Front. Comput. Sci.

LETTER

Supporting Document

Xiaohu LUO, Zili ZHANG(✉)

Faculty of Computer and Information Science, Southwest University, Chongqing 400715, China

Due to the limited space, we cannot discuss more in the original paper. To improve the completeness of the paper, we consider providing some supporting materials in this supporting document. To be specific, we first present the main proof of three theorems, and then consider conducting some experiments to further verify the effectiveness of the proposed model.

1 Proofs of three Theorems

Before moving on, we need the following Lemma.

Lemma 1. Let $S = supp(\hat{\mathbf{x}})$ and denote $\mathbf{h} = \hat{\mathbf{x}} - \mathbf{x}^\sharp$, where $\mathbf{x}^\sharp$ is the solution of

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

Then we have

$$\|\mathbf{h}_{S^c}\|_q^q \le w^q \|\mathbf{h}_S\|_q^q, \tag{2}$$

where $\mathbf{h}_S \in \mathbb{R}^S$ denotes the restriction of the vector $\mathbf{h} \in \mathbb{R}^n$ onto a subset of coordinates $S \subset \{1,\dots,n\}$.

Received month dd, yyyy; accepted month dd, yyyy

E-mail: zhangzl@swu.edu.cn

Proof. Since $\mathbf{x}^\sharp$ is the minimizer of (1), we have $\|\mathbf{w}\odot\mathbf{x}^\sharp\|_q \le \|\mathbf{w}\odot\hat{\mathbf{x}}\|_q$. By the triangle inequality, it yields that

$$\begin{aligned}\left\|\mathbf{w}\odot\mathbf{x}^\sharp\right\|_q^q &= \|\mathbf{w}\odot(\hat{\mathbf{x}}+\mathbf{h})\|_q^q \\ &= \|\mathbf{w}\odot(\hat{\mathbf{x}}_S+\mathbf{h}_S)\|_q^q + \|\hat{\mathbf{x}}_{S^c}+\mathbf{h}_{S^c}\|_q^q \\ &\ge \|\mathbf{w}\odot\hat{\mathbf{x}}\|_q^q - w^q\|\mathbf{h}_S\|_q^q + \|\mathbf{h}_{S^c}\|_q^q \\ &\ge \|\mathbf{w}\odot\hat{\mathbf{x}}\|_q^q - w^q\|\mathbf{h}_S\|_q^q + \|\mathbf{h}_{S^c}\|_q^q,\end{aligned}$$

using $\hat{\mathbf{x}}_S = \hat{\mathbf{x}}$, and $\hat{\mathbf{x}}_{S^c} = 0$. This completes the proof. □

1.1 Proof of Theorem 1

We would like to show that the recovery error $\mathbf{h}$ is zero. To do this, we decompose $\mathbf{h}$.

Step 1: Decomposing the support. Let T_0 be the support of $\hat{\mathbf{x}}$, T_1 index the ak largest coefficients of $\mathbf{h}_{T^c}$ in magnitude, T_2 index the next ak largest coefficients of $\mathbf{h}_{T^c}$, and so on. Denote $T_{01} = T_0 \cup T_1$. We decompose $\mathbf{A}\mathbf{h}$ as

$$0 = \|\mathbf{A}\mathbf{h}\|_q^q \ge \left\|A\mathbf{h}_{T_{01}}\right\|_q^q - \left\|A\mathbf{h}_{T_{01}^c}\right\|_q^q. \tag{3}$$

Next, we will examine the terms $\|A\mathbf{h}_{T_{01}}\|_q^q$ and $\|A\mathbf{h}_{T_{01}^c}\|_q^q$ one by one.

Step 2: Applying q-RIP. Since $\left|T_{0,1}\right| \leq k + ak$, applying q-RIP yields

$$\left\|A\mathbf{h}_{T_{01}}\right\|_q^q \geq (1 - \delta_{k+ak})\|\mathbf{h}_{T_0}\|$$

and the triangle inequality followed by q-RIP also gives

$$\begin{aligned}\left\|A\mathbf{h}_{T_{01}^c}\right\|_q^q &\leq \left\|\sum_{j\geq 2} \mathbf{A}\mathbf{h}_{T_j}\right\|_q^q \leq \sum_{j\geq 2} \left\|\mathbf{A}\mathbf{h}_{T_j}\right\|_q^q \\ &\leq (1+\delta_{ak}) \sum_{j\geq 2} \left\|\mathbf{h}_{T_j}\right\|_2^q .\end{aligned}$$

Now we need to control the size of $\left\|\mathbf{h}_{T_j}\right\|_2$. We aim to Applying Lemma 1 to estimate the ℓ_2 norm in terms of the ℓ_q norm. For each $t \in T_j$ and $s \in T_{j-1}$, $|h_t|^q \leq |h_s|^q$, so that $\left|(\mathbf{h}_{T_j})_t\right|^q \leq \frac{\left\|\mathbf{h}_{T_{j-1}}\right\|_q^q}{ak}$. Then $|(\mathbf{h}_{T_j})_t|^2 \leq \left\|\mathbf{h}_{T_{j-1}}\right\|_q^2 / L^{2/q}$, and thus

$$\left\|\mathbf{h}_{T_j}\right\|_2^q \leq (ak)^{q/2-1} \left\|\mathbf{h}_{T_{j-1}}\right\|_q^q .$$

Furthermore we can get

$$\begin{aligned}\sum_{j\geq 2} \left\|\mathbf{h}_{T_j}\right\|_2^q &\leq (ak)^{q/2-1} \sum_{j\geq 1} \left\|\mathbf{h}_{T_j}\right\|_q^q \\ &= L^{q/2-1} \left\|\mathbf{h}_{T_0^c}\right\|_q^q ,\end{aligned} \tag{4}$$

then convert back from ℓ_q to ℓ_2 by means of Hölder's inequality:

$$\begin{aligned}\left\|\mathbf{h}_{T_0}\right\|_q^q &= \sum_{t\in T_0} |h_t|^q \cdot 1 \leq \left(\sum_{T_0} |h_t|^2\right)^{\frac{q}{2}} \left(\sum_{T_0} 1\right)^{1-\frac{q}{2}} \\ &= \left\|\mathbf{h}_{T_0}\right\|_2^q k^{1-q/2}.\end{aligned}$$

Using Lemma 1 again, we obtain

$$\begin{aligned}\sum_{j\geq 2} \left\|\mathbf{h}_{T_j}\right\|_2^q &\leq (ak)^{q/2-1} \left\|\mathbf{h}_{T_0}\right\|_q^q w^q = w^q \left\|\mathbf{h}_{T_0}\right\|_2^q / a^{1-p/2} \\ &= w^q \left\|\mathbf{h}_{T_0}\right\|_2^q / b.\end{aligned} \tag{5}$$

