$2\tau\langle \nabla f(\mathbf{z}^*), \overline{\mathbf{z}} - \mathbf{z}^* \rangle < -\|\overline{\mathbf{z}} - \mathbf{z}^*\|^2.$$

Together with (17) and $0 < \tau \leq 1/(2\lambda_{\max})$ derives

$$\begin{aligned} f(\overline{\mathbf{z}}) - f(\mathbf{z}^*) \; & \leq \; \langle \nabla f(\mathbf{z}^*), \overline{\mathbf{z}} - \mathbf{z}^* \rangle + \lambda_{\max}\|\overline{\mathbf{z}} - \mathbf{z}^*\|^2 \\ & < \; (\lambda_{\max} - 1/(2\tau))\|\overline{\mathbf{z}} - \mathbf{z}^*\|^2 \leq 0. \end{aligned}$$

It contradicts the global optimality of $\mathbf{z}^*$. Therefore, $\mathbf{z}^*$ is a τ-stationary point with $0 < \tau \leq 1/(2\lambda_{\max})$.

Conversely, let $\mathbf{z}^*$ be a τ-stationary point with $\tau \geq 1/(2\lambda_{\min})$. The definition of $\Pi_{\mathcal{F}}$ and (20) imply

$$\|\mathbf{z}^* - (\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*))\|^2 \leq \|\mathbf{z} - (\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*))\|^2,$$

for any $\mathbf{z} \in \mathcal{F}$, delivering $2\tau\langle \nabla f(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq -\|\mathbf{z} - \mathbf{z}^*\|^2$. This and (17) yield

$$\begin{aligned} f(\mathbf{z}) - f(\mathbf{z}^*) \; & \geq \; \langle \nabla f(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle + \lambda_{\min}\|\mathbf{z} - \mathbf{z}^*\|^2 \\ & \geq \; (\lambda_{\min} - 1/(2\tau))\|\mathbf{z} - \mathbf{z}^*\|^2. \end{aligned}$$

Since $\tau \geq 1/(2\lambda_{\min})$, the above relation shows the global optimality of $\mathbf{z}^*$ to (9). □

D. Proof of Lemma 4.1

It follows from $\mathbf{z}^\ell(\tau) \in \Pi_{\mathcal{F}}(\mathbf{z}^\ell - \tau\nabla f(\mathbf{z}^\ell))$ that

$$\|\mathbf{z}^\ell(\tau) - (\mathbf{z}^\ell - \tau\nabla f(\mathbf{z}^\ell))\|^2 \leq \|\mathbf{z}^\ell - (\mathbf{z}^\ell - \tau\nabla f(\mathbf{z}^\ell))\|^2,$$

which results in

$$2\tau\langle \nabla f(\mathbf{z}^\ell), \mathbf{z}^\ell(\tau) - \mathbf{z}^\ell \rangle \leq -\|\mathbf{z}^\ell(\tau) - \mathbf{z}^\ell\|^2. \tag{43}$$

Combining with (17) leads to

$$\begin{aligned} & f(\mathbf{z}^\ell(\tau)) \\ \leq \; & f(\mathbf{z}^\ell) + \langle \nabla f(\mathbf{z}^\ell), \mathbf{z}^\ell(\tau) - \mathbf{z}^\ell \rangle + \lambda_{\max}\|\mathbf{z}^\ell(\tau) - \mathbf{z}^\ell\|^2 \\ \leq \; & f(\mathbf{z}^\ell) - \left(1/(2\tau) - \lambda_{\max}\right)\|\mathbf{z}^\ell(\tau) - \mathbf{z}^\ell\|^2 \\ \leq \; & f(\mathbf{z}^\ell) - \rho\|\mathbf{z}^\ell(\tau) - \mathbf{z}^\ell\|^2, \end{aligned}$$

where the last inequality is from $0 < \tau \leq 1/(2\rho + 2\lambda_{\max})$. Invoking the Armijo-type step size rule, one has $\tau_\ell \geq \beta/(2\rho + 2\lambda_{\max})$, which by $\tau_\ell \leq 1$ proves the desired assertion. □

