RESEARCH

# **Open Access**

# Semi-tensor product-based one-bit compressed sensing



Jingyao Hou<sup>1\*</sup> and Xinling Liu<sup>1</sup>

\*Correspondence: hjy17623226280@163.com

<sup>1</sup> Key Laboratory of Optimization Theory and Applications at China West Normal University of Sichuan Province, School of Mathematics and Information, Nanchong 637009, China

# Abstract

The area of one-bit compressed sensing (1-bit CS) focuses on the recovery of sparse signals from binary measurements. Over the past decade, this field has witnessed the emergence of well-developed theories. However, most of the existing literature is confined to fully random measurement matrices, like random Gaussian and random sub-Gaussian measurements. This limitation often results in high generation and storage costs. This paper aims to apply semi-tensor product-based measurements to 1-bit CS. By utilizing the semi-tensor product, this proposed method can compress high-dimensional signals using lower-dimensional measurement matrices, thereby reducing the cost of generating and storing fully random measurement matrices. We propose a regularized model for this problem that has a closed-form solution. Theoretically, we demonstrate that the solution provides an approximate estimate of the underlying signal with upper bounds on recovery error. Empirically, we conduct a series of experiments on both synthetic and real-world data to demonstrate the proposed method's ability to utilize a lower-dimensional measurement matrix for signal compression and reconstruction with enhanced flexibility, resulting in improved recovery accuracy.

Keywords: 1-Bit compressed sensing, Semi-tensor, L1-norm, Dither

# 1 Introduction

Compressed (CS) has drawn widely studied and applied in many fields such as moderate resolution imaging [1, 2] and magnetic resonance imaging (MRI). One-bit compressed sensing (1-bit CS) [3–9] has garnered significant attention in recent years as a unique paradigm of (CS) [10, 11]. In contrast to the classic compressed sensing method, which utilizes an infinite-precision real-valued measurement  $\boldsymbol{b} = A\boldsymbol{x}$  where  $A \in \mathbb{R}^{m \times n}$  is the measurement matrix and  $\boldsymbol{x} \in \mathbb{R}^n$  denotes a sparse signal, 1-bit CS only records the signs of real-valued measurements as  $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{b}) = \operatorname{sign}(A\boldsymbol{x})$ . Using binary measurement, 1-bit CS offers a cost-effective hardware implementation and high-speed sampling. Additionally, it exhibits greater resilience to various nonlinear distortions commonly encountered in input electronics [5]. Due to the aforementioned reasons, it possesses a wide range of potential applications in massive MIMO systems [12, 13], wireless sensor networks [14], synthetic aperture radar systems [15], and other scenarios where large-scale sparse data are typically involved [16].



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

Mathematically, standard 1-bit CS aims to recover the direction of the underlying sparse signal  $\mathbf{x} \in \mathbb{R}^n$  (i.e.,  $\mathbf{x}/||\mathbf{x}||_2$ ) from 1-bit measurements

$$y = \operatorname{sign}(Ax),\tag{1}$$

where sign(·) applies componentwise on real-valued measurement  $A\mathbf{x}$ , satisfying sign( $\nu$ ) = 1 if  $\nu \ge 0$ , and sign( $\nu$ ) = -1 if  $\nu < 0$ . The standard measurement model (1) fails to capture information on the magnitude of  $\mathbf{x}$  (i.e.,  $\|\mathbf{x}\|_2$ ), resulting in the aforementioned studies being limited to the unit sphere  $\mathbb{S}_n := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}$ . Jacques et al. [5] have proposed an  $\ell_0$ -minimization method for this problem, which has been proven to yield an optimal solution that provides a robust estimate of the original signal. Although a binary iterative hard threshold (BIHT) algorithm has been developed to approximate the underlying solution empirically, solving the  $\ell_0$ -minimization model exactly remains an NP-hard problem. Plan and Vershynin [6] utilize the  $\ell_1$ -norm as a means of promoting sparsity, proposing a convex programming approach that is also supported by theoretical guarantees. Moreover, their method demonstrates robustness against bit flips, a prevalent form of noise in binary measurements. Zhang et al. [7] suggest the following regularized model

$$\hat{\boldsymbol{u}} := \underset{\|\boldsymbol{u}\|_{2} \leq 1}{\arg\min} - \frac{1}{m} \langle \boldsymbol{A}^T \boldsymbol{y}, \boldsymbol{u} \rangle + \gamma \|\boldsymbol{u}\|_{1},$$
(2)

and present the existence of a closed-form solution, namely  $S_{\gamma}(A^T y)/\|S_{\gamma}(A^T y)\|_2$ where  $S_{\gamma}(\cdot)$  denotes the soft-thresholding operator [psee Eq. (22)]. However, the  $\ell_1$ norm is a loose approximation of  $\ell_0$ -quasi-norm, often resulting in an over-penalized problem [17]. In our prior research [8], we proposed an  $\ell_p(0 -minimization$  $method and demonstrated its superiority over the <math>\ell_1$ -norm based counterpart in terms of reducing the required number of measurements.

The aforementioned studies have solely focused on fully random Gaussian distributions, which limits their applicability in processing large-scale data. Unfortunately, Ai et al. [18] have shown that even with the commonly used fully random sub-Gaussian measurement matrices, it is impossible to recover any sparse unit signal from (1) without additional constraints (e.g., x is not too sparse). To mitigate this issue, it has recently been discovered that introducing dithers can be beneficial. Let

$$y = \operatorname{sign}(Ax + \tau), \tag{3}$$

where  $\tau \in \mathbb{R}^m$  denotes a random dither vector. Dirksen and Mendelson [19] demonstrated that a convex program can be applied to recover the signal  $x \in \mathbb{R}^n$  from (3) when  $\tau$  is generated from a uniform distribution, and A is a random sub-Gaussian measurement matrix. However, fully sub-Gaussian measurements still pose challenges in terms of high costs for generating and storing measurement matrices. Dirksen et al. [20] showed that randomly subsampled Gaussian circulant matrices can be utilized for 1-bit CS, and Liu et al. [21] provide an optimization algorithm based on it under generative prior. Dirksen et al. [22] further applied randomly subsampled sub-Gaussian circulant matrices to 1-bit CS and prove that the proposed algorithm is robust to analog prequantization noise as well as to adversarial bit corruptions in the quantization process. It should be noted that circulant measurement matrices are typically limited to subsampling, while oversampling is commonly employed in 1-bit CS despite being irrelevant in traditional CS problems. In practical applications of circulant matrices, we have to introduce the zero-padding technique on the original signal to accommodate oversampling, which lifts the original signal to a higher dimension. Additionally, a square measurement matrix with a matching dimension has to be constructed. In summary, the research on designing measurement matrices for 1-bit CS is still in its early stages.