Step 3: Summing up. By (3), (4) and (5), we have

$$\begin{aligned}0 &\geq (1-\delta_{ak+k}) \left\|\mathbf{h}_{T_{01}}\right\|_2^q - (1+\delta_{ak})\, w^q \left\|\mathbf{h}_{T_0}\right\|_2^q / b \\ &\geq (1-\delta_{ak+k} - (1+\delta_{ak})\, w^q / b) \left\|\mathbf{h}_{T_{01}}\right\|_2^q .\end{aligned} \tag{6}$$

Condition of Theorem 1 ensures that the scalar factor is positive, so $\mathbf{h}_{T_{01}} = 0$. In particular, $\mathbf{h}_{T_0} = 0$, then $\mathbf{h} = 0$ follows from (2).

1.2 Proof of Theorem 2

Let $\mathbf{r}_1 = \left(\mathbf{z}_1^{\mathrm{T}}, \mathbf{y}_1^{\mathrm{T}}\right)^{\mathrm{T}}, \mathbf{r}_2 = \left(\mathbf{z}_2^{\mathrm{T}}, \mathbf{y}_2^{\mathrm{T}}\right)^{\mathrm{T}} \in \mathbb{R}^{4n}$, then

$$\begin{aligned}\|Q(\mathbf{r}_1) - Q(\mathbf{r}_2)\|_2 &\leq (1+\lambda)\|\mathbf{y}_1 - \mathbf{y}_2\|_2 + \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\quad + \left\|[\mathbf{z}_1 - \nabla f(\mathbf{z}_1)]^+ - [\mathbf{z}_2 - \nabla f(\mathbf{z}_2)]^+\right\|_2 \\ &\leq (1+\lambda)\|\mathbf{y}_1 - \mathbf{y}_2\|_2 + \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\quad + \|\mathbf{z}_1 - \nabla f(\mathbf{z}_1) - [\mathbf{z}_2 - \nabla f(\mathbf{z}_2)]\|_2 \\ &\leq (1+\lambda)\|\mathbf{y}_1 - \mathbf{y}_2\|_2 + 2\|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\quad + L\|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\leq (3+\lambda+L)\|\mathbf{r}_1 - \mathbf{r}_2\|_2\end{aligned}$$

The third inequality is due to

$$\begin{aligned}&\|\mathbf{z}_1 - \nabla f(\mathbf{z}_1) - [\mathbf{z}_2 - \nabla f(\mathbf{z}_2)]\|_2 \\ &\leq \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + \|\mathbf{B}\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\quad + \tau \frac{\|(\mathbf{W})^q\|_2 \left\|(\mathbf{z}_1)^{q-1}\right\|_2}{\|\mathbf{W}\mathbf{z}_1\|_q^{q-1}} + \tau \frac{\|(\mathbf{W})^q\|_2 \left\|(\mathbf{z}_2)^{q-1}\right\|_2}{\|\mathbf{W}\mathbf{z}_2\|_q^{q-1}} \\ &\leq \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + \|\mathbf{B}\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + 2\tau H_1 \\ &\leq \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + \|\mathbf{B}\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + H_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\ &\leq \|\mathbf{z}_1 - \mathbf{z}_2\|_2 + L\|\mathbf{z}_1 - \mathbf{z}_2\|_2 ,\end{aligned}$$

where

$$\begin{aligned}H_1 &= \max \|(\mathbf{W})^q\|_2 \cdot \left\{\tau \frac{\left\|(\mathbf{z}_1)^{q-1}\right\|_2}{\|\mathbf{W}\mathbf{z}_1\|_q^{q-1}}, \tau \frac{\left\|(\mathbf{z}_2)^{q-1}\right\|_2}{\|\mathbf{W}\mathbf{z}_2\|_q^{q-1}}\right\} \\ &< \infty.\end{aligned}$$

It is easy to see that there exists $H_2 > 0$, such that $2\tau H_1 \leq H_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2$, and $L = 2\max\{\|\mathbf{B}\|_2, H_2\}$.

Hence $Q(\mathbf{r})$ is Lipschitz continuous on $\mathbb{R}^{4n}$ with Lipschitz constant $(3+\lambda+L)$. There exists a unique solution $\mathbf{r}(t)$ with initial condition $\mathbf{r}_0$ by the local existence theorem of the solution to ordinary differential equations. This completes the proof.

1.3 Proof of Theorem 3

If we take $\mathbf{z}^* \in S$ as the optimal solution of the variational inequality

$$p'(\alpha) = \nabla f\left(\mathbf{z}^*\right)^{\mathrm{T}}\left(\mathbf{z}-\mathbf{z}^*\right) \geq 0, \quad \forall \mathbf{z} \in S, \tag{7}$$

$\mathbf{z}(t)$, $t \in [0,+\infty)$ as the solution sequence of the proposed IPNN model, and consider the Lyapunov function

$$V(t) = \frac{1}{2}\left\|\mathbf{z}(t)-\mathbf{z}^*\right\|_2^2 = \frac{1}{2}\left(\mathbf{z}(t)-\mathbf{z}^*\right)^{\mathrm{T}}\left(\mathbf{z}(t)-\mathbf{z}^*\right)$$

then the following equality holds

$$\ddot{V}(t)+\lambda\dot{V}(t) = \|\mathbf{y}(t)\|_2^2 + \left\langle \mathbf{z}(t)-\mathbf{z}^*, \dot{\mathbf{y}}(t)+\lambda\mathbf{y}(t)\right\rangle. \tag{8}$$

Combining

$$\begin{cases} \dot{\mathbf{z}} = \mathbf{y}, \\ \dot{\mathbf{y}} = -\lambda\mathbf{y} + \left[\mathbf{z}-\mathbf{B}\mathbf{z}-\mathbf{c}-\tau\frac{(\mathbf{W})^q(\mathbf{z})^{q-1}}{\|\mathbf{W}\mathbf{z}\|_q^{q-1}}\right]^+ - \mathbf{z}. \end{cases}, \tag{9}$$

and the definition of operator $\mathbb{N}(\mathbf{z})$, it is not difficult to obtain

$$\dot{\mathbf{y}}(t)+\lambda\mathbf{y}(t) = \{\mathbf{z}(t)-\nabla f[\mathbf{z}(t)]\}^+ - \mathbf{z}(t) = -\mathbb{N}[\mathbf{z}(t)]. \tag{10}$$

