Optimization Letters
https://doi.org/10.1007/s11590-024-02106-z

ORIGINAL PAPER

On the implementation of ADMM with dynamically configurable parameter for the separable ℓ_1/ℓ_2 minimization

Jun Wang[1,2] · **Qiang Ma**[2]

Received: 26 April 2023 / Accepted: 28 February 2024

Abstract
In this paper, we propose a novel variant of the alternating direction method of multipliers (ADMM) approach for solving minimization of the rate of ℓ_1 and ℓ_2 norms for sparse recovery. We first transform the quotient of ℓ_1 and ℓ_2 norms into a new function of the separable variables using the least squares minimum norm solution of the linear system of equations. Subsequently, we employ the augmented Lagrangian function to formulate the corresponding ADMM method with a dynamically adjustable parameter. Additionally, each of its subproblems possesses a unique global minimum. Finally, we present some numerical experiments to demonstrate our results.

Keywords Sparse recovery · The rate of ℓ_1 and ℓ_2 · Alternating direction method of multipliers · The augmented Lagrangian · The least squares minimum norm solution

1 Introduction

In various fields of science and engineering, such as sparse optimization [1], machine learning [2], image processing [3] and geophysics [4], one aims to find the sparsest signal within an underdeterminant linear system of $\mathbf{Ax} = \mathbf{b}$. This problem is commonly known as compressed sensing (CS) [5–8], where the signal $\mathbf{x}$ is expected to exhibit sparsity or compressibility. Mathematically, the core problem in CS can be formulated as the cardinality minimization

✉ Jun Wang
jwang77@just.edu.cn

Qiang Ma
201900000093@just.edu.cn

1 School of Science, Jiangsu University of Science and Technology, Zhenjiang 212003, China

2 School of Materials Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China

Published online: 26 March 2024

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{A}\mathbf{x}=\mathbf{b} \tag{1}$$

where $\|\mathbf{x}\|_0$ denotes the number of nonzero entries in $\mathbf{x}$. Although problem (1) can be approximately solved by some greedy algorithms, such as OMP [9], CoSaMP [10] and others [11, 12], the ℓ_0 minimization (1) is known to be NP-hard [14].

Alternatively, there are numerous approximation/relaxation methods available to substitute the ℓ_0 quasi-norm with alternative convex/nonconvex metric functions. One popular convex relaxation in CS replaces the ℓ_0 quasi-norm by the ℓ_1 norm, leading to the basis pursuit (BP) problem [15, 16]

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

which can be efficiently solved by linear programming and other optimization algorithms, such as the fast iterative shrinkage-thresholding algorithm (FISTA) [17], the alternating direction method of multipliers (ADMM) [18] and others [11, 12]. Some other non-convex models used to approximately solve (1), including the ℓ_p $(0<p<1)$ minimization [19], the ratio and difference of ℓ_1 and ℓ_2 norms [20, 21], transformed ℓ_1 [22], capped-ℓ_1 [23] and mixed ℓ_1 and ℓ_0 norms [24], the ratio and difference of ℓ_1 and ℓ_∞ norms [25, 26] and others [13].

In this paper, we provide a novel method for solving the ratio ℓ_1 and ℓ_2 model

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\|\mathbf{x}\|_2} \text{ s.t. } \mathbf{A}\mathbf{x}=\mathbf{b} \tag{3}$$

where $\mathbf{A}\in\mathbb{R}^{m\times n}(m\ll n)$ is a known sensing matrix of full row rank and $\mathbf{b}\in\mathbb{R}^m$ is an observation vector.

Problem (3) is known as a scale-invariant [27] and parameter-free metric function, commonly used to approximate the desired scale-invariant ℓ_0 quasi-norm. The ratio was initially introduced by Hoyer [28] in the context of non-negative sparse coding and subsequently examined in the realm of sparse promoting regularizations [29]. Yin et. al. [20] theoretically established the equivalence between the ℓ_1/ℓ_2 and the ℓ_0 minimization for recovering non-negative signal under linear constraints. Recently, Rahimi [27] demonstrated that any sparse vector is a local minimum point of ℓ_1/ℓ_2 model by employing the strong null space property condition. Zeng, Yu and Pong [30] investigated a range of constrained ℓ_1/ℓ_2 regularizations, while Xu et. al. [31] provided a sufficient condition for an s-sparse signal to be the global minimum of ℓ_1/ℓ_2 minimization under linear equality constraints. Tao [32] derived a closed-form solution for the proximal operator of the ℓ_1/ℓ_2 function, designed a specific variable-splitting scheme of applying ADMM and then established its global convergence and linear convergence rate under certain conditions. Furthermore, the ℓ_1/ℓ_2 models have been successfully employed in blind deconvolution [33] and limited-angle CT reconstruction [34], demonstrating the superiority of applying ℓ_1/ℓ_2 on the gradient over classical total variation in enforcing image gradient sparsity.

Given that problem (3) is a non-convex optimization, the commonly employed approach to address it is the alternating direction method of multipliers (ADMM)

[18]. By introducing two auxiliary variables, we can rewrite (3) into its equivalent form

$$\begin{aligned} &\min_{\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{R}^n} \frac{\|\mathbf{z}\|_1}{\|\mathbf{y}\|_2} + \mathcal{I}(\mathbf{A}\mathbf{x}-\mathbf{b}) \\ &\text{s.t.} \begin{cases} \mathbf{x}-\mathbf{y} &= \mathbf{0} \\ \mathbf{x}-\mathbf{z} &= \mathbf{0} \end{cases} \end{aligned} \tag{4}$$

where $\mathcal{I}$ is an indicator function defined as

$$\mathcal{I}(\mathbf{t}) = \begin{cases} 0, & \mathbf{t} = \mathbf{0}, \\ +\infty, & \text{otherwise.} \end{cases}$$

The augmented Lagrangian of (4)

$$\begin{aligned} \mathcal{L}_{\rho_1,\rho_2}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{v},\mathbf{w}) = \frac{\|\mathbf{z}\|_1}{\|\mathbf{y}\|_2} + \mathcal{I}(\mathbf{A}\mathbf{x}-\mathbf{b}) + \langle \mathbf{v}, \mathbf{x}-\mathbf{y}\rangle + \frac{\rho_1}{2}\|\mathbf{x}-\mathbf{y}\|_2^2 \\ + \langle \mathbf{w}, \mathbf{x}-\mathbf{z}\rangle + \frac{\rho_2}{2}\|\mathbf{x}-\mathbf{z}\|_2^2 \end{aligned}$$

deduces the iterations of ADMM

$$\begin{cases} \mathbf{x}^{k+1} &= \operatorname{argmin}_{\mathbf{x}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x},\mathbf{y}^k,\mathbf{z}^k,\mathbf{v}^k,\mathbf{w}^k\right) \\ \mathbf{y}^{k+1} &= \operatorname{argmin}_{\mathbf{y}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x}^{k+1},\mathbf{y},\mathbf{z}^k,\mathbf{v}^k,\mathbf{w}^k\right) \\ \mathbf{z}^{k+1} &= \operatorname{argmin}_{\mathbf{z}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x}^{k+1},\mathbf{y}^{k+1},\mathbf{z},\mathbf{v}^k,\mathbf{w}^k\right) \\ \mathbf{v}^{k+1} &= \mathbf{v}^k + \rho_1\left(\mathbf{x}^{k+1}-\mathbf{y}^{k+1}\right) \\ \mathbf{w}^{k+1} &= \mathbf{w}^k + \rho_2\left(\mathbf{x}^{k+1}-\mathbf{z}^{k+1}\right) \end{cases} \tag{5}$$

which is summarized in [27, Algorithm 3.1].