The concept of semi-tensor product of matrices was proposed by Cheng et al., which is a generalization of conventional matrix product [23–25]. The introduction of semi-tensor product enables matrix multiplication even when two matrices do not meet the dimension matching condition. This concept has found applications in various engineering fields, including but not limited to information security [26], vehicle control [27], and gene regulation [28]. Specially, Xie et al. [29] proposed a new model of signal compression and reconstruction based on the semi-tensor product, which breaks the dimension matching conditions required by the traditional CS model. The novel approach employs lower-dimensional sensing matrices to compress high-dimensional signals, while simultaneously reducing total reconstruction time by conducting the reconstruction phase among multiple CS decoders. Inspired by these advantages, we are considering the potential application of semi-tensor product in 1-bit CS to tackle the challenge of storing large measurement matrices. The primary contributions of this paper are as follows.

- We first apply the semi-tensor product to the measurement model of 1-bit CS with employing the dither technique. The novel measurement model significantly reduces the dimensionality of the measurement matrix compared to existing fully random counterparts in 1-bit CS.
- We propose a regularized model with a closed-form solution and demonstrate that our proposed model, which employs an explicit selection strategy for the regularization parameter, can yield an approximate estimate of the underlying sparse signal with guaranteed convergence.

The rest of this paper is structured as follows. Section 2 introduces some notations, definitions, and useful lemmas. Section 3 clarifies the measurement model used in this paper. Section 4 elucidates the recovery model and then gives the corresponding theoretical guarantee. Section 5 presents numerical experiments, and Sect. 6 concludes this paper.

# 2 Preliminaries

# 2.1 Notations

We begin with some notations used throughout the paper. Scalars are denoted with lowercase letters, e.g., x, vectors with boldface lowercase letters, e.g., x and matrices with boldface capital letters, e.g., X. We denote  $x_i$  as the *i*-th element of the vector x. The fields of real numbers are denoted as  $\mathbb{R}$ . [n] denotes the index set  $\{1, 2, \ldots, n\}$ .  $I_n$  denotes identity matrix of size  $n \times n$ . The inner

product between matrices  $M = (m_{ij}) \in \mathbb{R}^{n_1 \times n_2}$  and  $N = (n_{ij}) \in \mathbb{R}^{n_1 \times n_2}$  of the same size is defined as  $\langle M, N \rangle = \operatorname{Tr}(M^T N) = \sum_{ij} m_{ij} n_{ij}$ . We define the  $\ell_p$  norm of  $\mathbf{x} \in \mathbb{R}^n$  as  $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$  where  $p \ge 1$ .  $\|\mathbf{x}\|_{\infty}$  denotes the absolute value of the elements of  $\mathbf{x}$  that has the biggest magnitude.

# 2.2 Semi-tensor product and some definitions

**Definition 1** (See [25]) Let  $u \in \mathbb{R}^{1 \times nl}$  be a row vector, and  $v \in \mathbb{R}^{l}$  be a column vector. Divide u into l contiguous blocks, named  $u_1, \ldots, u_l$ , which are all row vectors of length n. Then, the semi-tensor product (STP) between u and v, denoted by  $u \ltimes v$ , is defined as

$$\boldsymbol{u} \ltimes \boldsymbol{v} = \sum_{i=1}^{l} \boldsymbol{u}_i \boldsymbol{v}_i \in \mathbb{R}^{1 \times n}.$$

**Definition 2** (See [25]) Let  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times q}$ ,  $a_i$  denote the *i*-th row of A, and  $b^i$  denote the i-th column of B. The STP of A and B, denoted as  $C = A \ltimes B$ , consists of  $m \times q$  blocks as  $C = (c_{ij})$  where

- (i)  $c_{ii} = a_i \ltimes b^i$  is of length *p* when *p* is a factor of *n*, and
- (ii)  $\mathbf{c}_{ij} = (\mathbf{b}^{i^T} \ltimes \mathbf{a}_i^T)^T$  is of length *n* when *n* is a factor of *p*,

for  $i \in [m]$  and  $j \in [q]$ . Specially, when q = 1, C reduces to a vector of length mk where k := gcd(n, p) denotes the greatest common divisor(gcd).

We also introduce the Kronecker product, which is a mathematical operation that combines two matrices to form a larger matrix with block entries.

**Definition 3** The Kronecker product between matrices  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$  is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}$$

It is readily checked that the Kronecker product has the following property.

**Proposition 1** For any matrices  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$ , it holds that

 $(\boldsymbol{A}\otimes\boldsymbol{B})^T=\boldsymbol{A}^T\otimes\boldsymbol{B}^T.$ 

The STP of matrices can be equivalently defined through the Kronecker product.

**Definition 4** (See [24]) The STP of two matrices  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$  satisfies

 $A \ltimes B = (A \otimes I_{t/n})(B \otimes I_{t/p}),$ 

where *t* is the least common multiple of *n* and *p*, denoted as t = lcm(n, p).

The following lemma was proved in [7], which may facilitate our analysis.

**Lemma 1** (See [7]) Suppose  $\mathbf{b} = \theta(\mathbf{A}\mathbf{u})$  with  $\theta(\cdot)$  satisfying (10), and  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is populated by i.i.d. random standard Gaussian variables. With probability at least  $1 - e^{1-t}$ , it holds that

$$\left\|\frac{1}{m}\boldsymbol{A}^{T}\boldsymbol{b}-\lambda\boldsymbol{u}\right\|_{\infty}\leq O\left(\sqrt{\frac{t+\log(n)}{m}}\right).$$
(4)

# **3** Problem description

Let *A* be the measurement matrix of size  $m \times n$ . We suppose the signal x is of length nk where  $k \ge 1$  denotes some integer. As a result, the measurement matrix only requires 1/k of the parameters compared to traditional sensing models, which typically have a size of  $m \times nk$ . The standard measurement model for 1-bit CS under STP may be naturally given by

$$y = \operatorname{sign}(A \ltimes x), \tag{5}$$

where  $y \in \{-1, 1\}^{mk}$  denotes the sign measurements. Note that (5) loses any scaling information, and is restricted to narrow types of measurement such as random Gaussian distribution. To circumvent this issue, a mature strategy is to introduce random dither  $\tau \in \mathbb{R}^{mk}$  before quantization, leading to the dithered one-bit measurements

$$y = \operatorname{sign}(A \ltimes x + \tau), \tag{6}$$

where  $y \in \{-1, 1\}^{mk}$ . Please see Sect. 4.3 and [3] for more discussion on the dither technology. To facilitate discussion, we define the exactly sparse set

$$\mathbb{K}_{s}^{*} := \{ \boldsymbol{z} \in \mathbb{R}^{nK} | \| \boldsymbol{z} \|_{2} = 1, \| \boldsymbol{z} \|_{0} \le s \},$$
(7)

and the approximately sparse set

$$\mathbb{K}_{s} := \{ \boldsymbol{z} \in \mathbb{R}^{nk} | \| \boldsymbol{z} \|_{2} = 1, \| \boldsymbol{z} \|_{1} \le s \}.$$
(8)