From (8) and (10), we can get

$$\|\mathbf{y}(t)\|_2^2 = \ddot{V}(t)+\lambda\dot{V}(t)+\left\langle \mathbf{z}(t)-z^*, \mathbb{N}[\mathbf{z}(t)]\right\rangle.$$

Since $\mathbf{z}^*$ is the optimal solution of the variational inequality given as follows.

$$\mathbf{x}^* \in \Omega, \left\langle F(\mathbf{x}^*), \mathbf{x}-\mathbf{x}^*\right\rangle \geq 0, \text{ for all } \mathbf{x} \in \Omega, \tag{11}$$

where Ω is the nonempty closed convex subset of $\mathbb{R}^n$, F is a map from $\mathbb{R}^n$ to itself, $\langle\cdot,\cdot\rangle$ denotes the inner product in R^n, $\nabla f(\mathbf{z}^*) = 0$, and $\mathbb{N}(\mathbf{z}^*) = 0$, then

$$\|\mathbf{y}(t)\|_2^2 = \ddot{V}(t)+\lambda\dot{V}(t)+\left\langle \mathbf{z}(t)-\mathbf{z}^*, \mathbb{N}[\mathbf{z}(t)]-\mathbb{N}(\mathbf{z}^*)\right\rangle.$$

From the condition given in Theorem 3, we obtain

$$\|\mathbf{y}(t)\|_2^2 \geq \ddot{V}(t)+\lambda\dot{V}(t)+\frac{D}{2+L}\|\mathbb{N}[\mathbf{z}(t)]\|_2^2. \tag{12}$$

Combining (10) with (12) gives

$$\|\mathbf{y}(t)\|_2^2 \geq \ddot{V}(t)+\lambda\dot{V}(t)+\frac{D}{2+L}\|\dot{\mathbf{y}}(t)+\lambda\mathbf{y}(t)\|_2^2. \tag{13}$$

Expanding the right side of inequality (13), we have

$$\begin{aligned} u(t) = \ddot{V}(t)+\lambda\dot{V}(t)+\frac{D}{2+L}\|\dot{\mathbf{y}}\|_2^2 + \frac{D}{2+L}\lambda\frac{d\left(\|\mathbf{y}\|_2^2\right)}{dt} \\ +\left(\frac{D}{2+L}\lambda^2-1\right)\|\mathbf{y}\|_2^2 \leq 0. \end{aligned}$$

It is shown that the primitive function

$$\begin{aligned} U(t) = \dot{V}(t)+\lambda V(t)+\frac{D}{2+L}\lambda\|\mathbf{y}\|_2^2 \\ +\int_0^t \frac{D}{2+L}\|\dot{\mathbf{y}}(s)\|_2^2 ds \\ +\left(\frac{D}{2+L}\lambda^2-1\right)\int_0^t \|\mathbf{y}(s)\|_2^2 ds \end{aligned}$$

that $u(t)$ is monotone nonincreasing. Therefore, for any $t>0$, $U(t) \leq U(0)$, where

$$\begin{aligned} U(0) &= \dot{V}(0)+\lambda V(0)+\frac{D}{2+L}\lambda\|\mathbf{y}(0)\|_2^2 \\ &= \|\mathbf{z}(0)-\mathbf{z}^*\|_2\,\mathbf{y}(0)+\frac{\lambda}{2}\|\mathbf{z}(0)-\mathbf{z}^*\|_2^2 \\ &+\frac{D}{2+L}\lambda\|\mathbf{y}(0)\|_2^2. \end{aligned} \tag{14}$$

From the condition given in Theorem 3, we obtain

$$\left(\frac{D}{2+L}\lambda^2-1\right)\int_0^t \|y(s)\|_2^2 ds \geq 0,$$

Therefore,

$$\dot{V}(t) + \lambda V(t) \leq U(0). \tag{15}$$

Multiplying $e^{\lambda t}$ on both sides of inequality (15), we obtain

$$\dot{V}(t)e^{\lambda t} + \lambda V(t)e^{\lambda t} \leq U(0)e^{\lambda t},$$

which implies that

$$\frac{d}{dt}\left[V(t)e^{\lambda t}\right] = \dot{V}(t)e^{\lambda t} + \lambda V(t)e^{\lambda t} \leq U(0)e^{\lambda t}$$

By integrating the above inequalitiy from 0 to t, we can get $V(t) \leq V(0)\, e^{-\lambda t} + \frac{1}{\lambda}U(0)$. Therefore, $V(t)$ is bounded, and the trajectory $\mathbf{z}(t)$ of the neural network model (9) is bounded as well. Next, we prove that $\lim_{t\to\infty} \dot{\mathbf{y}}(t) = 0$ and $\lim_{t\to\infty} \mathbf{y}(t) = 0$. Since $V(t)$ is bounded and combine with inequality (14), we can get

$$\dot{V}(t) + \frac{D}{2+L}\lambda\|\mathbf{y}(t)\|_2^2 \leq U(0).$$

Then we have

$$\langle \mathbf{z}(t) - \mathbf{z}^*, y(t)\rangle + \frac{D}{2+L}\lambda\|\mathbf{y}(t)\|_2^2 \leq U(0). \tag{16}$$

From the boundedness of $\|\mathbf{z}(t) - \mathbf{z}^*\|_2$ and (16), we can deduce that $\|\mathbf{y}(t)\|_2$ is bounded. By (14) and (16), we can obtain

$$\int_0^{\infty} \|\dot{\mathbf{y}}(t)\|_2^2 dt < +\infty, \int_0^{\infty} \|\mathbf{y}(t)\|_2^2 dt < +\infty.$$

Suppose $\int_0^{+\infty} \|y(t)\|_2^2 dt = \mathbf{A} < +\infty$. Since

$$\int_0^{+\infty} \|\dot{y}(s)\|_2^2 ds < \infty,$$

it implies that there exists $M > 0$, such that $\|\dot{y}(t)\|_2 < M$ for any $t \in (0, +\infty)$. Therefore,

$$\begin{aligned}\int_0^{+\infty} \|y(t)\|_2^2\|\dot{y}(t)\|_2 ds &= \frac{1}{3}\lim_{t\to+\infty}\left[\|y(t)\|_2^3 - \|y(0)\|_2^3\right] \\ &\leq M\int_0^{+\infty} \|y(s)\|_2^2 ds < +\infty.\end{aligned}$$

According to the theory of calculus, it can be concluded that the value of $\lim_{t\to+\infty} \|y(t)\|_2^3$ exists. Therefore, the value of $\lim_{t\to+\infty} \|y(t)\|_2$ exists. Since $\int_0^{+\infty} \|y(t)\|_2^2 dt < +\infty$, we obtain $\lim_{t\to+\infty} \|y(t)\|_2 = 0$. Then according to the boundedness of $V(t)$, one can easily conclude that $\lim_{t\to+\infty} y(t) = 0$.