E. Proof of Lemma 4.2

a) By Lemma 4.1 and $\mathbf{u}^\ell = \mathbf{z}^\ell(\tau_\ell)$, we have

$$f(\mathbf{u}^\ell) \leq f(\mathbf{z}^\ell) - \rho\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2. \quad (44)$$

By the framework of Algorithm 1, if $\mathbf{z}^{\ell+1} = \mathbf{u}^\ell$, then the above condition implies,

$$\begin{aligned} f(\mathbf{z}^{\ell+1}) &\leq f(\mathbf{z}^\ell) - \rho\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2 \\ &= f(\mathbf{z}^\ell) - \rho\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\|^2. \end{aligned} \quad (45)$$

If $\mathbf{z}^{\ell+1} = \mathbf{v}^\ell$, then we obtain

$$\begin{aligned} f(\mathbf{z}^{\ell+1}) &= f(\mathbf{v}^\ell) \\ &\leq f(\mathbf{u}^\ell) - \rho\|\mathbf{z}^{\ell+1} - \mathbf{u}^\ell\|^2 \\ &\leq f(\mathbf{z}^\ell) - \rho\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2 - \rho\|\mathbf{z}^{\ell+1} - \mathbf{u}^\ell\|^2 \quad (46) \\ &\leq f(\mathbf{z}^\ell) - (\rho/2)\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\|^2, \end{aligned}$$

where the second and last inequalities used (44) and a fact $\|\mathbf{a}+\mathbf{b}\|^2 \leq 2\|\mathbf{a}\|^2 + 2\|\mathbf{b}\|^2$ for all vectors $\mathbf{a}$ and $\mathbf{b}$. Both cases lead to

$$\begin{aligned} f(\mathbf{z}^{\ell+1}) &\leq f(\mathbf{z}^\ell) - (\rho/2)\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\|^2, \quad (47) \\ f(\mathbf{z}^{\ell+1}) &\leq f(\mathbf{z}^\ell) - \rho\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2. \end{aligned}$$

Therefore, $\{f(\mathbf{z}^\ell)\}$ is a non-increasing sequence, and thus

$$\max\{\|A\mathbf{x}^\ell - \epsilon\mathbf{1} + \mathbf{y}^\ell\|^2, \eta\|\mathbf{x}^\ell\|^2\} \leq f(\mathbf{z}^\ell) \leq f(\mathbf{z}^0),$$

which indicates the boundedness of $\{\mathbf{x}^\ell\}$ and $\{\mathbf{y}^\ell\}$, and so that of $\{\mathbf{z}^\ell\}$. The non-increasing property in (47) and $f \geq 0$ also give rise to

$$\begin{aligned} &\sum_{\ell\geq 0} \max\{(\rho/2)\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\|^2, \rho\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2\} \\ \leq &\sum_{\ell\geq 0} \left[f(\mathbf{z}^\ell) - f(\mathbf{z}^{\ell+1})\right] \\ = &\ f(\mathbf{z}^0) - \lim_{\ell\to\infty} f(\mathbf{z}^{\ell+1}) \leq f(\mathbf{z}^0). \end{aligned}$$

The above condition suffices to $\lim_{\ell\to\infty}\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\| = \lim_{\ell\to\infty}\|\mathbf{u}^\ell - \mathbf{z}^\ell\| = 0$.

b) Let $\mathbf{z}^*$ be any accumulating point of $\{\mathbf{z}^\ell\}$. Then there exists a subset J of $\{0, 1, 2, \cdots\}$ such that $\lim_{\ell(\in J)\to\infty}\mathbf{z}^\ell = \mathbf{z}^*$. This further implies $\lim_{\ell(\in J)\to\infty}\mathbf{u}^\ell = \mathbf{z}^*$ by applying a). In addition, as stated in Lemma 4.1, we have $\{\tau_\ell\} \subseteq [\underline{\tau}, 1]$, which indicates that one can find a subsequence L of J and a scalar $\tau_* \in [\underline{\tau}, 1]$ such that $\{\tau_\ell : \ell \in L\} \to \tau_*$. To summarize, we have

$$\mathbf{z}^\ell \to \mathbf{z}^*, \quad \mathbf{u}^\ell \to \mathbf{z}^*, \quad \tau_\ell \to \tau_* \in [\underline{\tau}, 1], \quad \ell(\in L) \to \infty. \quad (48)$$