Motivated by the effective performance of ADMM and its variants in solving (3), this paper introduces a simplified version of ADMM, termed the reduced form (8), achieved by decomposing the objective function of (3) into a separable continuous function of variables. The main contributions of this paper consist of three schemes for minimizing the ℓ_1/ℓ_2 model, namely:

(1) By utilizing the least squares minimum norm solution $\mathbf{x}^+$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$, we transform the constrained optimization problem (3) into a separable problem (6) with a linear equality constraint through the introduction of constrained auxiliary variables.
(2) By utilizing the least squares minimum norm solution $\mathbf{x}^+$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$, we transform the constrained optimization problem (3) into a separable problem (6) with a linear equality constraint through the introduction of constrained auxiliary variables.
(3) For the $\mathbf{y}$-subproblem, we efficiently obtain its optimal solution by using either Gauss-Newton's method 1 or solving the linear equations (20), an advantage not found in other variations of ADMM. Moreover, instead of specifying the initial point for the $\mathbf{x}$-subproblem, we only need to provide an initial point $\mathbf{y}^0 = \mathbf{0} \in \mathbb{R}^n$

for the $\mathbf{y}$-subproblem, thereby significantly reducing the impact of initial points on Algorithm 2.

The structure of the remaining paper is as follows. In Sect. 2, we employ the orthogonal decomposition method of the subspace to derive a variant of ADMM. Section 3 presents a set of experiments designed to verify the validity and effectiveness of Algorithm 2 in comparison to ADMM 5 for sparse recovery.

2 Main Results

Before proceeding, we introduce the notations used in this paper. $\mathbb{R}^n$ represents the n-dimensional Euclidean space, while $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. Matrices are represented by bold capital letters, vectors by bold lowercase letters, and scalars or entries by regular font. The transpose of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is denoted as $\mathbf{A}^\top$. In particular, $\mathbf{x} = \left(x_1, x_2, \dots, x_n\right)^\top \in \mathbb{R}^n$ is a vector, where x_i is the i-th component. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, their inner product is represented by $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^\top \mathbf{y}$.

Denoting the least square solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ by $\mathbf{x}^+ = \mathbf{A}^+\mathbf{b}$, where $\mathbf{A}^+$ is the Moore–Penrose pseudoinverse, we obtain that $\mathbf{x} = \mathbf{y} + \mathbf{x}^+$ is the general solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $\mathbf{A}\mathbf{y} = \mathbf{0}$. It follows from

$$\langle \mathbf{y}, \mathbf{x}^+\rangle = \langle \mathbf{y}, \mathbf{A}^+\mathbf{b}\rangle = \langle \mathbf{y}, \mathbf{A}^+\mathbf{A}\mathbf{x}\rangle = \langle \left(\mathbf{A}^+\mathbf{A}\right)^\top \mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{A}^+\mathbf{A}\mathbf{y}, \mathbf{x}\rangle = 0$$

that $\|\mathbf{x}\|_2^2 = \|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2$ and therefore problem (3) can be expressed as

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\sqrt{\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \mathcal{I}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mathcal{I}(\mathbf{A}\mathbf{y}) \text{ s.t. } \mathbf{x} - \mathbf{y} = \mathbf{x}^+. \tag{6}$$

We introduce a Lagrangian multiplier λ and the parameter $\rho > 0$ and therefore formed the augmented Lagrangian function of (6) as

$$\begin{aligned}\mathcal{L}_\rho(\mathbf{x}, \mathbf{y}, \lambda) = & \frac{\|\mathbf{x}\|_1}{\sqrt{\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \mathcal{I}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \frac{\rho}{2}\|\mathbf{x} - \mathbf{y} - \mathbf{x}^+\|_2^2 \\ & + \left\langle \lambda, \mathbf{x} - \mathbf{y} - \mathbf{x}^+ \right\rangle + \mathcal{I}(\mathbf{A}\mathbf{y})\end{aligned} \tag{7}$$

where $\lambda \in \mathbb{R}^n$ and $\rho > 0$.

The ADMM consists of the following iterations

$$\begin{cases} \mathbf{x}^{k+1} &= \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n} \mathcal{L}_\rho\left(\mathbf{x}, \mathbf{y}^k, \lambda^k\right) \\ \mathbf{y}^{k+1} &= \operatorname{argmin}_{\mathbf{y}\in\mathbb{R}^n} \mathcal{L}_\rho\left(\mathbf{x}^{k+1}, \mathbf{y}, \lambda^k\right) \\ \lambda^{k+1} &= \lambda^k + \rho\left(\mathbf{x}^{k+1} - \mathbf{y}^{k+1} - \mathbf{x}^+\right) \end{cases} \tag{8}$$

with the initial points $\mathbf{y}^0 = \mathbf{0} \in \mathbb{R}^n$ and $\lambda^0 = \mathbf{0} \in \mathbb{R}^n$.

For the $\mathbf{x}$-subproblem, we obtain the following smoothed basis pursuit problem

$$\begin{aligned} \mathbf{x}^{k+1} &= \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n} \mathcal{L}_\rho\left(\mathbf{x}, \mathbf{y}^k, \lambda^k\right) \\ &= \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\sqrt{\|\mathbf{y}^k\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \frac{\rho}{2}\|\mathbf{x} - \mathbf{y}^k - \mathbf{x}^+\|_2^2 \\ &\quad + \mathcal{I}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \langle \lambda^k, \mathbf{x} - \mathbf{y}^k - \mathbf{x}^+ \rangle \\ &= \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\sqrt{\|\mathbf{y}^k\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \mathcal{I}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \frac{\rho}{2}\|\mathbf{x} - \mathbf{u}^k\|_2^2, \end{aligned} \tag{9}$$

where $\mathbf{u}^k = \mathbf{y}^k + \mathbf{x}^+ - \lambda^k/\rho$. Observe that, by the fact that problem (9) is a strongly convex programming with $\rho > 0$, if (9) has a solution, then we obtain that it has a unique optimal solution and further solve it by using the Matlab tool CVX [39] with the solver Mosek. In addition, it is worth noting that the optimal solution $\mathbf{x}^1$ of the $\mathbf{x}$ -subproblem is not a solution of BP. In fact, if $k = 0$, problem (9) is equivalent to the following problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\|\mathbf{x}^+\|_2} + \frac{\rho}{2}\|\mathbf{x} - \mathbf{x}^+\|_2^2 \ \text{ s.t. } \mathbf{A}\mathbf{x} = \mathbf{b}$$

which is clearly different from BP (2). The above analysis about $\mathbf{x}$-subproblem shows that the proposed iterations of ADMM (8) does not start from a solution of BP. Hence, to some extent, this algorithm should have stronger robustness compared to algorithms that start with the solution of the basis pursuit (BP) problem.

Now, let us consider the $\mathbf{y}$-subproblem.

$$\begin{aligned} \mathbf{y}^{k+1} &= \operatorname{argmin}_{\mathbf{y}\in\mathbb{R}^n} \mathcal{L}_\rho\left(\mathbf{x}^{k+1}, \mathbf{y}, \lambda^k\right) \\ &= \operatorname{argmin}_{\mathbf{y}\in\mathbb{R}^n} \frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \mathcal{I}(\mathbf{A}\mathbf{y}) + \langle \lambda^k, \mathbf{x}^{k+1} - \mathbf{y} - \mathbf{x}^+ \rangle \\ &\quad + \frac{\rho}{2}\|\mathbf{x}^{k+1} - \mathbf{y} - \mathbf{x}^+\|_2^2 \\ &= \operatorname{argmin}_{\mathbf{A}\mathbf{y}=\mathbf{0}} \frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \frac{\rho}{2}\|\mathbf{y} - \mathbf{v}^k\|_2^2 =: \varphi_k(\mathbf{y}), \end{aligned} \tag{10}$$

where $\mathbf{v}^k = \mathbf{x}^{k+1} - \mathbf{x}^+ + \lambda^k/\rho$.