Next,we prove $\lim_{t\to\infty} \dot{\mathbf{y}}(t) = 0$.

Define $\mathbf{e}_\varepsilon(t) = \frac{1}{\varepsilon}[\mathbf{y}(t+\varepsilon) - \mathbf{y}(t)]$ and bring $\mathbf{e}_\varepsilon(t)$ into (10), then

$$\dot{\mathbf{e}}_\varepsilon(t) + \mathbf{e}_\varepsilon(t) = -\frac{1}{\varepsilon}\{\mathbb{N}[\mathbf{z}(t+\varepsilon)] - \mathbb{N}[\mathbf{z}(t)]\}. \tag{17}$$

Due to the condition in Theorem 3

$$\begin{aligned}&\{\mathbb{N}[\mathbf{z}(t+\varepsilon)] - \mathbb{N}[\mathbf{z}(t)]\}^{\mathrm{T}}[\mathbf{z}(t+\varepsilon) - \mathbf{z}(t)] \\ &\geq \frac{D}{(2+L)}\|\mathbb{N}[\mathbf{z}(t+\varepsilon)] - \mathbb{N}[\mathbf{z}(t)]\|_2^2,\end{aligned}$$

then

$$\begin{aligned}&-\frac{\varepsilon(2+L)}{D}\left[\dot{\mathbf{e}}_\varepsilon(t) + \mathbf{e}_\varepsilon(t)\right]^{\mathrm{T}}\left[\mathbf{z}(t+\varepsilon) - \mathbf{z}(t)\right] \\ &\geq \|\mathbb{N}[\mathbf{z}(t+\varepsilon)] - \mathbb{N}[\mathbf{z}(t)]\|_2^2.\end{aligned}$$

Then we can get

$$\frac{\varepsilon(2+L)}{D}\sup_{s\in[t,+\infty)}\|\mathbf{y}(s)\|_2 \geq \|\mathbb{N}[\mathbf{z}(t+\varepsilon)] - \mathbb{N}[\mathbf{z}(t)]\|_2.$$

Integrating (17), we can get $\lim_{t\to\infty} \sup \|\mathbf{e}_\varepsilon(t)\|_2 = 0$. Since $\|\dot{\mathbf{y}}(t)\|_2 \leq \sup\|\mathbf{e}_\varepsilon(t)\|_2$, it follows that

$$\lim_{t\to\infty} \dot{\mathbf{y}}(t) = 0.$$

From (9), we can conclude

$$\lim_{t\to\infty}\left(\dot{\mathbf{y}} + \lambda\mathbf{y} + \mathbf{z} - \left[\mathbf{z} - \mathbf{B}\mathbf{z} - \mathbf{c} - \tau\frac{(\mathbf{W})^q(\mathbf{z})^{q-1}}{\|\mathbf{W}\mathbf{z}\|_q^{q-1}}\right]^+\right) = 0.$$

So, we can prove that

$$\begin{aligned}&\lim_{t\to\infty}\left(\mathbf{z} - \left[\mathbf{z} - \mathbf{B}\mathbf{z} - \mathbf{c} - \tau\frac{(\mathbf{W})^q(\mathbf{z})^{q-1}}{\|\mathbf{W}\mathbf{z}\|_q^{q-1}}\right]^+\right) \\ &= \lim_{t\to\infty}\{\mathbb{N}[\mathbf{z}(t)]\} = 0 = \mathbb{N}(\mathbf{z}^*),\end{aligned}$$

which implies that $\lim_{t\to+\infty} \mathbf{z}(t) = \mathbf{z}^*$. In other words, the solution of model (9) converges to the optimal solution of

$$p'(\alpha) = \nabla f(\mathbf{z}^*)^{\mathrm{T}}(\mathbf{z} - \mathbf{z}^*) \geq 0, \quad \forall \mathbf{z} \in S. \tag{18}$$

This certifies the proof.

2 Numerical experiments

In this section, several concrete examples will be given to further demonstrate the performance of the IPNN algorithm when it is used to solve the weighted L_q minimization problem with prior information.

The whole experiment process is mainly divided into the following five steps:

1. Generate the desired k-sparse signal $\mathbf{x}$ randomly with length n and sparsity k. The indices of the non-zero entries are generated at random and there non-zero entries are generated independently from a standard normal distribution.
2. Generate the desire sensing matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ whose elements are generated independently from a standard normal distribution.
3. Compute the measurement signal $\mathbf{b}$ by $\mathbf{Ax}$.
4. Set the prior information (i.e., the correct partially known support set T, $|T| = s$), and the initial value $\mathbf{x}_0$.
5. Recover the original sparse signal by the IPNN algorithm.

We define the relative recovery error as

$$\mathrm{RE} = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2},$$

where$\hat{\mathbf{x}}$ is denoted by the recovered signal, and we also use the Matlab command *ode45* to solve the ordinary differential equation..

2.1 Parameter selection

In this experiment, we first consider the problem of selecting the regularization parameter τ. We set $n = 64, m = 32$, $s = 0$, and the sparsity $k = 4, 5, 6, 7, 8$. Then a test procedure is carried out to determine the effect of τ on the relative error for $q = 0.3, 0.5, 0.7, 1$. In Fig. 1 we plot the obtained results under different q's. It first shows that the regularization parameter τ has a great influence on the signal recovery error, and the best choice of τ mainly depends on the value of q. For a fixed q, when we recover the signals with different sparsity, the value of τ that minimizes the recovery error always appears in a fixed range. Specifically, for $q = 0.3$, the optimal range of τ is $0.0001 \sim 0.001$. In other words, when the regular parameter τ takes the middle value, the relative error is the smallest. Similarly, we can get the selection of τ values for other values of q. Therefore, among all the following noise-free experiments, for different types of signals, we choose the appropriate τ values according to the values of q.