In practice, there will always be adversarial noise in the measurements during acquisition and transmission, which usually flip some signs [5]. In this case, the noisy measurements may be modeled as

$$\mathbf{y} = \boldsymbol{\xi} \odot \operatorname{sign}(\boldsymbol{A} \ltimes \boldsymbol{x} + \boldsymbol{\tau}), \tag{9}$$

where  $\odot$  denotes the dot product between vectors,  $\boldsymbol{\xi} \in \{-1, 1\}^{mk}$  denotes random bit flips, which are populated by i.i.d. Bernoulli random variables satisfying  $P(\xi_i = 1) = p$ , meaning that each sign measurement  $y_i$  is only correct with probability p.

To illustrate how to construct a loss function for the noisy measurements (9), inspired by [6], we begin with a more general measurement model. We assume

$$E(y_i|z_i) = \theta(z_i), \quad \forall i \in [m], \tag{10}$$

where  $z_i = (A \ltimes x)_i$ ,  $x \in \mathbb{R}^n$  denotes the needed-to-be-recovered signal, and  $\theta(\cdot)$  is some function unknown or unspecified, which automatically must satisfy  $-1 \le \theta(z) \le 1$ . We also assume

$$E(\theta(g)g) := \lambda, \tag{11}$$

to capture the relation between  $y_i$  and  $(A \ltimes x + \tau)_i$ , where g is a standard normal random variable. In (9), since  $y_i = \xi_i \operatorname{sign}((A \ltimes x)_i + \tau_i)$  for each  $i \in [m]$ , then it holds that  $\theta(z) = E_{\tau_i}(E_{\xi_i}(y_i|z,\tau_i)) = 2(p-1/2)(1-2P(\tau \leq -z))$  in (10), and the correlation coefficient  $\lambda = E(\theta(g)g)$  can be evaluated using integration by parts, which gives

$$\lambda = 2(p - 1/2)E[\theta'(g)] = 4(p - 1/2)E[f(-g)] > 0.$$

For instance, if  $\tau_i(\forall i \in [mk])$  are random variables generated from the uniform distribution

$$Unif(-u, u) \tag{12}$$

where c > 0 denotes a constant, then

$$\lambda = 2(p - 1/2)\frac{2\Phi(u) - 1}{u},\tag{13}$$

where  $\Phi(\cdot)$  denotes the distribution function of standard Gaussian variables.

We formulate theories based on model (10) in the following section as it naturally encompasses the model presented in model (9).

## 4 Main results

This part introduces the optimization model for recovering the underlying signal from model (10) and subsequently establishes its corresponding theoretical guarantee.

# 4.1 The proposed model

Suppose  $x^* \in \mathbb{R}^{nk}$  is the underlying signal in  $\mathbb{K}_s$  where k is a positive integer, and y satisfies (10). We reorder its entries to define  $\tilde{x}_*$  such that

$$\tilde{\boldsymbol{x}}_{*}^{(l)} = \begin{pmatrix} \boldsymbol{x}_{l} \\ \boldsymbol{x}_{k+l} \\ \vdots \\ \boldsymbol{x}_{(n-1)k+l} \end{pmatrix} \in \mathbb{R}^{n}, \quad l \in [k],$$
(14)

and also reorder entries of y to define  $\tilde{y}$  such that

$$\tilde{\boldsymbol{y}}^{(l)} = \begin{pmatrix} \boldsymbol{y}_l \\ \boldsymbol{y}_{k+l} \\ \vdots \\ \boldsymbol{y}_{(m-1)k+l} \end{pmatrix} \in \mathbb{R}^m, \quad l \in [k].$$
(15)

A direction application of Lemma 4 in [7] to each  $\tilde{y}^{(l)}$  indicates the following result.

**Lemma 2** Suppose  $A \in \mathbb{R}^{m \times n}$  is populated by i.i.d. random standard Gaussian variables and  $a_i$  denotes the *i*-th row of A. Following the definitions given by (10) and (11), it holds that

$$E\left[\tilde{\boldsymbol{y}}_{i}^{(l)}\boldsymbol{a}_{i}\right] = \theta(\tilde{\boldsymbol{x}}_{*}^{(l)})^{T}, \,\forall i \in [m], \quad \forall l \in [k],$$

$$(16)$$

where  $\tilde{\mathbf{x}}_{*}^{(l)}$ ,  $\tilde{\mathbf{y}}^{(l)}$  are, respectively, defined as in (14), (15), and  $\tilde{y}_{i}^{(l)}$  represents the *i*-th element of  $\tilde{\mathbf{y}}^{(l)}$ .

By taking the inner product in both sides of (16) with respect to  $\tilde{x}_*^{(l)}$  and then summing over all  $l \in [m]$ , we can derive that

$$E\left[\sum_{l=1}^{k} \tilde{y}_{i}^{(l)} \left\langle \boldsymbol{a}_{i}, \tilde{\boldsymbol{x}}_{*}^{(l)} \right\rangle\right] = \theta \|\boldsymbol{x}_{*}\|_{2}^{2} = \theta, \quad \forall i \in [m].$$

$$(17)$$

When  $\theta(\cdot) = \text{sign}(\cdot)$ ,  $\lambda = E(|g|) = \sqrt{2/\pi}$  achieves the maximum relation [6]. According to the law of large numbers, Eq. (17) suggests that the optimal solution  $\hat{x}$  is expected to also achieve a large value for the relation function

$$\frac{1}{m}\sum_{i=1}^{m}\sum_{l=1}^{k}\tilde{y}_{i}^{(l)}\langle \boldsymbol{a}_{i},\hat{\boldsymbol{x}}\rangle$$

under the constraint  $\|\hat{x}\|_2 \leq 1$ . It is not hard to check that

$$\frac{1}{m}\sum_{i=1}^{m}\sum_{l=1}^{k}\tilde{y}_{i}^{(l)}\langle \boldsymbol{a}_{i},\hat{\boldsymbol{x}}\rangle = \frac{1}{m}(\boldsymbol{A}\ltimes\hat{\boldsymbol{x}})^{T}\boldsymbol{y}.$$
(18)