Let $\overline{\mathbf{z}}^\ell := \mathbf{z}^\ell - \tau_\ell\nabla f(\mathbf{z}^\ell)$. Algorithm 1 implies

$$\mathbf{u}^\ell \in \Pi_{\mathcal{F}}(\overline{\mathbf{z}}^\ell), \quad \lim_{\ell(\in L)\to\infty}\overline{\mathbf{z}}^\ell = \mathbf{z}^* - \tau_*\nabla f(\mathbf{z}^*) =: \overline{\mathbf{z}}^*. \quad (49)$$

The first condition means $\mathbf{u}^\ell \in \mathcal{F}$ for any $\ell \geq 1$. Note that $\mathcal{F}$ is closed and $\mathbf{z}^*$ is the accumulating point of $\{\mathbf{u}^\ell\}$ by (48). Therefore, $\mathbf{z}^* \in \mathcal{F}$, which results in

$$\min_{\mathbf{z}\in\mathcal{F}}\|\mathbf{z} - \overline{\mathbf{z}}^*\| \leq \|\mathbf{z}^* - \overline{\mathbf{z}}^*\|. \quad (50)$$

If the strict inequality holds in the above condition, then there is an $\varepsilon_0 > 0$ such that

$$\begin{aligned} \|\mathbf{z}^* - \overline{\mathbf{z}}^*\| - \varepsilon_0 &= \min_{\mathbf{z}\in\mathcal{F}}\|\mathbf{z} - \overline{\mathbf{z}}^*\| \\ &\geq \min_{\mathbf{z}\in\mathcal{F}}(\|\mathbf{z} - \overline{\mathbf{z}}^\ell\| - \|\overline{\mathbf{z}}^\ell - \overline{\mathbf{z}}^*\|) \\ &= \|\mathbf{u}^\ell - \overline{\mathbf{z}}^\ell\| - \|\overline{\mathbf{z}}^\ell - \overline{\mathbf{z}}^*\| \end{aligned}$$

where the last equality is from (49). Taking the limit of both sides of the above condition along $\ell(\in L) \to \infty$ yields $\|\mathbf{z}^* - \overline{\mathbf{z}}^*\| - \varepsilon_0 \geq \|\mathbf{z}^* - \overline{\mathbf{z}}^*\|$ by (48) and (49), a contradiction with $\varepsilon_0 > 0$. Therefore, we must have the equality holds in (50), showing that

$$\mathbf{z}^* \in \Pi_{\mathcal{F}}(\overline{\mathbf{z}}^*) = \Pi_{\mathcal{F}}\left(\mathbf{z}^* - \tau_*\nabla f(\mathbf{z}^*)\right).$$

The above relation means the conditions in (21) hold for $\tau = \tau_*$, then these conditions must hold for any $0 < \tau \leq \underline{\tau}$ due to $\underline{\tau} \leq \tau_*$ from (48), namely,

$$\mathbf{z}^* \in \Pi_{\mathcal{F}}\left(\mathbf{z}^* - \tau\nabla f(\mathbf{z}^*)\right),$$

displaying that $\mathbf{z}^*$ is a τ-stationary point of (9), as desired. □

F. Proof of Theorem 4.1

As shown in Lemma 4.2, one can find a subsequence of $\{\mathbf{z}^\ell\}$ that converges to the τ-stationary point $\mathbf{z}^*$ with $0 < \tau \leq \underline{\tau}$ of (9). Recall that a τ-stationary point $\mathbf{z}^*$ that satisfies (21) also meets (18), which by Lemma 3.1 indicates that $\mathbf{z}^*$ is the unique global minimizer if $\mathbf{z}^* \in \mathrm{int}\mathcal{F}$ and the unique local minimizer otherwise. In other words, $\mathbf{z}^*$ is an isolated local minimizer of (9). Finally, it follows from $\mathbf{z}^*$ being isolated, [43, Lemma 4.10] and $\lim_{\ell\to\infty}\|\mathbf{z}^{\ell+1} - \mathbf{z}^\ell\| = 0$ by Lemma 4.2 that the whole sequence converges to $\mathbf{z}^*$. □