2.2 Weight selection

In the analysis process of the algorithm, it can be found that the recovery effectiveness of the algorithm has a certain dependence on the selection of the weight value. Therefore, the selection of the weight value is very important. Next, we will set up different weight experiments to compare, and find the appropriate weight. In the experiment, we set $n = 256, m = 128, k = 20, s = 20, \lambda = 10$, and different q values with suitable parameter value τ. When q is chosen from 0.3, 0.5, 0.7, 1, then τ will be set as 0.0001, 0.005, 0.05, and 0.1, correspondingly. For different weights, the relationship between relative error and weight obtained via

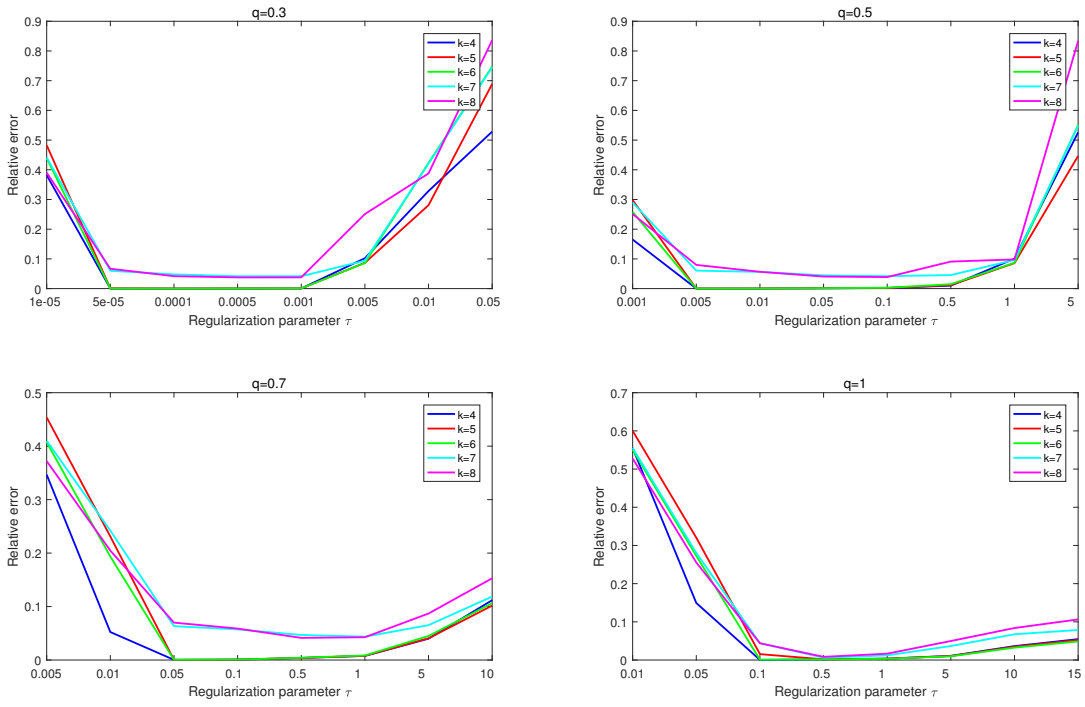


Fig. 1 Relative error with varying regularization parameter τ under the condition of different sparsity k for $q = 0.3, 0.5, 0.7, 1$.

the IPNN algorithm is shown in Fig. 2. It can be

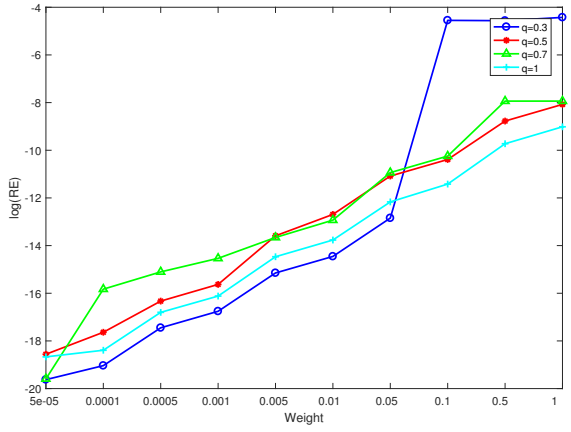


Fig. 2 Relative error variation with different values of weight w for $q = 0.3, 0.5, 0.7, 1$.

seen that when $q = 0.3, 0.5, 0.7, 1$, the relative error increases with an increase in the weight. However, from Fig. 3 and Fig. 4 for extremely small weights (i.e. $w \leq 0.0001$), the relative error oscillates, which leads to instability of the algorithm. Furthermore, the smaller the weight is , the more obvious the oscillation can be found. So, in the next experiments, we uniformly set the weight w to 0.0005.

2.3 Sparse signal recovery

Firstly, we present an experiment of sparse vector recovery without noise. We set $n = 500, m = 150, k = 40, s = 20$, $q = 0.5$, $\tau = 0.005$ and, $\lambda = 10$. Fig. 5 shows the effectiveness of recovery, where the recovered signal completely covers the original signal, which implies that the IPNN algorithm has good performance in solving the weighted L_q minimization problem with prior information. Fig. 6 depicts the relationship between the number of iterations and the recovery relative error, and it can be observed that the relative error tends to