On the other hand, given the sparsity of x, it is imperative that  $\hat{x}$  also exhibits sparsity, which can be achieved through regularization using the  $\ell_1$ -norm  $\|\cdot\|_1$ . By combining (18) with the  $\ell_1$ -norm regularizer via a regularization parameter  $\gamma > 0$ , we propose the optimization model

$$\min_{\|\boldsymbol{u}\|_{2}\leq 1} -\frac{1}{m} (\boldsymbol{A} \ltimes \boldsymbol{u})^{T} \boldsymbol{y} + \gamma \|\boldsymbol{u}\|_{1}.$$
(19)

Since  $A \ltimes u = (A \otimes I_k)u$  by Definition 4 and Proposition 1, we can equivalently rewrite (19) as

$$\min_{\|\boldsymbol{u}\|_{2} \leq 1} - \frac{1}{m} \boldsymbol{u}^{T} (\boldsymbol{A}^{T} \otimes \boldsymbol{I}_{k}) \boldsymbol{y} + \gamma \|\boldsymbol{u}\|_{1}.$$
(20)

### 4.2 Theoretical guarantee

The following result shows that the model (20) has a closed-form solution.

**Lemma 3** (See [7]) *The optimal solution to* (20) *is given by* 

$$\hat{\boldsymbol{x}} = \frac{S_{\gamma}\left(\left(\boldsymbol{A}^{T} \otimes \boldsymbol{I}_{k}\right)\boldsymbol{y}\right)}{\left\|S_{\gamma}\left(\left(\boldsymbol{A}^{T} \otimes \boldsymbol{I}_{k}\right)\boldsymbol{y}\right)\right\|_{2}},\tag{21}$$

where the soft-thresholding operator  $S_{\gamma}(\cdot)$ 

$$S_{\gamma}(u) = \begin{cases} 0, & \text{if } |u| \le \gamma, \\ \operatorname{sign}(u)(|u| - \gamma), & \text{otherwise.} \end{cases}$$
(22)

applies componentwise.

To prove the main theorem, we first provide the following lemma.

**Lemma 4** Suppose  $A \in \mathbb{R}^{m \times n}$  is populated by independent and identically distributed random standard Gaussian variables,  $\mathbf{x}^* \in \mathbb{R}^{nk}$  where k is a positive integer, and  $\mathbf{y} = \theta(\mathbf{A} \ltimes \mathbf{x}_*)$  with  $\theta(\cdot)$  satisfying (10). Then, with probability at least  $1 - ke^{1-t}$ ,

$$\left\|\frac{1}{m}(\boldsymbol{A}^{T}\otimes\boldsymbol{I}_{k})\boldsymbol{y}-\lambda\boldsymbol{x}_{*}\right\|_{\infty}\leq\kappa\sqrt{\frac{t+\log(n)}{m}},$$
(23)

where  $\kappa > 0$  is a constant, and  $\lambda$  is specified by (11).

**Proof** We define  $\tilde{x}_*$  and  $\tilde{y}_*$  as given in (14) and (15), respectively. For each  $l \in [k]$ , we have  $\tilde{y}^{(l)} = A\tilde{x}_*^{(l)}$  and

$$\left\|\frac{1}{m}(\boldsymbol{A}^T \otimes \boldsymbol{I}_k)\boldsymbol{y} - \lambda \boldsymbol{x}_*\right\|_{\infty} = \max_{l \in [k]} \left\|\frac{1}{m}\boldsymbol{A}^T \tilde{\boldsymbol{y}}^{(l)} - \lambda \tilde{\boldsymbol{x}}_*^{(l)}\right\|_{\infty}.$$

Since Lemma 1 indicates that

$$\left\|\frac{1}{m}\boldsymbol{A}^{T}\tilde{\boldsymbol{y}}^{(l)} - \lambda \tilde{\boldsymbol{x}}_{*}^{(l)}\right\|_{\infty} \leq O\left(\sqrt{\frac{t + \log(n)}{m}}\right)$$

holds with a probability at least  $1 - e^{1-t}$  for each  $l \in [k]$ , taking a union bound over all l implies that

$$\left\|\frac{1}{m} \left(\boldsymbol{A}^{T} \otimes \boldsymbol{I}_{k}\right) \boldsymbol{y} - \lambda \boldsymbol{x}_{*}\right\|_{\infty} \leq O\left(\sqrt{\frac{t + \log(n)}{m}}\right)$$
(24)

holds with a probability at least  $1 - ke^{1-t}$ . Therefore, we complete our proof.  $\Box$ 

Now, we can present the main theoretical result.

**Theorem 1** Assume  $A \in \mathbb{R}^{m \times n}$  is populated by independent and identically distributed random standard Gaussian variables, y is observed via (9),  $\lambda$  is defined in Eq. (13), and

$$\gamma = c\sqrt{\frac{t + \log(n)}{m}},\tag{25}$$

where  $c = 2\kappa$  is a constant with  $\kappa$  coming from (23). If  $\mathbf{x}^* \in \mathbb{R}^{nk}$  belongs to (7) where k is a positive integer, with a probability at least  $1 - ke^{1-t}$ , the estimation  $\hat{\mathbf{x}}$  given by (21) satisfies

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \le \frac{3\gamma}{\lambda}\sqrt{s} = O\left(\sqrt{\frac{s\log(n)}{m}}\right).$$

If  $\mathbf{x}^* \in \mathbb{R}^{nk}$  belongs to (8), with probability exceeding  $1 - ke^{1-t}$  the estimation  $\hat{\mathbf{x}}$  given by (21) satisfies

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \le \sqrt{\frac{3\gamma\sqrt{s}}{\lambda}} = O\left(\sqrt[4]{\frac{s\log(n)}{m}}\right).$$

*Proof* Since  $\hat{x}$  is the optimal solution, we have

$$\begin{aligned} &-\frac{1}{m} \hat{\boldsymbol{x}}^T (\boldsymbol{A}^T \otimes \boldsymbol{I}_k) \boldsymbol{y} + \gamma \, \| \hat{\boldsymbol{x}} \|_1 \\ &\leq -\frac{1}{m} \boldsymbol{x}_*^T (\boldsymbol{A}^T \otimes \boldsymbol{I}_k) \boldsymbol{y} + \gamma \, \| \boldsymbol{x}_* \|_1 \end{aligned}$$