Clearly, the function $\varphi_k(\mathbf{y})$ defined by (10) is twice continuously differentiable. From the Lagrangian function corresponding to problem (10)

$$\mathscr{L}_k(\mathbf{y}, \mathbf{r}) \triangleq \varphi_k(\mathbf{y}) + \langle \mathbf{r}, \mathbf{A}\mathbf{y} \rangle,$$

we obtain that the first-order optimality conditions for $\bar{\mathbf{y}}^{k+1}$ to be an optimal solution to problem (10), which state that there exist multipliers $\bar{\mathbf{r}}^{k+1}$ such that $\left(\bar{\mathbf{y}}^{k+1}, \bar{\mathbf{r}}^{k+1}\right)$ is a solution to the nonlinear system of equations

$$\begin{cases} \nabla_{\mathbf{y}}\mathscr{L}_k(\mathbf{y},\mathbf{r}) = \nabla\varphi_k(\mathbf{y}) + \mathbf{A}^\top\mathbf{r} & = \mathbf{0}, \\ \nabla_{\mathbf{r}}\mathscr{L}_k(\mathbf{y},\mathbf{r}) = \mathbf{A}\mathbf{y} & = \mathbf{0}. \end{cases} \tag{11}$$

Substituting the derivative of $\varphi_k(\mathbf{y})$ with respect to $\mathbf{y}$

$$\nabla\varphi_k(\mathbf{y}) = \left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^{\frac{3}{2}}}\right)\mathbf{y} - \rho\mathbf{v}^k \tag{12}$$

into (11) yields

$$\left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^{\frac{3}{2}}}\right)\mathbf{y} + \mathbf{A}^\top\mathbf{r} = \rho\mathbf{v}^k. \tag{13}$$

Multiplying both sides of (13) by $\mathbf{A}$ gives

$$\left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^{\frac{3}{2}}}\right)\mathbf{A}\mathbf{y} + \mathbf{A}\mathbf{A}^\top\mathbf{r} = \rho\mathbf{A}\mathbf{v}^k$$

which admits a unique multiplier

$$\bar{\mathbf{r}}^{k+1} = \rho\left(\mathbf{A}\mathbf{A}^\top\right)^{-1}\mathbf{A}\mathbf{v}^k,$$

since $\mathbf{A}\mathbf{y} = \mathbf{0}$ and $\mathbf{A}$ is of row-full rank. From the above formula and (13), we obtain

$$\left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^{\frac{3}{2}}}\right)\mathbf{y} = \rho\mathbf{P}_A\mathbf{v}^k \tag{14}$$

where $\mathbf{P}_A \triangleq \mathbf{I}_n - \mathbf{A}^\top\left(\mathbf{A}\mathbf{A}^\top\right)^{-1}\mathbf{A}$, and therefore solve $\bar{\mathbf{y}}^{k+1}$ by using the Gauss-Newton's method [35], trust-region(-dogleg) [37] or Levenberg-Marquardt algorithm [38].

However, the condition (11) is only necessary, which does not guarantee that the solution $\bar{\mathbf{y}}^{k+1}$ is the optimal solution of problem (10). Therefore, we need to further consider its second-order sufficient condition. With the help of $\nabla\varphi_k(\mathbf{y})$ defined by (12), we obtain that the Hessian matrix of $\varphi_k(\mathbf{y})$ is

$$\begin{aligned}\nabla^2\varphi_k(\mathbf{y}) = {} & \left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^3}}\right)\mathbf{I}_n \\ & + \frac{3\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^5}}\mathbf{y}\mathbf{y}^\top,\end{aligned} \tag{15}$$

which comes from

$$\frac{\mathbf{y}}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}=\left(\frac{y_1}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}},\dots,\frac{y_n}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}\right)^\top$$

and

$$\frac{\partial}{\partial y_j}\left(\frac{y_i}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}\right)=\begin{cases}\frac{3y_iy_j}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^5}}, & i\neq j,\\ \frac{3y_i^2}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^5}}-\frac{1}{\sqrt{\left(\|\mathbf{y}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}, & i=j,\end{cases}$$

with the indices $i, j = 1, 2, \dots, n$.

Assume that there exists two nonzero vectors $\bar{\mathbf{y}}^{k+1}$ and $\bar{\mathbf{r}}^{k+1}$ such that

$$\begin{cases}\nabla_{\mathbf{y}}\mathscr{L}_k\left(\bar{\mathbf{y}}^{k+1},\bar{\mathbf{r}}^{k+1}\right) &= \mathbf{0},\\ \nabla_{\mathbf{r}}\mathscr{L}_k\left(\bar{\mathbf{y}}^{k+1},\bar{\mathbf{r}}^{k+1}\right) &= \mathbf{0},\end{cases}$$

if the Hessian matrix

$$\nabla^2_{\mathbf{y}}\mathscr{L}_k\left(\bar{\mathbf{y}}^{k+1},\bar{\mathbf{r}}^{k+1}\right)=\nabla^2_{\mathbf{y}}\varphi_k\left(\bar{\mathbf{y}}^{k+1}\right)$$

is a symmetric positive definite matrix, then we obtain that $\bar{\mathbf{y}}^{k+1}$ is a strictly local minimum of problem (10). Fortunately, this sufficient condition can be achieved by setting suitable parameter ρ. It follows from $\mathscr{L}_k(\mathbf{y},\mathbf{r})$ and (15) that

$$\begin{aligned}\mathbf{z}^\top\nabla^2_{\mathbf{y}}\mathscr{L}_k\left(\bar{\mathbf{y}}^{k+1},\bar{\mathbf{r}}^{k+1}\right)\mathbf{z}=\mathbf{z}^\top\nabla^2\varphi_k\left(\bar{\mathbf{y}}^{k+1}\right)\mathbf{z}&\\ =\left(\rho-\frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\bar{\mathbf{y}}^{k+1}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}\right)\|\mathbf{z}\|_2^2&\\ +\frac{3\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\bar{\mathbf{y}}^{k+1}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^5}}\left(\mathbf{z}^\top\bar{\mathbf{y}}^{k+1}\right)^2,&\end{aligned}$$

for any nonzero vector $\mathbf{z}\in\mathbb{R}^n$. Combining $\|\mathbf{z}\|_2^2>0$ and $\left(\mathbf{z}^\top\bar{\mathbf{y}}^{k+1}\right)^2\geq 0$ gives that if

$$\rho-\frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\bar{\mathbf{y}}^{k+1}\|_2^2+\|\mathbf{x}^+\|_2^2\right)^3}}>0, \tag{16}$$

then $\mathbf{z}^\top\nabla^2\varphi_k\left(\bar{\mathbf{y}}^{k+1}\right)\mathbf{z}>0$. In order to make (16) holds, we can set

$$\rho > \max_{\mathbf{Ay}=\mathbf{0}} \frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^3}} = \max_{\mathbf{Ay}=\mathbf{0}} \frac{\|\mathbf{x}^{k+1}\|_1}{\sqrt{\left(\|\mathbf{x}^+\|_2^2\right)^3}}$$

and therefore have

$$\rho > \frac{\|\mathbf{x}^{k+1}\|_1}{\|\mathbf{x}^+\|_2^3}. \tag{17}$$

Hence, the above analysis yields the following important result.

Proposition 1 *Suppose that $\bar{\mathbf{y}}^{k+1}$ satisfies nonlinear system of equations* (14). *If the condition* (17) *holds, then $\bar{\mathbf{y}}^{k+1}$ is a unique global minimum of problem* (10).