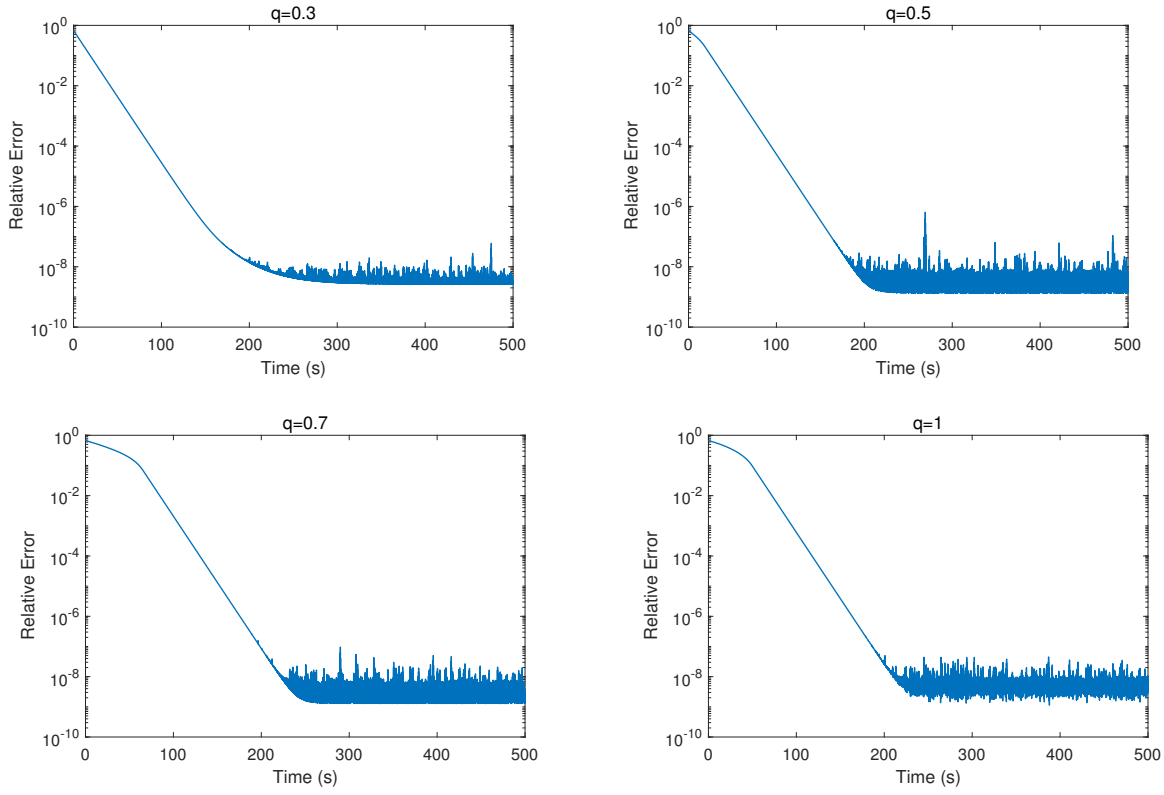


Fig. 3 Relative recovery error variation for $w = 0.00005$ and $q = 0.3, 0.5, 0.7, 1$.

stabilize. The solution of the neural network is shown in Fig. 7, each line represents the trend of each component in the solution vector over time. We can see that each component tends to be stable gradually, which enough to illustrate the solution is asymptotically stable and converges to the local optimal solution. Fig. 6 and Fig. 7 show the real-time performance of the IPNN algorithm. Moreover, for other values of q, the same conclusion can be obtained.

We also consider to compare the relative recovery error (for noiseless case) against the number of measurements. Fixed $n = 500, k = 40, s = 20, \lambda = 10$, for $q = 0.3, 0.5, 0.7, 1$, Fig. 8 shows that the relative error decreases with the increase of the number of measurements.

In order to explore the influence of different amounts of prior information on the recovery error, we fixed a sparse signal and draw a diagram of the recovery relative error variation with the increase of the number of known supports s at $q = 0.3, 0.5, 0.7, 1$, respectively, as shown in Fig.9. As expected, the relative error decreases with the increase of s.

Secondly, we consider the problem of sparse signal recovery with noise and compare with other algorithms. The main algorithms for comparison are the seq-L_q algorithm (a method for solving the L_q problem based on the sequence iterative format, which can be found on the Foucart homepage) [1], and IRLS-Lq (the iterative least squares method for solving the L_q problem) [2]. In addition, we also consider the most classical PD-L_1 solver (from the Matlab software released by Candès) [3], ADMM-L_1 (a classical algorithm for solving sparse problems, the full name is alternating direction method of multipliers) [4], ISTA (the iterative soft thresh-

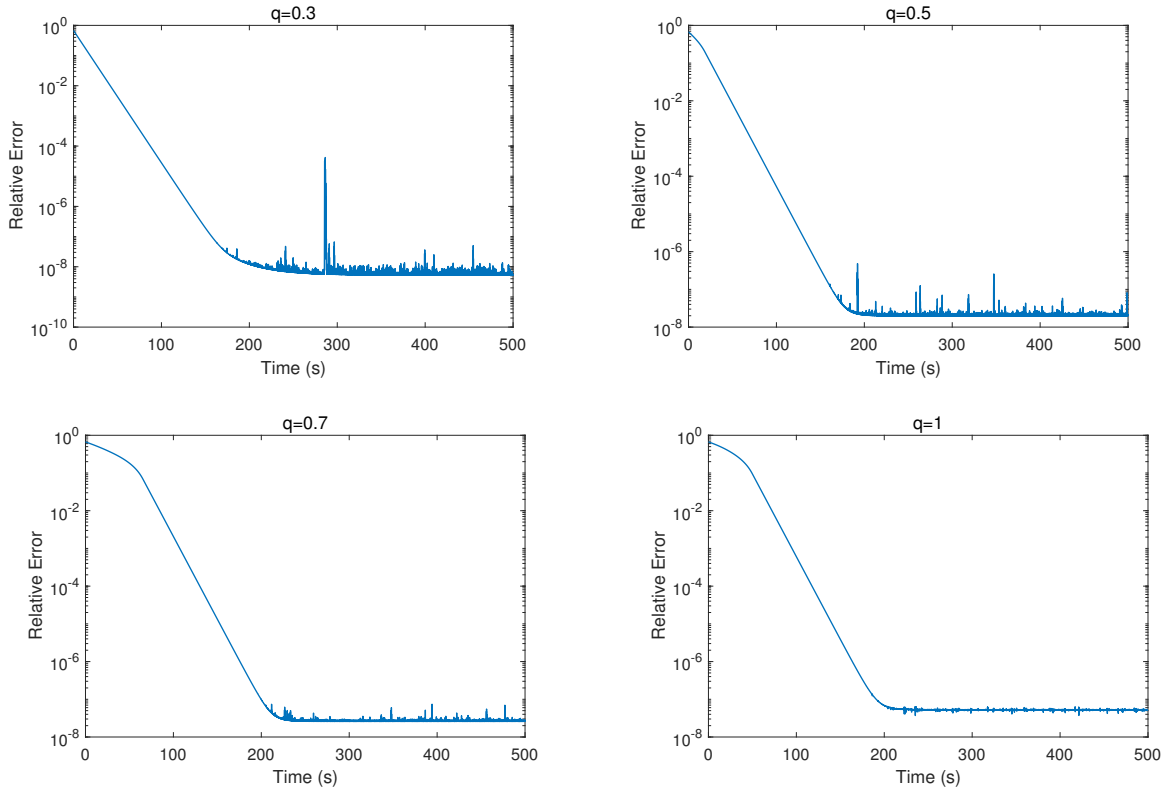


Fig. 4 Relative recovery error variation for $w = 0.0001$ and $q = 0.3, 0.5, 0.7, 1$.