G. Proof of Theorem 4.2

i) Theorem 4.1 states that the whole sequence $\{\mathbf{z}^\ell\}$ converges to $\mathbf{z}^*$. So does $\{\mathbf{u}^\ell\}$ by Lemma 4.2. Then it is easy to show that, for sufficiently large ℓ,

$$\mathrm{supp}(\mathbf{z}^*) \subseteq \mathrm{supp}(\mathbf{u}^\ell).$$

In addition, it follows from $\mathbf{u}^\ell := \mathbf{z}^\ell(\tau_\ell) = (\mathbf{x}^\ell(\tau_\ell); \mathbf{y}^\ell(\tau_\ell)) \in \Pi_{\mathcal{F}}(\mathbf{z}^\ell - \tau_\ell\nabla f(\mathbf{z}^\ell))$ that

$$\mathbf{x}^\ell(\tau_\ell) \in \Pi_S(\mathbf{x}^\ell - \tau_\ell\nabla_{\mathbf{x}} f(\mathbf{z}^\ell)), \ \ \mathbf{y}^\ell(\tau_\ell) \in \Pi_K(\mathbf{y}^\ell - \tau_\ell\nabla_{\mathbf{y}} f(\mathbf{z}^\ell)),$$

which by (12) and (15) results in

$$\begin{aligned} \mathbf{x}^\ell(\tau_\ell) &= ((\mathbf{x}^\ell - \tau_\ell\nabla_{\mathbf{x}} f(\mathbf{z}^\ell))_T; 0), \\ \mathbf{y}^\ell(\tau_\ell) &= ((\mathbf{y}^\ell - \tau_\ell\nabla_{\mathbf{y}} f(\mathbf{z}^\ell))_\Gamma; 0), \end{aligned} \quad (51)$$

where $T \in \Sigma(\mathbf{x}^\ell - \tau_\ell\nabla_{\mathbf{x}} f(\mathbf{z}^\ell); s)$ and $\Gamma \in \Theta(\mathbf{y}^\ell - \tau_\ell\nabla_{\mathbf{y}} f(\mathbf{z}^\ell); k)$. Therefore,

$$\begin{aligned} \mathrm{supp}(\mathbf{z}^*) \subseteq \mathrm{supp}(\mathbf{u}^\ell) &= \mathrm{supp}((\mathbf{x}^\ell(\tau_\ell); \mathbf{y}^\ell(\tau_\ell))) \\ &\subseteq T \cup (n + \Gamma) := W, \end{aligned} \quad (52)$$

for sufficiently large ℓ, where $n + \Gamma := \{n + i : i \in \Gamma\}$, which leads to

$$\mathbf{z}^*_{\overline{W}} = \mathbf{u}^\ell_{\overline{W}} = 0, \quad \mathbf{u}^\ell_W = (\mathbf{z}^\ell - \tau_\ell\nabla f(\mathbf{z}^\ell))_W, \quad (53)$$

where the last equality is from (51). Those conditions contribute to

$$\begin{aligned} & \langle \mathbf{z}^* - \mathbf{u}^\ell, \mathbf{u}^\ell - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))\rangle \\ = \; & \langle \mathbf{z}^*_W - \mathbf{u}^\ell_W, \mathbf{u}^\ell_W - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))_W\rangle \\ + \; & \langle \mathbf{z}^*_{\overline{W}} - \mathbf{u}^\ell_{\overline{W}}, \; \mathbf{u}^\ell_{\overline{W}} - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))_{\overline{W}}\rangle \\ = \; & 0. \end{aligned} \tag{54}$$