Proof Firstly, the constrained problem

$$\min_{\mathbf{y}\in\mathbb{R}^n} \varphi_k(\mathbf{y}) \text{ s.t. } \mathbf{Ay} = \mathbf{0}$$

is equivalent to the following unconstrained minimization

$$\min_{\mathbf{y}\in\mathbb{R}^n} \phi_k(\mathbf{y}) \triangleq \varphi_k(\mathbf{y}) + \mathcal{I}(\mathbf{Ay})$$

with the help of the indicator function $\mathcal{I}$. From the condition (17) and the fact that

$$\nabla^2_{\mathbf{y}}\phi_k(\mathbf{y}) = \nabla^2_{\mathbf{y}}\varphi_k(\mathbf{y})$$

a symmetric positive definite matrix, we obtain that $\phi_k(\mathbf{y})$ is a strongly convex function [36]. Then, we obtain from (14) that $\mathbf{A}\bar{\mathbf{y}}^{k+1} = \mathbf{0}$ and then $\bar{\mathbf{y}}^{k+1}$ is a unique global minimum of problem (10).

Finally, the proof is completed. □

The parameter ρ defined by (17) ensures that $\mathbf{y}^{k+1} = \bar{\mathbf{y}}^{k+1}$ is a unique solution of problem (10) and meanwhile there is no ill-condition case for solving $\bar{\mathbf{y}}^{k+1}$ by nonlinear equations (18). In addition, Proposition 1 gives us one idea of solving problem (10). With (17) and symmetric positive definite matrix $\nabla^2\varphi_k(\mathbf{y})$, we provide the Gauss-Newton's method of solving the system of nonlinear equations (14) which is expanded as

$$\mathbf{F}_k(\mathbf{y}) \triangleq \left(\rho - \frac{\|\mathbf{x}^{k+1}\|_1}{\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}^+\|_2^2\right)^{3/2}}\right)\mathbf{y} - \rho\mathbf{P}_A\mathbf{v}^k = \mathbf{0} \tag{18}$$

and its Jacobi matrix $J_{\mathbf{F}_k}(\mathbf{y}) = \nabla^2\varphi_k(\mathbf{y})$. In order to find a solution $\bar{\mathbf{y}}^{k+1} = \bar{\mathbf{y}}^{k_{\ell+1}}$ to a multivariable nonlinear equation (18), we consider

$$\min_{\mathbf{y}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{F}_k(\mathbf{y})\|_2^2 \tag{19}$$

and then construct Gauss-Newton's method which is summarized in Algorithm 1.

Algorithm 1 The Gauss-Newton's method for (19)

Input: $\bar{\mathbf{y}}^{k_0} \in \mathbb{R}^n$, $\varepsilon_k = 10^{-8}$ and $\ell := 0$.

while $\|\mathbf{F}_k\left(\bar{\mathbf{y}}^{k_\ell}\right)\|_2 > \varepsilon_k$ **do**

(1) Compute $\bar{\mathbf{d}}^{(\ell)}$ of

$$J_{\mathbf{F}_k}^{\top}\left(\bar{\mathbf{y}}^{k_\ell}\right) J_{\mathbf{F}_k}\left(\bar{\mathbf{y}}^{k_\ell}\right) \mathbf{d} = -J_{\mathbf{F}_k}(\mathbf{y})\,\mathbf{F}_k\left(\bar{\mathbf{y}}^{k_\ell}\right),$$

(2) Calculus the step α_{k_ℓ} with the line search of Wolfe-Powell rule,

(3) Update

$$\begin{cases} \bar{\mathbf{y}}^{k_{\ell+1}} & = \bar{\mathbf{y}}^{k_\ell} + \alpha_{k_\ell}\bar{\mathbf{d}}^{(\ell)}, \\ \ell & = \ell + 1. \end{cases}$$

end

Output: Finally solution $\bar{\mathbf{y}}^{k_{\ell+1}}$.

Remark 1 The key aspects of Algorithm 1 involve the selection of an appropriate initial point $\bar{\mathbf{y}}^{k_0}$ and ensuring that the coefficient matrix $J_{\mathbf{F}_k}(\mathbf{y})$ is invertible.

(*i*) For the initial point $\bar{\mathbf{y}}^{k_0}$, a tactic is suggested to convert the system of nonlinear equations (18) into linear equations. Referring to our model (6), additional constraints are incorporated namely, $\mathbf{x} = \mathbf{y} + \mathbf{x}^+$ and $\mathbf{A}\mathbf{y} = \mathbf{0}$. This approach ensures that the solution $\mathbf{y}^{k+1}$ of the $\mathbf{y}$-subproblem (10) satisfies $\mathbf{x}^{k+1} \approx \mathbf{y}^{k+1} + \mathbf{x}^+$ and $\mathbf{A}\mathbf{y}^{k+1} = \mathbf{0}$ as closely as possible. By the equation $\|\mathbf{y}^{k+1}\|_2^2 + \|\mathbf{x}^+\|_2^2 \approx \|\mathbf{x}^{k+1}\|_2^2$, (18) can be approximated by

$$\widehat{\mathbf{y}}^{k+1} = \left(1 - \frac{\|\mathbf{x}^{k+1}\|_1}{\rho\|\mathbf{x}^{k+1}\|_2^3}\right)^{-1} \mathbf{P}_A \mathbf{v}^k. \tag{20}$$

Indeed, it is straightforward to calculate:

$$\frac{\|\mathbf{x}^{k+1}\|_1}{\|\mathbf{x}^{k+1}\|_2^3} \leq \frac{\|\mathbf{x}^{k+1}\|_1}{\|\mathbf{x}^{+}\|_2^3} < \rho$$

using the conditions $\|\mathbf{x}^{k+1}\|_2 \geq \|\mathbf{x}^+\|_2$ and $\mathbf{A}\mathbf{x}^{k+1} = b$. As a result, we establish:

$$\bar{\mathbf{y}}^{k_0} = \widehat{\mathbf{y}}^{k+1}.$$

(*ii*) By applying (17), it is understood that $\nabla^2\varphi_k(\bar{\mathbf{y}}^{\ell})$ is a symmetric positive definite matrix. This validates that Gauss-Newton's step can be effectively computed, circumventing any ill-condition cases.

In summary, we will add the condition (17) of ρ into the ADMM iterations and summarize them in Algorithm 2.

Algorithm 2 The $L1/L2$ minimization via ADMM with dynamic parameters (ADMMdp)

Data: Initial $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, IterMax, ε
Initialization :
$k = 0$, $\rho_x^0 = 100$, $\mathbf{x}^+ = \mathbf{A}^+\mathbf{b}$, $\lambda^0 = \mathbf{0} \in \mathbb{R}^n$, $\mathbf{y}^0 = \mathbf{0} \in \mathbb{R}^n$;
while $k \le$ *IterMax or* $\frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|_2}{\|\mathbf{x}^{k-1}\|_2} > \varepsilon$ **do**
Step 1 Compute

$$\mathbf{x}^{k+1} = \mathrm{argmin}_{\mathbf{Ax}=\mathbf{b}} \frac{\|\mathbf{x}\|_1}{\sqrt{\|\mathbf{y}^k\|_2^2 + \|\mathbf{x}^+\|_2^2}} + \frac{\rho_x^k}{2}\|\mathbf{x} - \mathbf{u}^k\|_2^2,$$

where $\mathbf{u}^k = \mathbf{y}^k + \mathbf{x}^+ - \lambda^k/\rho_x^k$.
Step 2 Select $\rho_y^{k+1} = \frac{\|\mathbf{x}^{k+1}\|_2}{\|\mathbf{x}^+\|_2^3}$ and then set
(i) $\mathbf{y}^{k+1} = \bar{\mathbf{y}}^{k_{\ell+1}}$ by Algorithm 1;
(ii) $\mathbf{y}^{k+1} = \hat{\mathbf{y}}^{k+1}$ with $\rho = \rho_y^{k+1}$.
Step 3 Compute

$$\lambda^{k+1} = \lambda^k + \rho_y^{k+1}\left(\mathbf{x}^{k+1} - \mathbf{y}^{k+1} - \mathbf{x}^+\right)$$