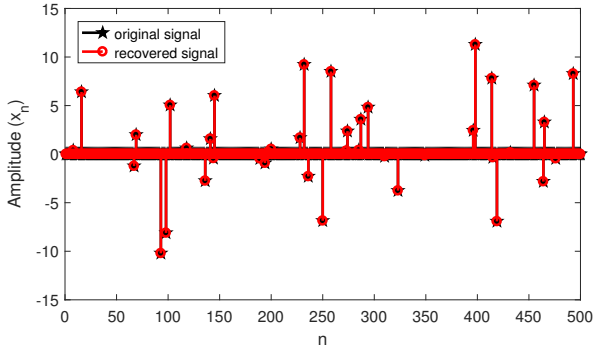


Fig. 5 Original signal and recovered signal via the IPNN algorithm, noiseless case.

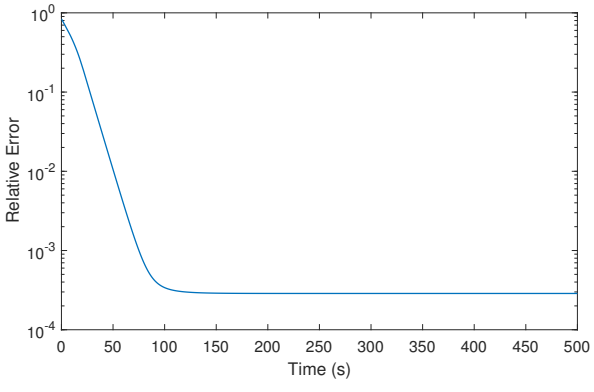


Fig. 6 Stability of relative error of signal via the IPNN algorithm.

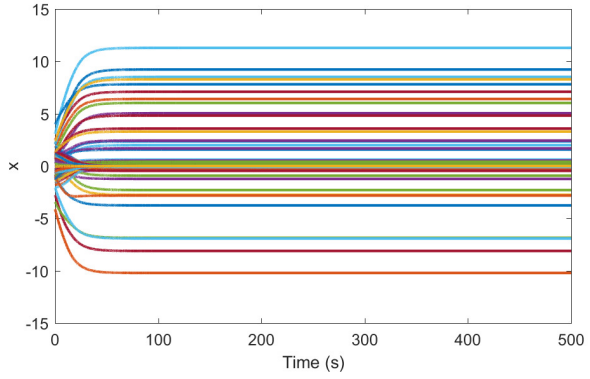


Fig. 7 Stability and convergence of solutions via the IPNN algorithm.

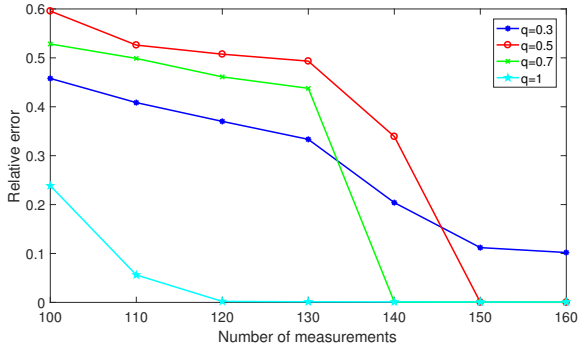


Fig. 8 Relative error with different measurement numbers for $q = 0.3, 0.5, 0.7, 1$.

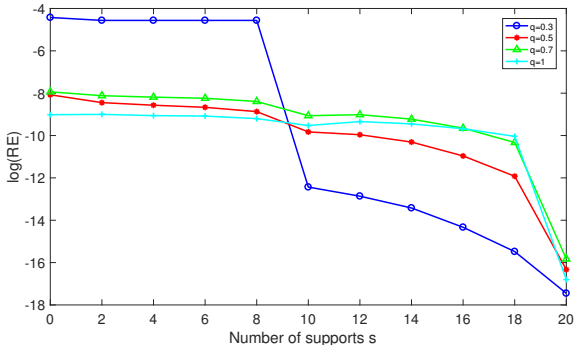


Fig. 9 Relative error with different the number of known supports s for $q = 0.3, 0.5, 0.7, 1$.

old algorithm for solving basis pursuit de-noising) [5], and IPNN-L_{1-2} [6].

In this experiment, Gaussian white noise is introduced into the original data and the signal-to-noise ratio (SNR) of the recovered signal is calculated. In MATLAB, the noise is introduced into the clean data $\mathbf{Ax}$ by $\mathbf{b} = \text{awgn}(\mathbf{A} * \mathbf{x}; \text{snr})$, where snr represents SNR measured in dB. The SNR of the recovered signal is computed by

$$\text{SNR} = 10 \log_{10} \frac{||\mathbf{x}||_2^2}{||\hat{\mathbf{x}} - \mathbf{x}||_2^2}.$$

The higher the SNR value, the better the recovery performance of the algorithm. We set $n = 128, m = 64, k = 10$, and $s = 0, 5, 10$, for $q = 0.3, 0.5, 0.7$. Fig. 10 shows the recovery results of adding white

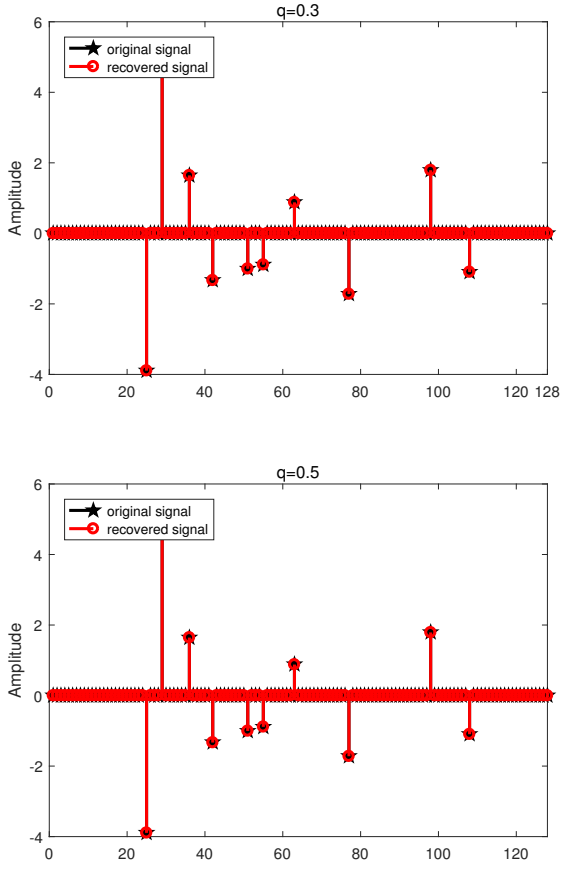


Fig. 10 Original signal and the recovered signal via IPNN algorithm with SNR= 30dB

Gaussian noise with SNR = 30 dB. It can be seen that the IPNN algorithm is superior in solving the optimization problem with noise, which verifies the robustness of the algorithm. Also, with the addition of different levels of noise, 50 random experiments were carried out, and the average SNR of the recovered signals is obtained. As shown in Table 1, the left most column represents the SNR of added measurement noise and the other columns represent the SNR of the recovered signals by each algorithm. In each case, our IPNN consistently beats the others, and even when there is more prior information, the effect of the algorithm is better.