Using the above fact, we obtain

$$\begin{aligned} & \|\mathbf{z}^* - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))\|^2 \\ = \; & \|\mathbf{z}^* - \mathbf{u}^\ell + \mathbf{u}^\ell - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))\|^2 \\ = \; & \|\mathbf{z}^* - \mathbf{u}^\ell\|^2 + \|\mathbf{u}^\ell - (\mathbf{z}^\ell - \tau_\ell \nabla f(\mathbf{z}^\ell))\|^2, \end{aligned}$$

which after the simple manipulating results in

$$\begin{aligned} & \langle \nabla f(\mathbf{z}^\ell), \mathbf{u}^\ell - \mathbf{z}^\ell\rangle + \frac{1}{2\tau_\ell}\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2 \qquad (55) \\ = \; & \langle \nabla f(\mathbf{z}^\ell), \mathbf{z}^* - \mathbf{z}^\ell\rangle + \frac{1}{2\tau_\ell}\left[\|\mathbf{z}^* - \mathbf{z}^\ell\|^2 - \|\mathbf{z}^* - \mathbf{u}^\ell\|^2\right]. \end{aligned}$$

By Lemma 4.1, the Armijo-type step size rule indicates $\underline{\tau} \leq \tau_\ell \leq 1/(2\rho + 2\lambda_{\max}) < 1/(2\lambda_{\max})$. Therefore, it follows from (17) that

$$\begin{aligned} f(\mathbf{u}^\ell) \; \leq \; & f(\mathbf{z}^\ell) + \langle \nabla f(\mathbf{z}^\ell), \mathbf{u}^\ell - \mathbf{z}^\ell\rangle + \lambda_{\max}\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2 \\ \leq \; & f(\mathbf{z}^\ell) + \langle \nabla f(\mathbf{z}^\ell), \mathbf{u}^\ell - \mathbf{z}^\ell\rangle + \frac{1}{2\tau_\ell}\|\mathbf{u}^\ell - \mathbf{z}^\ell\|^2 \\ \overset{(55)}{=} \; & f(\mathbf{z}^\ell) + \langle \nabla f(\mathbf{z}^\ell), \mathbf{z}^* - \mathbf{z}^\ell\rangle \\ + \; & \frac{1}{2\tau_\ell}\left[\|\mathbf{z}^* - \mathbf{z}^\ell\|^2 - \|\mathbf{z}^* - \mathbf{u}^\ell\|^2\right] \\ \overset{(17)}{\leq} \; & f(\mathbf{z}^*) - \lambda_{\min}\|\mathbf{z}^* - \mathbf{z}^\ell\|^2 \\ + \; & \frac{1}{2\tau_\ell}\left[\|\mathbf{z}^* - \mathbf{z}^\ell\|^2 - \|\mathbf{z}^* - \mathbf{u}^\ell\|^2\right]. \end{aligned} \tag{56}$$

Recall that a τ-stationary point $\mathbf{z}^*$ that satisfies (21) also meets (18). By Lemma 3.1, there exits a $\delta_* > 0$ satisfying (19) for any $\mathbf{z} \in \mathcal{F} \cap N(\mathbf{z}^*, \delta_*)$. Thus, the fact $\mathbf{u}^\ell \to \mathbf{z}^*$ indicates that for sufficiently large ℓ, we have $\mathbf{u}^l \in \mathcal{F} \cap N(\mathbf{z}^*, \delta_*)$ and hence

$$f(\mathbf{u}^\ell) \geq f(\mathbf{z}^*) + \|\mathbf{z}^* - \mathbf{u}^\ell\|_H^2 \geq f(\mathbf{z}^*) + \lambda_{\min}\|\mathbf{z}^* - \mathbf{u}^\ell\|^2.$$