Step 4 Update

$$\begin{cases} \rho_x^{k+1} & = \rho_y^{k+1} \\ k & = k+1 \end{cases}$$

end
Result: $\mathbf{x}^* = \mathbf{x}^{k+1}$, $\mathbf{y}^* = \mathbf{y}^{k+1}$ and $\rho_{\text{end}} = \rho_x^{k+1}$

3 Numerical experiments

In this section, some experiments are given to illustrate the efficiency and performance of Algorithm 2 for solving optimization problem (3). The numerical testing is carried out on a Dell Inc. Precision 5820 Tower X-Series with Intel(R) Core(TM) i9-10900X (3.7GHz) and 32 (4×8) GB of UDIMM memory and Matlab R2020b installed in the Linux Mint 20.1 Ulyssa (x86_64 Linux 5.14.0-1058-oem).

Regarding the redundantly introduced variable z in [27, Algorithm 3.1], it is only used to simplify the computation of the subproblem and is not enforced to satisfy the linear equation constraint $Az = b$. Hence, we give the following modifications to iterations forms (5).

$$\begin{cases} \mathbf{x}^{k+1} & = \mathrm{argmin}_{\,\mathbf{x}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x}, \mathbf{y}^k, \mathbf{z}^k, \mathbf{v}^k, \mathbf{w}^k\right) \\ \mathbf{y}^{k+1} & = \mathrm{argmin}_{\,\mathbf{y}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x}^{k+1}, \mathbf{y}, \mathbf{z}^k, \mathbf{v}^k, \mathbf{w}^k\right) \\ \mathbf{z}^{k+1} & = \mathrm{argmin}_{\,\mathbf{Az}=\mathbf{b}} \mathcal{L}_{\rho_1,\rho_2}\left(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{z}, \mathbf{v}^k, \mathbf{w}^k\right) \\ \mathbf{v}^{k+1} & = \mathbf{v}^k + \rho_1\left(\mathbf{x}^{k+1} - \mathbf{y}^{k+1}\right) \\ \mathbf{w}^{k+1} & = \mathbf{w}^k + \rho_2\left(\mathbf{x}^{k+1} - \mathbf{z}^{k+1}\right) \end{cases} \tag{21}$$

The only difference between (5) and (21) is that we additionally require an equality constraint $\mathbf{Az} = \mathbf{b}$ about $\mathbf{z}$-subproblems. The reason for this consideration is that when we introduce an auxiliary variable z, there is no guarantee of the constraint $\mathbf{Az} = \mathbf{b}$. Simplifying (21), we obtain the following iterations:

$$
\begin{cases}
\mathbf{x}^{k+1} & = \left(\mathbf{I}_n - \mathbf{A}^{+}\mathbf{A}\right)\mathbf{f}^k + \mathbf{x}^{+} \\
\mathbf{y}^{k+1} & = \begin{cases} \mathbf{e}^k, & \mathbf{d}^k = \mathbf{0} \\ \tau^k \mathbf{d}^k, & \mathbf{d}^k \neq \mathbf{0} \end{cases} \\
\mathbf{z}^{k+1} & = \operatorname{argmin}_{\mathbf{Az}=\mathbf{b}} \|\mathbf{z}\|_1 + \frac{\beta_k}{2}\left\|\mathbf{z} - \left(\mathbf{x}^{k+1} + \frac{\mathbf{w}^k}{\rho_2}\right)\right\|_2^2 \\
\mathbf{v}^{k+1} & = \mathbf{v}^k + \rho_1\left(\mathbf{x}^{k+1} - \mathbf{y}^{k+1}\right) \\
\mathbf{w}^{k+1} & = \mathbf{w}^k + \rho_2\left(\mathbf{x}^{k+1} - \mathbf{z}^{k+1}\right)
\end{cases}
\tag{22}
$$

where

$$
\mathbf{f}^k = \frac{\rho_1}{\rho_1 + \rho_2}\left(\mathbf{y}^k - \frac{\mathbf{v}^k}{\rho_1}\right) + \frac{\rho_2}{\rho_1 + \rho_2}\left(\mathbf{z}^k - \frac{\mathbf{w}^k}{\rho_1}\right),
$$

the vector $\mathbf{d}^k = \mathbf{x}^{k+1} + \mathbf{v}^k/\rho_1$, $\beta_k = \left(\rho_2\|\mathbf{y}^{k+1}\|_2\right)/2$ and $\mathbf{e}^k, \tau_k$ defined in [27, Eq 3.9, pp. A3655], which denoted by `mADMMrw` and different from Algorithm 2 in this paper.

Firstly, we generate the ground truth signal $\mathbf{x}^{\mathrm{t}} \in \mathbb{R}^{1024}$ as a s-sparse vector following with s i.i.d. standard Gaussian entries at random positions with the line of data setting in [27, Sect. 4, Numerical experiments]. Next, we consider the sensing matrix $\mathbf{A} \in \mathbb{R}^{64\times 1024}$ with an oversampled discrete cosine transform (DCT), defined as

$$
\mathbf{a}_j \triangleq \frac{1}{\sqrt{64}} \cos\left(\frac{2\pi\omega j}{F}\right),\ j = 1, 2, \dots, 1024,
$$

where $\mathbf{a}_j$ is the j-column of $\mathbf{A}$ and ω is a random vector that is uniformly distributed in $[0, 1]^{64}$ and a positive parameter F to control the coherence in a way that a large F yields a more coherent matrix. Finally, the initial points of ADMM [27] are important and should be well chosen. The selected initial point directly affects the performance of the algorithm. The initial guess of `mADMMrw` is the solution to BP in [27]. However, this actually provides a very good start for `mADMMrw`, due to the fact that the solution of BP is to some extent available as a sparse solution. Hence, this potentially ignores the true recovery performance of `mADMMrw`. Compared with `mADMMrw`, we construct `ADMMdp` in such a way that there is no additional consideration of how to give the initial solution to the $\mathbf{y}$-subproblem, due to the fact that it has a default desirable initial point $\mathbf{y}^0 = \mathbf{0} \in \mathbb{R}^{1024}$. With the help of the above analysis, we set the initial point of `ADMMdp` and `mADMMrw` to be $\mathbf{0} \in \mathbb{R}^{1024}$. The stopping criterion is when the relative error is smaller than the given relative error $\varepsilon = 10^{-8}$ or the iterative number exceeds $10n$. Then, we consider the corresponding sparsity s of 12 and the parameter F of 10, 20 respectively. In addition, we set the value $\rho_x^0 = 100$ of `ADMMdp` and $\rho_1 = \rho_2 = \rho_{\mathrm{end}}$ and $\mathbf{y}^0, \mathbf{z}^0, \mathbf{v}^0, \mathbf{w}^0 = \mathbf{0} \in \mathbb{R}^{1024}$ in `mADMMrw`.