Thus one has

$$\begin{split} & \gamma \| \boldsymbol{x}_* \|_1 \\ \geq \left\langle \boldsymbol{x}_* - \hat{\boldsymbol{x}}, \frac{1}{m} (\boldsymbol{A}^T \otimes \boldsymbol{I}_k) \boldsymbol{y} \right\rangle + \gamma \| \hat{\boldsymbol{x}} \|_1 \\ \geq \left\langle \boldsymbol{x}_* - \hat{\boldsymbol{x}}, \lambda \boldsymbol{x}_* \right\rangle + \left\langle \boldsymbol{x}_* - \hat{\boldsymbol{x}}, \frac{1}{m} (\boldsymbol{A}^T \otimes \boldsymbol{I}_k) \boldsymbol{y} - \lambda \boldsymbol{x}_* \right\rangle \\ & + \gamma \| \hat{\boldsymbol{x}} \|_1 \\ \geq \lambda (1 - \hat{\boldsymbol{x}}^T \boldsymbol{x}_*) - \| \boldsymbol{x}_* - \hat{\boldsymbol{x}} \|_1 \left\| \frac{1}{m} (\boldsymbol{A}^T \otimes \boldsymbol{I}_k) \boldsymbol{y} - \lambda \boldsymbol{x}_* \right\|_{\infty} \\ & + \gamma \| \hat{\boldsymbol{x}} \|_1 \end{split}$$

Combining it with Lemma 4, we have that

$$\lambda \Big( 1 - \hat{\boldsymbol{x}}^T \boldsymbol{x}_* \Big) + \gamma \| \hat{\boldsymbol{x}} \|_1 \le \frac{\gamma}{2} \| \boldsymbol{x}_* - \hat{\boldsymbol{x}} \|_1 + \gamma \| \boldsymbol{x}_* \|_1.$$
(26)

The remain discussion considers the two distinct types of sparse signals, respectively. (1) When  $x_*$  is exactly sparse, i.e.,  $x_* \in \mathbb{K}_s^*$ , we define *S* as the support set of  $x_*$  and  $S_{\perp} = [nk] \setminus S$  be its complement set. We also define  $P_S(u)$  as

$$P_S(\boldsymbol{u}) = \boldsymbol{u}_S,$$

denoting the sub-vector of u indexed by the set *S*. Then (26) implies

$$\begin{split} \lambda \Big( 1 - \hat{\boldsymbol{x}}^T \boldsymbol{x}_* \Big) &+ \gamma \| P_S(\hat{\boldsymbol{x}}) \|_1 + \gamma \| P_{S_\perp}(\hat{\boldsymbol{x}}) \|_1 \\ &\leq \gamma \| \boldsymbol{x}_* \|_1 + \frac{\gamma}{2} \| P_S(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_1 + \frac{\gamma}{2} \| P_{S_\perp}(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_1 \\ &= \gamma \| \boldsymbol{x}_* \|_1 + \frac{\gamma}{2} \| P_S(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_1 + \frac{\gamma}{2} \| P_{S_\perp}(\hat{\boldsymbol{x}}) \|_1. \end{split}$$

Therefore, using triangle inequality yields

$$\begin{split} \lambda \Big( 1 - \hat{\boldsymbol{x}}^T \boldsymbol{x}_* \Big) &+ \frac{\gamma}{2} \| P_S(\hat{\boldsymbol{x}}) \|_1 \\ &\leq \gamma \| \boldsymbol{x}_* \|_1 - \gamma \| P_S(\hat{\boldsymbol{x}}) \|_1 + \frac{\gamma}{2} \| P_S(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_1 \\ &\leq \frac{3\gamma}{2} \| P_S(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_1 \leq \frac{3\gamma}{2} \sqrt{\| \boldsymbol{x}_* \|_0} \| P_S(\boldsymbol{x}_* - \hat{\boldsymbol{x}}) \|_2, \end{split}$$

where the final inequality follows the Cauchy–Schwarz inequality. Recalling  $\|\boldsymbol{x}_* - \hat{\boldsymbol{x}}\|_2^2 \leq 2(1 - \hat{\boldsymbol{x}}^T \boldsymbol{x}_*)$ , we obtain that

$$\|\boldsymbol{x}_* - \hat{\boldsymbol{x}}\|_2^2 \leq \frac{3\gamma}{\lambda}\sqrt{s}\|\boldsymbol{x}_* - \hat{\boldsymbol{x}}\|_2,$$

i.e.,  $\|\boldsymbol{x}_* - \hat{\boldsymbol{x}}\|_2 \leq \frac{3\gamma}{\lambda}\sqrt{s}.$ 

(2) When  $x_*$  is approximately sparse, i.e.,  $x_* \in \mathbb{K}_s$ , from (26), we have

$$egin{aligned} &\lambda \Big(1-\hat{oldsymbol{x}}^T oldsymbol{x}_* \Big) \ &\leq \gamma \|oldsymbol{x}_*\|_1 - \gamma \|\hat{oldsymbol{x}}\|_1 + rac{\gamma}{2} \|oldsymbol{x}_*\|_1 + rac{\gamma}{2} \|\hat{oldsymbol{x}}\|_1 \ &\leq rac{3\gamma}{2} \|oldsymbol{x}_*\|_1, \end{aligned}$$

which indicates

$$\|m{x}_* - \hat{m{x}}\|_2^2 \leq 2\Big(1 - \hat{m{x}}^Tm{x}_*\Big) \leq rac{3\gamma}{\lambda}\|m{x}_*\|_1 \leq rac{3\gamma}{\lambda}\sqrt{s}.$$

We complete our proof.

**Remark 1** Theorem 1 offers a practical recommendation for selecting the regularization parameter  $\gamma$  in model (20). As demonstrated experimentally in Sect. 5.1, when combined with finely tuned *c* in (25), the regularization parameter in (20) performs well on general sparse signals.

**Remark 2** According to the suggestion put forward by [7], we have selected u = 1 in (12) for our methodology. In this case, we have  $\lambda \approx 1.365(p - 1/2)$ , meaning that recovering the signal  $x_*$  is still possible even if each measurement is flipped with probability nearly 1/2. In particular, when k = 1, the measurement model (9) reduces to the traditional 1-bit CS model. Therefore, the semi-tensor-based 1-bit CS can be considered as a generalization of traditional 1-bit CS, implying that Theorem 1 extends Theorem 1 in [7].