Together with (56) gives rise to

$$\begin{aligned} & 2\tau_\ell \lambda_{\min}\|\mathbf{z}^* - \mathbf{u}^\ell\|^2 \\ \leq \; & -2\tau_\ell \lambda_{\min}\|\mathbf{z}^* - \mathbf{z}^\ell\|^2 + \|\mathbf{z}^* - \mathbf{z}^\ell\|^2 - \|\mathbf{z}^* - \mathbf{u}^\ell\|^2. \end{aligned}$$

Note that $\tau_\ell \geq \underline{\tau}$. The desired assertion then follows immediately from

$$(1 + 2\underline{\tau}\lambda_{\min})\|\mathbf{z}^* - \mathbf{u}^\ell\|^2 \leq (1 - 2\underline{\tau}\lambda_{\min})\|\mathbf{z}^* - \mathbf{z}^\ell\|^2.$$

ii) If there are infinitely many ℓ such that $\mathbf{z}^{\ell+1} = \mathbf{u}^\ell$, then (34) for sufficiently large ℓ can be derived by (33) immediately. Otherwise, there is an $\tilde{\ell} \geq 1$ such that $\mathbf{z}^{\ell+1} = \mathbf{v}^\ell$ for any $\ell \geq \tilde{\ell}$. The updating (27) of $\mathbf{v}^\ell$ indicates

$$T^{\ell+1} \subseteq T^\ell, \quad \Gamma_+^{\ell+1} \subseteq \Gamma_+^\ell, \quad \Gamma_0^{\ell+1} \supseteq \Gamma_0^\ell, \quad \Gamma_-^{\ell+1} \subseteq \Gamma_-^\ell.$$

Note that these sets have finite elements. Therefore, the sequences $\{T^\ell\}, \{\Gamma_+^\ell\}, \{\Gamma_0^\ell\}, \{\Gamma_-^\ell\}$ converge. In other words, there is an $\hat{\ell} \geq \tilde{\ell}$ such that, for any $\ell \geq \hat{\ell}$, it holds

$$T^{\ell+1} = T^\ell, \quad \Gamma_+^{\ell+1} = \Gamma_+^\ell, \quad \Gamma_0^{\ell+1} = \Gamma_0^\ell, \quad \Gamma_-^{\ell+1} = \Gamma_-^\ell, \tag{57}$$

which implies $\Omega(\mathbf{z}^{\ell+1}) = \Omega(\mathbf{z}^\ell)$. This and(27) yield

$$\begin{aligned} \mathbf{z}^{\ell+2} = \mathbf{v}^{\ell+1} \; = \; & \operatorname{argmin}\{f(\mathbf{z}) : \mathbf{z} \in \Omega(\mathbf{z}^{\ell+1})\} \\ = \; & \operatorname{argmin}\{f(\mathbf{z}) : \mathbf{z} \in \Omega(\mathbf{z}^\ell)\} \\ = \; & \mathbf{v}^\ell = \mathbf{z}^{\ell+1}. \end{aligned}$$

Overall, for any $\ell > \hat{\ell}$, we have $\mathbf{z}^\ell = \mathbf{v}^{\hat{\ell}}$. Recall Theorem 4.1 that the whole sequence $\{\mathbf{z}^\ell\}$ converges to $\mathbf{z}^*$, which suffices to $\mathbf{z}^* = \lim_{\ell\to\infty} \mathbf{z}^\ell = \mathbf{v}^{\hat{\ell}}$.

iii) Note that both $\mathbf{z}^\ell (\in \mathcal{F}) \to \mathbf{z}^*$ and $\mathbf{u}^\ell (\in \mathcal{F}) \to \mathbf{z}^*$. We must have $T^\ell = \widetilde{T}^\ell$ and $\Gamma_+^\ell = \widetilde{\Gamma}_+^\ell$ for sufficiently large ℓ if $\|\mathbf{x}^*\|_0 = s$ and $\|\mathbf{y}_+^*\|_0 = k$. Therefore, the framework of of Algorithm 1 allows us to assert that $\mathbf{z}^{\ell+1} = \mathbf{v}^\ell$ for all sufficiently large ℓ due to