The empirical convergence and behavior of the proposed `ADMMdp` and `mADMMrw` are demonstrated in Fig. 1, Fig. 2, Fig. 3, and Fig. 5. These figures collectively show that both methods can optimally converge to the solution of problem (3), and, moreover, that `ADMMdp` outperforms `mADMMrw`. Primarily, `ADMMdp` requires approximately half the number of iterations that `mADMMrw` does. As

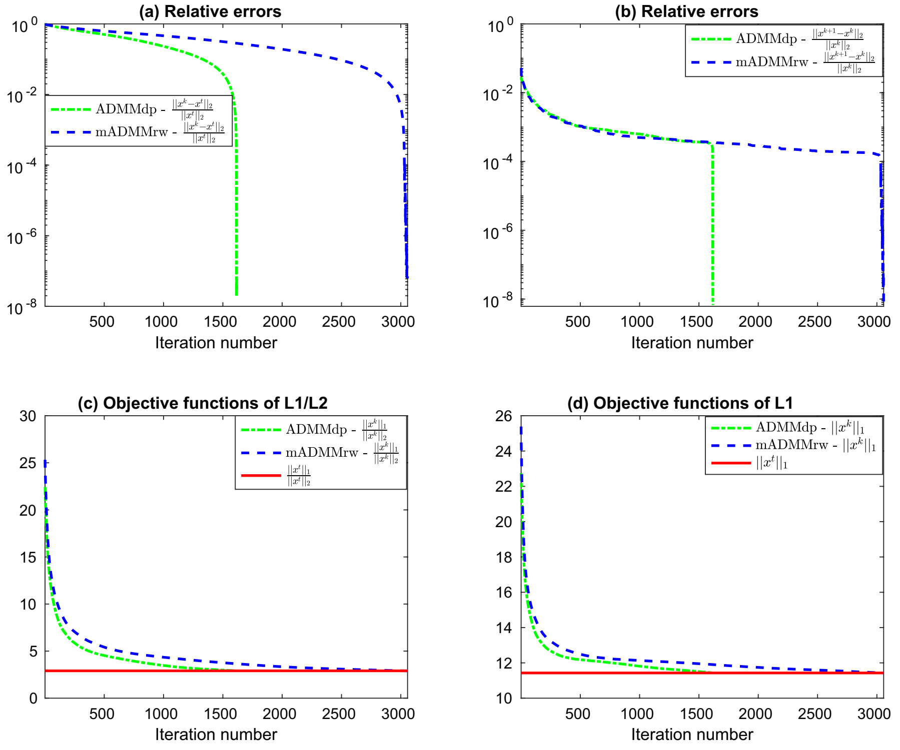

Fig. 1 Plots of residual errors and objective functions values for demonstrating the convergence and behaviors of the proposed `ADMMdp` and `mADMMrw` about variable $\mathbf{x}$ with the parameters $m = 64$, $n = 1024$ and $s = 12$, $F = 10$

shown in subsections (a) and (b) of Fig. 1 when $F = 10$, the iteration count for `ADMMdp` stands at 1618, whereas it is 3057 for `mADMMrw`.

Secondly, the major portion of computational time consumed by both algorithms is devoted to solving the $\mathbf{x}$-subproblems in `ADMMdp` and the $\mathbf{z}$-subproblems in `mADMMrw`. Given that these subproblems belong to the same class of optimization problem, albeit with different parameters, an equal number of iterations would lead to roughly equivalent CPU time being required to solve these subproblems.

With the number of iterations showed in four Figures, we should be able to obtain that the CPU time of `ADMMdp` is also roughly half of that of `mADMMrw`. In fact, the computation time of `ADMMdp` is 830.5787(*sec.*), while that of `mADMMrw` is 1573.9687(*sec.*). Hence, the rate 0.5277 of the computation time of `ADMMdp` and `mADMMrw` is in line with our previous speculation.

Furthermore, for these auxiliary variables $\mathbf{y}$ of `ADMMdp` and $\mathbf{y},\mathbf{z}$ of `mADMMrw` introduced in the ℓ_1/ℓ_2 model, the (e) of Fig. 2 and 5 indicate that the values of $\|\mathbf{x}^k - \mathbf{y}^k\|_2$ of `ADMMdp` drops rapidly to a constant value $\|\mathbf{x}^+\|_2 \approx 1.002537$ within five steps, while the values of $\|\mathbf{x}^k - \mathbf{y}^k\|_2$ of `mADMMrw` is rapidly decreasing, but

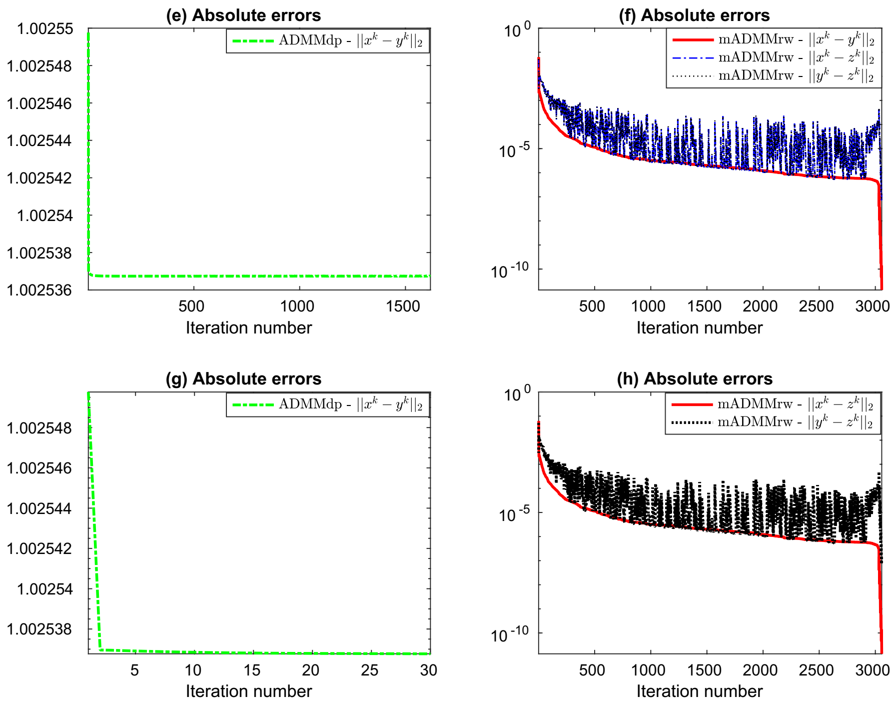

Fig. 2 Plots of residual errors and objective functions values for demonstrating the convergence and behaviors of the proposed `ADMMdp` and `mADMMrw` about variable $\mathbf{y}$ with the parameters $m = 64$, $n = 1024$ and $s = 12$, $F = 10$

the values of $\|\mathbf{x}^k - \mathbf{z}^k\|_2$ and $\|\mathbf{y}^k - \mathbf{z}^k\|_2$ of `mADMMrw` oscillate down throughout the iterations. On the other hand, this also illustrates that the introduction $\mathbf{x} = \mathbf{z}$ in problem (4) is redundant and result in an increase about the number of iterations, which further slows down the convergence rate of `mADMMrw`. As a consequence, it is reasonable to consider the following model:

$$\min_{\mathbf{x},\mathbf{y}\in\mathbb{R}^n} \frac{\|\mathbf{x}\|_1}{\|\mathbf{y}\|_2} + \mathcal{I}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mathcal{I}(\mathbf{A}\mathbf{y} - \mathbf{b}) \text{ s.t. } \mathbf{y} = \mathbf{x} \tag{23}$$

which is more difficult to be solved by a corresponding variant of ADMM than problem (4).