# 4.3 Discussion on dither technique

If  $a_i$  denotes the *i*-th row of the measurement matrix A, then  $q_i = \text{sign}(\langle a_i, x \rangle)$  indicates which side of the hyperplane  $H_{a_i} : \langle a_i, u \rangle = 0$  ( $i \in [m]$ ) the point x lies on. As all  $a_i$  are generated randomly, the sign vector q = sign(Ax) encodes the cell of a random hyperplane



**Fig. 1** Illustration on random hyperplane tessellation in  $\mathbb{R}^2$ . **a** corresponds to homogeneous hyperplanes in which case any two points on a ray from the origin cannot be separated even if one uses all possible hyperplanes. **b** Corresponds to hyperplanes with random parallel shifts (resulted by random dithers) in which case we can approximate distances between signals

tessellation [30] in which the signal x is located. A popular strategy used for recovering x is searching for a vector  $x^{\#}$  that is quantization consistent, i.e.,  $q = \operatorname{sign}(Ax^{\#})$ . However, due to the homogeneity of these hyperplanes, it is impossible to separate any two points on a ray from the origin even when all possible hyperplanes are utilized (see Fig. 1a). In other words, the measurement  $q = \operatorname{sign}(Ax)$  loses information regarding the signal's scale. Fortunately, the issue can be circumvented by introducing a dither vector prior to quantization. In the context of random tessellations, dithering can be geometrically interpreted as adding random parallel shifts to homogeneous hyperplanes, resulting in nonhomogeneous hyperplanes  $H_{a_i,\tau_i} : \langle a_i, u \rangle + \tau_i = 0$  ( $i \in [m]$ ).

As illustrated in Fig. 1, the random parallel shifts enable the appropriate separation between two points on a ray originating from the origin, provided that  $\tau$  is suitably selected. Although we focus on unit signals in this paper, our proof process reveals that under the circumstances of semi-tensor product measurement, sparse signals are restored by segments. This necessitates coding the signal scale in quantization to prevent magnitude distortion during recovery. Consequently, we have employed the dithered measurement (6) and its noisy variant (9) in this study. Indeed, Chen et al. [31] proved the following result:

**Proposition 2** (Lemma 1, [31]) Let  $x, \tau$  be two independent random variables satisfying  $|x| \le B, \tau \sim uni([-\gamma, \gamma])$  where  $\gamma \ge B$ , then we have

 $\mathbb{E}(\gamma \cdot \operatorname{sign}(x + \tau)) = \mathbb{E}(x).$ 

Proposition 2 states that if all elements of  $\tau$  are drawn from the uniform distribution  $[-\gamma, \gamma]$ , where the bound  $\gamma > 0$  is chosen such that  $\gamma \ge ||A\mathbf{x}||_{\infty}$ , then the dithered measurements sign $(\langle \mathbf{a}_i, \mathbf{x} \rangle + \tau_i)$  contain most scale information on  $\langle \mathbf{a}_i, \mathbf{x} \rangle$ . By Cauchy–Schwartz inequality, it is not hard to obtain



**Fig. 2** Recovery errors obtained with different values of the constant *c* in (25). **a** shows the results for signals of different lengths when SR = 20, and **b** shows the results under different sampling rates when N = 1000

$$\|A\boldsymbol{x}\|_{\infty} = \max_{i \in [m]} \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \leq \max_{i \in [m]} \|\boldsymbol{a}_i\|_2 \|\boldsymbol{x}\|_2.$$

Since we assume  $\|\mathbf{x}\|_2 = 1$  in this paper, we suggest setting  $\gamma = \max_{i \in [m]} \|\mathbf{a}_i\|_2$  to generate uniformly distributed dithers  $\tau$ .

## 5 Simulations

This section initially conducts experiments on synthetic data to facilitate the selection of regularization parameters. Subsequently, we apply the proposed method to ECG signals, MRI data, and hyperspectral images (HSIs) for evaluating its effectiveness. All experiments are conducted using MATLAB (2016a) on a platform equipped with AMD Ryzen 7 5800 H 3.20-GHz CPU and 16 GB memory.

Practical analog-to-digital converters (ADCs) not only sample signals, but also quantize each measurement to a finite number of bits. The performance of ADCs is primarily limited by the quantizer, resulting in an exponential decrease in sampling rate as the resolution increases linearly [32]. This implies that higher resolution per measurement results in slower sampling rates and increased costs for the ADC. The conventional CS framework provides a solution to mitigate the quantization bottleneck by reducing the ADC sampling rate, while the 1-bit CS framework directly addresses this issue in the quantization domain by decreasing the number of bits per measurement. In other words, while a sampling rate SR > 1 is not significant in conventional compressed sensing, it may prove highly practical in 1-bit (binary) systems that can acquire sign measurements at extra-low bit rates [5]. We define the recovery error (RE) as RE :=  $\|\hat{x}/\|\hat{x}\|_2 - x/\|x\|_2\|_2$ , where x denotes the original signal and  $\hat{x}$  denotes its estimate.



**Fig. 3** Recovery errors obtained with different values of factor k, where k = 1 corresponds to full random standard Gaussian measurement scheme

# 5.1 Parameter selection

Theorem 1 has indicated that the regularization parameters  $\lambda$  in (20) should be set as (25). However, determining the appropriate value for constant *c* remains unclear. To address this issue, we conducted a series of numerical experiments. Firstly, we generate *N*-dimensional signals with sparsity *s* randomly and evaluate the RE of the proposed method while keeping SR fixed at 20. We consider different values of *c* for N = 500, s = 5, N = 1000, s = 10; and N = 2000, s = 20. The average results from repeating the experiment for a hundred times are presented in Fig. 2a. Next, we generate 1000-dimensional signals with sparsity 10 randomly and evaluate the proposed method's REs at different SRs. In addition, we conducted 100 trials and presented the average results in Fig. 2b. It can be observed that c = 0.3 is a suitable setting for Eq. (20), as it yields relatively small REs across almost all cases. As the constant *c* does not depend on the structure of the signal, the parameter setting obtained can be applied universally. Therefore, we will maintain these settings throughout all remaining experiments in this article.

Next, we will examine the impact of factor k on recovery performance. We randomly generate signals with a length of N = 1000 with sparsity s = 10, and consider various values for k, including k = 1, k = 2, k = 4, k = 5, and k = 10. We repeat 100 trails and report the average REs at different SRs in Fig. 3. In addition to reducing the scale of the measurement matrix, we unexpectedly found from the experimental results that incorporating the semi-tensor product can enhance recovery accuracy. As revealed by our proof process, the semi-tensor product-based measurement is equivalent to the segmented measurement scheme of sparse signals. This implies that an excessively large value of k may result in measuring some zero vectors. To avoid the trivial situation, we suggest empirically selecting k such that it does not exceed the sparsity s of the signal.