$$\begin{aligned} f(\mathbf{v}^\ell) \; \overset{(29)}{\leq} \; & f(\mathbf{u}^\ell) - \lambda_{\min}\|\mathbf{u}^\ell - \mathbf{v}^\ell\|^2 \\ \leq \; & f(\mathbf{u}^\ell) - \rho\|\mathbf{u}^\ell - \mathbf{v}^\ell\|^2, \end{aligned}$$

where the second inequality is from $0 < \rho \leq \lambda_{\min}$. Then similar reasoning to prove ii) can claim the conclusion. □

REFERENCES

[1] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

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

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

[4] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in *2008 42nd Annual Conference on Information Sciences and Systems*. IEEE, 2008, pp. 16–21.

[5] A. Movahed and M. C. Reed, "Iterative detection for compressive sensing: Turbo CS," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 4518–4523.

[6] W. Tang, W. Xu, X. Zhang, and J. Lin, "A low-cost channel feedback scheme in mmwave massive mimo system," in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, 2017, pp. 89–93.

[7] Z. Zhou, X. Chen, D. Guo, and M. L. Honig, "Sparse channel estimation for massive MIMO with 1-bit feedback per dimension," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2017, pp. 1–6.

[8] J. Meng, H. Li, and Z. Han, "Sparse event detection in wireless sensor networks using compressive sensing," in *2009 43rd Annual Conference on Information Sciences and Systems*. IEEE, 2009, pp. 181–185.

[9] C. Feng, S. Valaee, and Z. Tan, "Multiple target localization using compressive sensing," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*. IEEE, 2009, pp. 1–6.

[10] J. Xiong and Q. Tang, "1-bit compressive data gathering for wireless sensor networks," *Journal of Sensors*, vol. 2014, 2014.

[11] C.-H. Chen and J.-Y. Wu, "Amplitude-aided 1-bit compressive sensing over noisy wireless sensor networks," *IEEE Wireless Communications Letters*, vol. 4, no. 5, pp. 473–476, 2015.

[12] D. Lee, T. Sasaki, T. Yamada, K. Akabane, Y. Yamaguchi, and K. Uehara, "Spectrum sensing for networked system using 1-bit compressed sensing with partial random circulant measurement matrices," in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*. IEEE, 2012, pp. 1–5.

[13] N. Fu, L. Yang, and J. Zhang, "Sub-nyquist 1 bit sampling system for sparse multiband signals," in *2014 22nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2014, pp. 736–740.

[14] J. Haboba, M. Mangia, R. Rovatti, and G. Setti, "An architecture for 1-bit localized compressive sensing with applications to EEG," in *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2011, pp. 137–140.

[15] X. Dong and Y. Zhang, "A map approach for 1-bit compressive sensing in synthetic aperture radar imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 6, pp. 1237–1241, 2015.

[16] Z. Li, W. Xu, X. Zhang, and J. Lin, "A survey on one-bit compressed sensing: Theory and applications," *Frontiers of Computer Science*, vol. 12, no. 2, pp. 217–230, 2018.

[17] J. N. Laska, Z. Wen, W. Yin, and R. G. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5289–5301, 2011.

[18] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[19] P. T. Boufounos, "Greedy sparse signal reconstruction from sign measurements," in *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*. IEEE, 2009, pp. 1305–1309.

[20] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.

[21] X. Fu, F.-M. Han, and H. Zou, "Robust 1-bit compressive sensing against sign flips," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 3121–3125.

[22] X. Cai, Z. Zhang, H. Zhang, and C. Li, "Soft consistency reconstruction: a robust 1-bit compressive sensing algorithm," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 4530–4535.

[23] H. Wang, X. Huang, Y. Liu, H. Sabine Van, and W. Qun, "Binary reweighted 11-norm minimization for one-bit compressed sensing," in *In Procedding of 8th International Joint Conference on Biomedical Engineering Systems and Technologies*, 2015.