When the proposed Algorithm 2 is stable after about 2000 iterations, we also seem to get the optimal solution $\mathbf{y}^*$ to the following problem

$$\mathcal{K}(\mathbf{y}^*) \triangleq \min_{\mathbf{y}\in\mathbb{R}^n} \left\{ \frac{\|\mathbf{y}\|_1}{\|\mathbf{y}\|_2} : \mathbf{A}\mathbf{y} = \mathbf{0}, \mathbf{y} \neq \mathbf{0} \right\}. \tag{24}$$

In fact, solving this optimization problem directly is very difficult, and the technique provided in this paper would not be adaptable.The result is illustrated in Fig. 4.

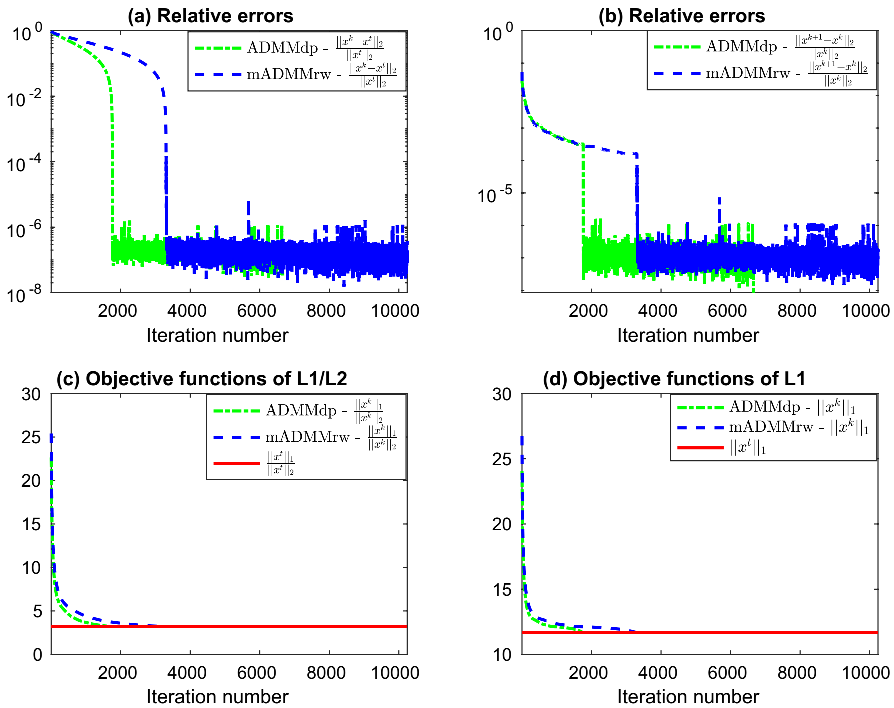

Fig. 3 Plots of residual errors and objective functions values for demonstrating the convergence and behaviors of the proposed `ADMMdp` and `mADMMrw` algorithm about variable $\mathbf{x}$ with the parameters $m = 64$, $n = 1024$ and $s = 12$, $F = 20$

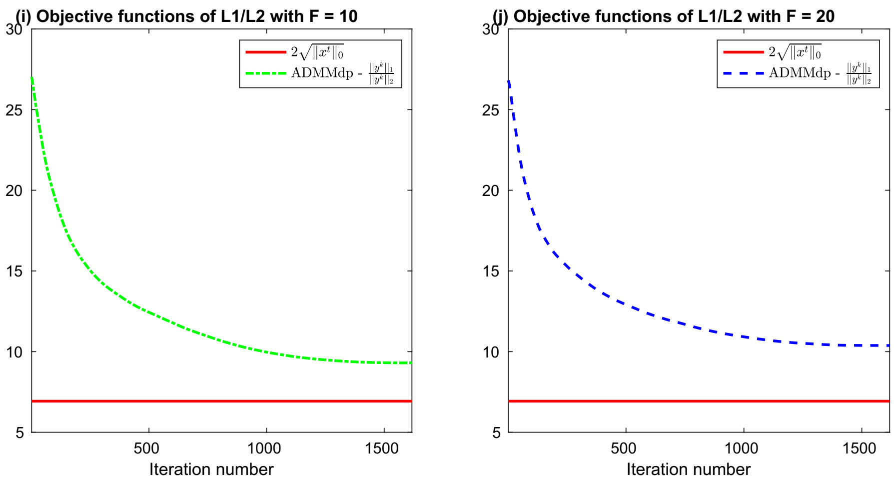

Fig. 4 Plots of the values $\frac{\|\mathbf{y}^k\|_1}{\|\mathbf{y}^k\|_2}$ for demonstrating the convergence and behaviors of the proposed `ADM-Mdp` algorithm about variable $\mathbf{y}$ with the parameters $m = 64$, $n = 1024$ and $s = 12$, $F = 10$(left), 20(right)

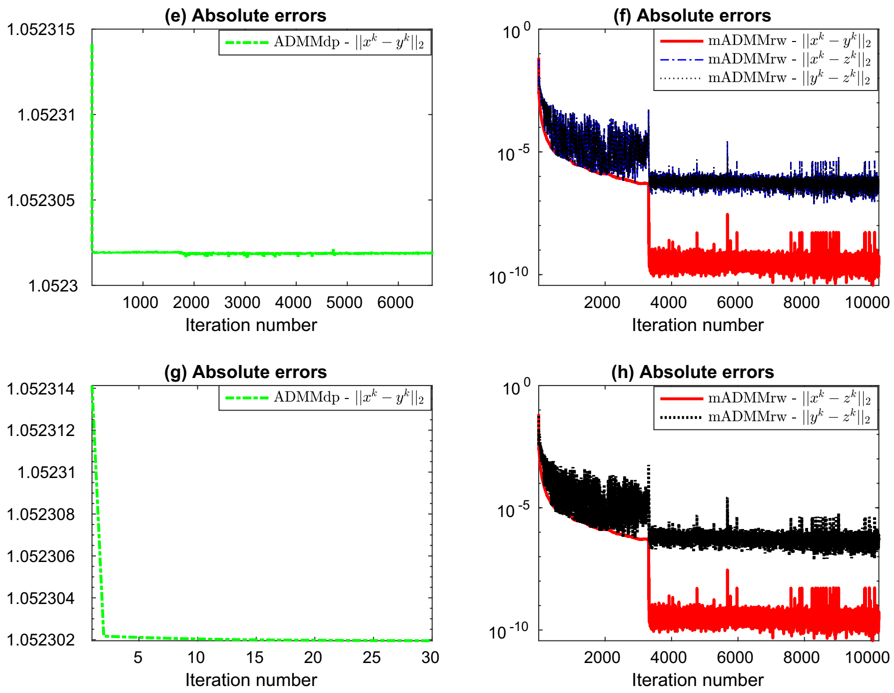

Fig. 5 Plots of residual errors and objective functions values for demonstrating the convergence and behaviors of the proposed `ADMMdp` and `mADMMrw` algorithm about variable $\mathbf{y}$ with the parameters $m = 64$, $n = 1024$ and $s = 12$, $F = 20$

Moreover, it also verifies a weaker sufficient condition for the exact ℓ_1 recovery was proved by Zhang [40]. It is stated that if a vector $\mathbf{x}^*$ satisfies $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ and

$$2\sqrt{\|\mathbf{x}^*\|_0} < \mathcal{K}(\mathbf{y}^*) \tag{25}$$

then $\mathbf{x}^*$ is the unique solution of (1). From Fig. 4 and (24), we have

$$2\sqrt{\|\mathbf{x}^*\|_0} \approx 6.9282 < \mathcal{K}(\mathbf{y}^*) \approx 9.3075$$

and

$$2\sqrt{\|\mathbf{x}^*\|_0} \approx 6.9282 < \mathcal{K}(\mathbf{y}^*) \approx 10.3890$$

for different value $F = 10$ and $F = 20$, respectively. Conversely, from Fig. 4 and $\mathbf{x}^{k+1}$ satisfying $\mathbf{A}\mathbf{x}^{k+1} = \mathbf{b}$, it follows that $\mathbf{x}^{k+1}$ is a unique sparse solution to the ℓ_0 minimization (1) and the ℓ_1/ℓ_2 model (3).