#### 5.2 ECG signal recovery

In this section, we apply the proposed method to recover a 1000-length ECG signal segment randomly selected from the MIT-BIH Arrhythmia Database [33]. We compare the semi-tensor product-based measurement with other representative types of measurements in 1-bit CS, including Gaussian measurement, symmetric Bernoulli measurement,



Fig. 4 Recovery errors obtained by different methods at various sampling rates with noiseless measurements

and circulant Gaussian measurement. Denoted by *Gaussian* [7], we refer to a scenario where all entries of the measurement are independently and identically distributed according to the standard Gaussian distribution  $\mathcal{N}(0, 1)$ . Denoted by *Symm-Bernoulli* [19], we mean that each element is independently sampled from a Bernoulli distribution with equal probabilities of 1 and -1. Referred to as *Circ-Gaussian* [20], the measurement matrix was generated by shifting one element of a random row Gaussian vector to the right relative to the preceding row vector through rotation. Besides, we also compare the proposed algorithm with methods based on other sparsity regularization. For  $\ell_0$  regularization, we compare with a Hard Thresholding-based method [34], denoted as *HT*, which aims to solve a non-convex sparsity-constrained program. For weighted  $\ell_1$ -norm regularization, we compare our method with the Inverse Weights-based approach [35] (denotes as Inverse-WT) and the  $\ell_1$ -norm Shannon Entropy Weights-based method [35]

The quantitative results are plotted in Fig. 4 for different sampling rates. The reconstructed results of a fraction with a length of 250, obtained at SR = 40, are depicted in Fig. 5. The measurements based on semi-tensor product have consistently demonstrated superior performance in 1-bit CS compared to commonly used measurements, both in terms of indicators and recovery outcomes, encompassing reconstruction performance and execution time. To evaluate the robustness of the proposed approach, we introduced random bit flips to the measurements by flipping each entry of binary measurements with a probability of 0.2. The quantitative results are plotted in Fig. 6 for different sampling rates, and the reconstructed results of a fraction with a length of 250 at SR = 40 are illustrated in Fig. 7. Although the reconstructed results are slightly degraded compared to noise-free situation, our approach still demonstrates a significant advantage over its competitors. As is shown, though weighted  $\ell_1$ -norm based methods, Inverse-WT and  $\ell_1$ -SE-WT, obtain better reconstruction performance than most other  $\ell_1$ -norm based algorithms, it is still inferior to our method. We believe that our advantage primarily stems from the utilization of semi-tensor product-based measurements. If other sparse



Fig. 5 Reconstructed results obtained by different schemes from noiseless measurements (SR of 40)



Fig. 6 Recovery errors obtained by different methods at various sampling rates with noisy measurements

regularization methods are incorporated, it is possible to achieve improved recovery outcomes, which necessitates further exploration in future discussions.

# 5.3 MRI data recovery

We also utilize the proposed method to conduct further test on MRI data, which usually exhibit sparsity under discrete cosine transform (DCT) or wavelet basis. We used two different datasets-Data A and Data B. Data A is a cardiac cine MRI dataset distributed by the 2013 ISMRM Recon Challenge committee,<sup>1</sup> which was collected using a 2D cine breathheld bSSFP sequence with 32-channel cardiac receiver coils. Data B is a publicly available myocardial perfusion MRI dataset from [36], which was acquired using a saturation recovery FLASH sequence. We introduce random bit flips to the measurements by flipping each entry of binary measurements with a probability of 0.1. To facilitate the implementation of the measurement process and enhance solving speed, we employed a block-based sampling scheme. Empirically, we partitioned the original image into  $32 \times 32$ -sized parts and subsequently perform DCT on each part to obtain sparse signals through vectorization. The reconstruct is conducted in blocks. Readers concerning on further information on block-based samplings may refer to [37, 38]. Two quantitative indicators, peak signal-tonoise ratio (PSNR) and structural similarity (SSIM) [39], and execution times (in seconds) were calculated at SRs of 5, 10, 20, and 40 and are presented in Table 1, where the bolded data represent the optimal outcomes. Figure 8 displays the twentieth frame along with the reconstructed results from various methods for each of the MRI datasets obtained at SR = 40. It is evident, both quantitatively and visually, that our recovery method achieves better reconstruction performance. In addition, the reduced computational complexity of matrix and signal multiplication operations in semi-tensor product-based measurement results in significantly shorter execution time for our algorithm compared to other algorithms. This aligns with the MRI's requirement for fast sampling and reconstruction speed.

# 5.4 Hyperspectral images recovery

We also apply the proposed method to perform compressed sensing of HSIs. We select two typical HSI datasets, pure HYDICE Urbanpart and AVIRIS Moffett Field, both of



Fig. 7 Reconstructed results obtained by different schemes from noisy measurements (SR of 40)

	מ הור לממוי					ממומ מר מוווכו כ		410.0				
Recovery method	SR = 5			SR = 10			SR = 20			SR = 40		
	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
Data A												
Gaussian	27.04	0.643	2.983	28.99	0.722	5.871	30.74	0.788	11.685	32.34	0.836	24.456
Sym-Bern	26.26	0.619	5.899	27.87	0.693	11.894	29.35	0.751	24.522	30.89	0.788	48.548
Circ-Gauss	27.22	0.651	5.916	28.94	0.728	9.656	30.69	0.792	19.234	32.38	0.845	36.902
HT	26.61	0.614	3.663	29.01	0.706	6.097	31.03	0.781	12.877	32.35	0.823	24.326
Inverse-WT	24.21	0.521	3.394	26.11	0.606	7.523	28.19	0.694	13.302	29.98	0.778	27.711
ℓ <sub>T</sub> SE-WT	27.29	0.652	3.226	29.30	0.729	6.775	30.95	0.790	13.302	32.56	0.836	26.526
Ours	28.84	0.718	0.612	30.71	0.783	1.322	32.25	0.837	2.666	33.57	0.870	5.396
Data B												
Gaussian	25.24	0.518	1.802	27.02	0.607	4.033	28.38	0.677	7.491	29.83	0.740	14.442
Sym-Bern	24.53	0.487	3.528	25.94	0.563	6.991	27.00	0.626	13.282	28.61	0.708	23.763
Circ-Gauss	25.39	0.519	2.813	26.90	0.609	5.694	28.36	0.680	10.209	29.70	0.754	21.537
HT	24.27	0.486	1.639	26.56	0.586	4.084	28.39	0.665	7.172	29.76	0.715	14.234
Inverse-WT	23.11	0.444	1.933	25.08	0.556	3.715	26.55	0.654	8.411	28.85	0.737	16.107
ℓ <sub>T</sub> SE-WT	23.80	0.457	1.866	26.29	0.567	4.739	28.24	0.662	7.199	29.64	0.722	15.421
Ours	26.68	0.592	0.363	28.09	0.668	0.890	29.34	0.733	1.432	30.65	0.801	2.894