Table 1 Output SNR of recovered signal for IPNN, IRLS, and seq under Gaussian measurement noise with SNR= $\mathrm{SNR_M}$

	$q = 0.3$				
$\mathrm{SNR_M}$	IPNN ($s = 0$)	IPNN ($s = 5$)	IPNN ($s = 10$)	IRLS-L_q	seq-L_q
−30	−6.855	−6.628	−6.622	−24.889	−21.606
−20	2.033	4.767	4.829	−14.450	−13.456
−10	21.300	21.638	22.297	−4.598	−1.499
10	38.159	38.169	38.185	13.957	13.541
20	48.101	48.136	48.185	25.594	25.934
30	57.857	58.029	58.185	39.375	38.542
	$q = 0.5$				
$\mathrm{SNR_M}$	IPNN ($s = 0$)	IPNN ($s = 5$)	IPNN ($s = 10$)	IRLS-L_q	seq-L_q
−30	−6.633	−6.195	−6.020	−24.788	−21.592
−20	8.048	9.720	13.4916	−15.698	−13.477
−10	21.300	22.634	22.297	−5.209	−1.601
10	35.825	37.451	38.642	14.247	13.549
20	44.111	45.369	48.184	26.996	25.877
30	55.914	56.471	58.182	39.570	38.433
	$q = 0.7$				
$\mathrm{SNR_M}$	IPNN ($s = 0$)	IPNN ($s = 5$)	IPNN ($s = 10$)	IRLS-L_q	seq-L_q
−30	−6.379	−6.409	−5.676	−26.744	−21.596
−20	5.469	7.499	7.529	−14.585	−13.581
−10	20.423	22.666	23.361	−5.764	−1.4902
10	33.558	34.605	38.182	14.109	13.376
20	44.894	45.242	48.183	27.118	25.945
30	53.803	54.649	55.373	39.543	38.217

We also test the algorithm with colored noise, which is generated by

$$e(i) = \frac{\xi(i) + 0.5\xi(i-1) + 0.2\xi(i-2)}{\xi(i) - 1.5\xi(i-1) + 0.7\xi(i-2) + 0.1\xi(i-3)},$$

where $\xi(i)$ is a Gaussian white noise sequence with mean 0 and variance 1. Then, we can superimpose the colored noise sequence satisfying different SNR into the measurement signal $\mathbf{Ax}$, and get the measurement signal with colored noise. We set $n = 128, m = 64, k = 10, s = 5$, for $q = 0.3$. Similar to the Gaussian white noise experiment, we can get the results in Table 2. Most of our IPNN algorithm is better than other algorithms.

Finally, in order to further prove the performance of our IPNN algorithm in some practical applications, we apply our algorithm, together with the algorithms used above like IRLS-L_q and set-L_q to recover the FECG signals [7]. We only intercept a segment of the FECG signal and use the Gaussian random matrix of the size $125 * 250$ as the measurement matrix. Fig. 11 plots the recovery results of three different algorithms for $q = 0.3$. It is easy to see that our algorithm is closer to the original segment.

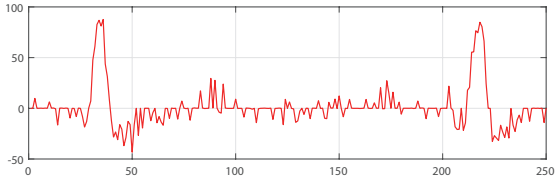
References

1. Foucart S, Lai M J. Sparsest solutions of underdetermined linear systems via ℓ_p-minimization for $0 < q \leq 1$. Applied and Computational Harmonic Analysis, 2009, 26(3): 395–407
2. Chartrand R, Yin W. Iteratively reweighted algorithms for compressive sensing. In: 2008 IEEE international conference on acoustics, speech and signal processing. 2008, 3869–3872
3. Candès E J, Plan Y. Near-ideal model selection by ℓ_1 minimization. The Annals of Statistics, 2009, 37(5A): 2145–2177

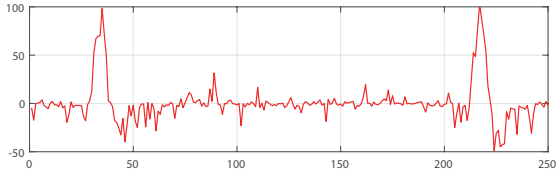
Table 2 Output SNR of recovered signal for various algorithms under colored measurement noise with SNR= $\mathrm{SNR_M}(q = 0.3)$

$\mathrm{SNR_M}$	IPNN ($s = 5$)	IRLS-L_q	seq-L_q	IPNN-L_{1-2}	PD-L_1	ADMM-lasso	ISTA
−30	−32.143	−33.066	−34.951	−29.841	−29.567	−31.367	−26.091
−20	−20.041	−24.986	−24.693	−19.923	−19.660	−20.926	−16.544
−10	−6.270	−15.621	−14.810	−8.583	−8.884	−12.632	−6.510
10	24.194	2.882	7.019	12.991	9.143	7.114	9.988
20	34.183	13.940	18.245	20.301	16.846	16.850	12.412
30	44.145	22.848	26.242	33.049	30.050	26.046	12.150

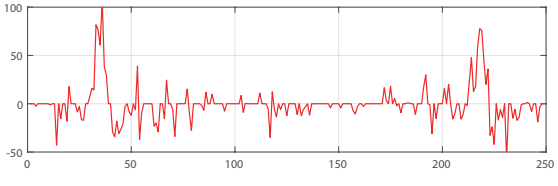
(a) Segment from FECG signals



(b) IPNN with RE= 0.3494



(c) IRLS-L_q with RE=0.3895



(d) set-L_q with RE=0.3996

Fig. 11 recovery results of the real signal via different algorithms for $q = 0.3$.

4. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J, others . Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine learning, 2011, 3(1): 1–122
5. Donoho D L. De-noising by soft-thresholding. IEEE transactions on information theory, 1995, 41(3): 613–627
6. Zhu L, Wang J, He X, Zhao Y. An inertial projection neural network for sparse signal reconstruction via l1- 2 minimization. Neurocomputing, 2018, 315: 89–95
7. Moor B D. Daisy: Database for the identification of systems, 2011