The codes for figures 1 – 4 can be downloaded in Github. Once the download is complete, it is only necessary to change the parameters in `CODE.m` to obtain the behaviors of `ADMMdp` and `mADMMrw` with different coherent factor F and sparsity s under the different row number m and column number n of the sensing matrix

$\mathbf{A} \in \mathbb{R}^{m\times n}$ you wanted before you should install `CVX` [39] with the default solver `Mosek`.

4 Conclusions

In this paper, we transform the ratio ℓ_1 and ℓ_2 norms into an equivalent objective function of separable variables by using the general solution of the inhomogeneous system of linear equations. Then we construct the corresponding ADMM algorithm with dynamic configurable parameters via the corresponding augmentation Lagrangian function (7). Further, we set the positive parameter ρ controlled by (17) in each iteration of `ADMMdp` to ensure that its $\mathbf{y}$-subproblem has a unique global minimum point. Finally, we also verified the convergence and effectiveness of the proposed `ADMMdp` in the section of numerical experiments. In addition, in order to improve the rate of convergence of `ADMMdp`, it is worthwhile to construct a fast method for solving the smoothed basis pursuit problem in the future.

Acknowledgements We are very grateful to anonymous reviewers for their constructive suggestions which improved this paper significantly. The research of Jun Wang has been supported by Scientific Startup Foundation for Doctors of Jiangsu University of Science and Technology (CN) (No. 1052931903). The work of Qiang Ma has been supported by Scientific Startup Foundation for Doctors of Jiangsu University of Science and Technology (CN) (No. 1062931902), the National Natural Science Foundation of China (NSFC) under Grants (No. 52105350) and State Laboratory of Advanced Welding and Joining in HIT (CN) (AWJ-23R01).

Data Availability Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. Berg, E.V.D., Friedlander, M.P.: Sparse optimization with least-squares constraints. SIAM J. Opt. **21**(4), 1201–1229 (2011)
2. Wright, J., Yang, A.Y., Arvind, G., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 210–227 (2008)
3. Bruckstein, A.M., Donoho, D.L., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev. **51**, 34–81 (2009)
4. Yu, Siwei, Ma, Jianwei: Deep learning for geophysics: current and future trends. Rev. Geophys. **59**(3), e2021RG000742 (2021)
5. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**, 489–509 (2006)
6. Candés, E.J., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. **59**, 1207–1233 (2006)
7. Candès, E.J., Tao, T.: Near-optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans. Inf. Theory **52**, 5406–5425 (2006)
8. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
9. Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S.: Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Proc. Asilomar Conf. Signals, Syst. Comput. pp. 40–44 (1993)

10. Needell, D., Tropp, J.A.: CoSaMP: iterative signal recovery from incomplete and inaccurate samples. App. Comput. Harmonic Anal. **26**(3), 301–321 (2009)
11. Eldar, Y.C., Kutyniok, G.: Compressed Sensing: Theory and Applications. Cambridge University Press, Cambridge (2015)
12. Foucart, S., Rauhut, H.: A Mathematical Introduction to Compressive Sensing. Birkhauser, Cambridge MA, USA (2013)
13. Boche, H., Calderbank, R., Kutyniok, G. , Vybiral, J.: Compressed sensing and its applications. Appl. Numer. Harmon. Anal. (2015)
14. Natarajan, B.K.: Sparse approximate solutions to linear systems. SIAM J. Comput. **24**, 227–234 (1995)
15. Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic decomposition by basis pursuit. SIAM J. Sci. Comput. **43**(1), 129–159 (2006)
16. Candès, E. J., Tao, T.: Decoding by linear programming. IEEE Trans. Inform. Theory **51** (2005)
17. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imag. Sci. **2**, 183–202 (2009)
18. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. **3**, 1–122 (2011)
19. Lai, M., Wang, J.: An unconstrained ℓ_q minimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems. SIAM J. Opt. **21**, 82–101 (2010)
20. Yin, P., Lou, Y., He, Q., Xin, J.: Minimization of ℓ_{1-2} for compressed sensing. SIAM J. Sci. Comput. **37**, A536–A563 (2015)
21. Lou, Y., Yin, P., He, Q., Xin, J.: Computing sparse representation in a highly coherent dictionary based on difference of ℓ_1 and ℓ_2. J. Sci. Comput. **64**, 178–196 (2015)
22. Zhang, S., Xin, J.: Minimization of transformed L_1 penalty: theory, difference of convex function algorithm, and robust application in compressed sensing. Math. Program. **169**(1), 307–336 (2018)
23. Shen, X., Pan, W., Zhu, Y.: Likelihood-based selection and sharp parameter estimation. J. Am. Stat. Assoc. **107**(497), 223–232 (2012)
24. Wang, J.: Sparse reconstruction via the mixture optimization model with iterative support estimate. Inf. Sci. **574**, 1–11 (2021)
25. Wang, J.: The proximal gradient methods for the $\ell_{1-\infty}$ minimization problem with the sharp estimate, submitted (2022)
26. Wang, J.: A wonderful triangle in compressed sensing. Inf. Sci. **611**, 95–106 (2022)
27. Rahimi, Y., Wang, C., Dong, H., Lou, Y.: A scale invariant approach for sparse signal recovery. SIAM J. Sci. Comput. **41**(6), A3649–A3672 (2019)
28. Hoyer, P. O. : Non-negative sparse coding. In: Proceedings of the IEEE Workshop on Neural Networks for Signal, Martigny, Switzerland, pp. 557–565 (2002)
29. Hurley, N., Rickard, S.: Comparing measures of sparsity. IEEE Trans. Inform. Theory **55**, 4723–4741 (2009)
30. Zeng, L., Yu, P., Pong, T.K.: Analysis and algorithms for some compressed sensing models based on L1/L2 minimization. SIAM J. Opt. **31**(2), 1576–1603 (2021)
31. Xu, Y., Narayan, A., Tran, H., Webster, C.G.: Analysis of the ratio of ℓ_1 and ℓ_2 norms in compressed sensing. Appl. Comput. Harmon. Anal. **55**, 486–511 (2021)
32. Tao, Min: Minimization of L1 over L2 for sparse signal recovery with convergence guarantee. SIAM J. Sci. Comput. **44**(2), A770–A797 (2022)
33. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: Proc. IEEE comput. Vis. Pattern recongit., pp. 233–240 (2011)
34. Wang, C., J. G, Min Tao, Lou, Nagy, Y.: Limited-angle CT reconstruction via the L1/L2 minimization. SIAM M. Imag. Sci. **14**(2), 749–777 (2021)
35. Coleman, T.F., Li, Y.: On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds. Math. Program. **67**(2), 189–224 (1994)
36. Mordukhovich, B., Nam, N.M.: An easy path to convex analysis and applications. Synth. Lect. Math. Stat. **6**(2), 1–218 (2016)
37. Powell, M.J.D.: A Fortran Subroutine for Solving Systems of Nonlinear Algebraic Equations, Numerical Methods for Nonlinear Algebraic Equations, P. Rabinowitz, (ed.), Ch.7 (1970)

38. Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters. SIAM J. Appl. Math. **11**, 431–441 (1963)
39. Grant, M.: Stephen Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 2.2. http://cvxr.com/cvx, January (2020)
40. Zhang, Y.: Theory of compressive sensing via L1-minimization : a non-RIP analysis and extensions. J. Oper. Res. Soc. China **1**, 79–105 (2013)