[24] X. Huang, L. Shi, M. Yan, and J. A. Suykens, "Pinball loss minimization for one-bit compressive sensing: Convex models and algorithms," *Neurocomputing*, vol. 314, pp. 275–283, 2018.

[25] J. Huang, Y. Jiao, X. Lu, and L. Zhu, "Robust decoding from 1-bit compressive sampling with ordinary and regularized least squares," *SIAM Journal on Scientific Computing*, vol. 40, no. 4, pp. A2062–A2086, 2018.

[26] Q. Fan, C. Jia, J. Liu, and Y. Luo, "Robust recovery in 1-bit compressive sensing via l_q-constrained least squares," *Signal Processing*, vol. 179, p. 107822, 2021.

[27] M. Yan, Y. Yang, and S. Osher, "Robust 1-bit compressive sensing using adaptive outlier pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3868–3875, 2012.

[28] A. Movahed, A. Panahi, and G. Durisi, "A robust rfpi-based 1-bit compressive sensing reconstruction algorithm," in *2012 IEEE Information Theory Workshop*. IEEE, 2012, pp. 567–571.

[29] A. Movahed, A. Panahi, and M. C. Reed, "Recovering signals with variable sparsity levels from the noisy 1-bit compressive measurements," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 6454–6458.

[30] D.-Q. Dai, L. Shen, Y. Xu, and N. Zhang, "Noisy 1-bit compressive sensing: models and algorithms," *Applied and Computational Harmonic Analysis*, vol. 40, no. 1, pp. 1–32, 2016.

[31] Y. Plan and R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 482–494, 2012.

[32] F. Li, J. Fang, H. Li, and L. Huang, "Robust one-bit bayesian compressed sensing with sign-flip errors," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 857–861, 2014.

[33] L. Zhang, J. Yi, and R. Jin, "Efficient algorithms for robust one-bit compressive sensing," in *International Conference on Machine Learning*. PMLR, 2014, pp. 820–828.

[34] H.-J. M. Shi, M. Case, X. Gu, S. Tu, and D. Needell, "Methods for quantized compressed sensing," in *2016 Information Theory and Applications Workshop (ITA)*. IEEE, 2016, pp. 1–9.

[35] L. Flodin, V. Gandikota, and A. Mazumdar, "Superset technique for approximate recovery in one-bit compressed sensing," *arXiv preprint arXiv:1910.13971*, 2019.

[36] L. Rencker, F. Bach, W. Wang, and M. D. Plumbley, "Sparse recovery and dictionary learning from nonlinear compressive measurements," *IEEE Transactions on Signal Processing*, vol. 67, no. 21, pp. 5659–5670, 2019.

[37] S. Gopi, P. Netrapalli, P. Jain, and A. Nori, "One-bit compressed sensing: Provable support and vector recovery," in *International Conference on Machine Learning*. PMLR, 2013, pp. 154–162.

[38] X. Huang and M. Yan, "Nonconvex penalties with analytical solutions for one-bit compressive sensing," *Signal Processing*, vol. 144, pp. 341–351, 2018.

[39] P. Xiao, B. Liao, and J. Li, "One-bit compressive sensing via schur-concave function minimization," *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4139–4151, 2019.

[40] S. Khobahi and M. Soltanalian, "Model-based deep learning for one-bit compressive sensing," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5292–5307, 2020.

[41] W. Xu, Y. Tian, S. Wang, and Y. Cui, "Feature selection and classification of noisy proteomics mass spectrometry data based on one-bit perturbed compressed sensing," *Bioinformatics*, vol. 36, no. 16, pp. 4423–4431, 2020.

[42] A. Beck and Y. C. Eldar, "Sparsity constrained nonlinear optimization: Optimality conditions and algorithms," *SIAM Journal on Optimization*, vol. 23, no. 3, pp. 1480–1509, 2013.

[43] J. J. Moré and D. C. Sorensen, "Computing a trust region step," *SIAM Journal on Scientific and Statistical Computing*, vol. 4, no. 3, pp. 553–572, 1983.