S	
U.	
片	
2	
σ	2
Ē	
Ω	
2	
ਛੋ	
ŝ	
+	
Ū,	
ົດ	
Æ	
÷=	
0	
片	
10	
ਯੋ	
ō	
Ľ Ľ	
$\geq$	
2	
Z	
0	
Ş	
ğ	
ğ	
노	
ā	
Ē	
2	
Ę	
5	
2	
U,	
Ŧ	
÷	
~	
6	
~	
Q	
Ψ	
m,	
Ħ	
H	
5	
8	
÷=	
2	
.'=	
d)	
~	
÷	
- m	
- E	
5	
ਯ	
σ	
0	
he a	
the a	
of the a	
of the a	
n of the a	
on of the a	
ison of the a	
arison of the a	
parison of the d	
nparison of the d	
mparison of the a	
omparison of the a	
Comparison of the a	
Comparison of the a	
1 Comparison of the a	
e 1 Comparison of the d	



**Fig. 8** Visual display of the reconstructed results obtained by different approaches on frame 20 of Data A (upper), and Data B (lower) at SR = 40. **a** Original image, **b** Gaussian, **c** Sym-Bern, **d** Circ-Gauss, **e** HT, **f** Inverse-WT, **g**  $\ell_{T}$ SE-WT, **h** Ours

which are sized to  $256 \times 256 \times 128$  for our experiments. To facilitate the implementation of the measurement process, we also employed a block-based sampling scheme. Empirically, we partitioned the original data into  $16 \times 16 \times 16$ -sized blocks and subsequently perform three-dimensional DCT on each block to obtain sparse signals through vectorization. The recovery process is conducted in a block-wise manner. We evaluate the recovery performance by calculating the PSNR, SSIM values and execution times at SRs of 10 and 20. The quantitative indices are reported in Table 2, where the bolded data represent the optimal outcomes, and the images of all methods with bands 6-45-155 as R-G-B at SR = 20 are depicted in Fig. 9. The superiority of both  $\ell_1$ -SE-WT and our method in terms of recovery performance is evident, with  $\ell_1$ -SE-WT exhibiting a slightly better performance than ours. However, our method exhibits a significant advantage in execution time compared to the former.

# 6 Conclusion

This paper proposes a model for data compression and reconstruction from one-bit measurements using semi-tensor product. Unlike traditional measurements, this new method does not require the number of columns in the sensing matrix to be equal to the length of the underlying signal, thus breaking the dimension matching condition. Leveraging the theory of semi-tensor product, this approach is capable of compressing high-dimensional signals with lower-dimensional sensing matrices. Theoretically, we provide an upper bound for the recovery error of this method. We also apply the proposed method to the compressed sensing of real-world data such as ECG signals, MRI data, and HSIs to show the low time complexity of our approach.

The selection of an optimal value for k may be influenced by the characteristics of the underlying signal, and this article presents empirical approaches to aid in its determination. In future research, we intend to explore a theoretical solution to this issue. Additionally, the current compressed sensing MRI method relies on a sampling matrix derived from partial Fourier transform. However, due to the limitations imposed by one-bit quantization, this paper proposes the utilization of a semi-tensor product measurement based on Gaussian distribution, which distinguishes it from real-world scenarios. Further research is needed on how to apply 1-bit quantization and semi-tensor product-based measurements to the real application scenarios of MRI in future. Additionally, there is still much to be

Recovery method	HYDICE ui	rbanpart					AVIRIS Mc	offett field				
	SR = 10			SR = 20			SR = 10			SR = 20		
	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
Gaussian	27.39	0.798	1264.38	28.88	0.850	2667.12	34.58	0.821	1000.40	35.88	0.866	1904.12
Sym-Bern	27.41	0.794	3012.19	28.89	0.849	6022.14	32.10	0.706	1957.48	32.90	0.747	3843.99
Circ-Gauss	27.48	0.799	1535.26	28.92	0.851	2907.79	34.35	0.820	1393.31	35.89	0.862	2882.96
НТ	27.35	0.773	1268.63	28.22	0.815	2458.66	34.41	0.781	999.37	35.37	0.823	1954.25
Inverse-WT	26.62	0.731	1404.11	27.63	0.790	2868.52	31.89	0.748	994.56	32.52	0.823	1995.26
ℓ <sub>T</sub> SE-WT	29.75	0.859	1426.32	31.44	0.898	2536.04	36.68	0.914	1020.42	38.32	0.940	2090.91
Ours	29.87	0.860	219.90	31.02	0.892	447.74	36.37	0.911	188.93	37.96	0.936	409.24

rates
npling
rent sar
at diffe
on HSIs
nethods c
different r
ed by (
obtain
indices
titative
e quant
n of th
ompariso
<b>7</b>
Table



Fig. 9 Recovered images of all competing methods with bands 6-45-155 as R-G-B at SR = 20. The upper are results for HYDICE Urbanpart and the lower are results for AVIRIS Moffett Field. **a** Original image, **b** Gaussian, **c** Sym-Bern, **d** Circ-Gauss, **e** HT, **f** Inverse-WT, **g** ℓ<sub>T</sub>SE-WT, **h** Ours

investigated regarding the application of semi-tensor product-based measurement in highdimensional scenarios, particularly with respect to compressed sensing of low-rank matrices and tensors from binary measurements.

#### Acknowledgements

Not applicable.

#### Author contributions

JYH proposed the framework of the whole method; performed the simulations, analysis and interpretation of the results; and drafted the manuscript. XLL has participated in the conception and revision of this research. The authors read and approved the final manuscript.

#### Funding

This research was supported in part by the National Natural Science Foundation of China (Grant Nos. 12201505, 12301594), in part by the Sichuan Science and Technology Program (Grant No. 2023NSFSC0060), and in part by the Initiative Projects for Ph.D. in China West Normal University (Grant No. 22kE030).

## Availability of data and materials

Please contact any of the authors for data and materials.

#### Declarations

#### **Competing interests**

The authors declare that they have no competing interests.

## Received: 18 June 2023 Accepted: 14 October 2023 Published online: 03 November 2023

#### References

- H.F. Shen, X.H. Li, L.P. Zhang, D.C. Tao, C. Zeng, Compressed sensing-based inpainting of aqua moderate resolution imaging spectroradiometer Band 6 using adaptive spectrum-weighted sparse Bayesian dictionary learning. IEEE Trans. Geosci. Remote Sens. 51(2), 894–906 (2014)
- M. Lustig, D.L. Donoho, J.M. Santos, J.M. Pauly, Compressed sensing MRI. IEEE Signal Process. Mag. 25(2), 72–82 (2008)
- S. Dirksen, S. Mendelson, Non-Gaussian hyperplane tessellations and robust one-bit compressed sensing. J. Eur. Math. Soc. 23, 2913–2947 (2021)
- P.T. Boufounos, R.G. Baraniuk, 1-bit compressive sensing, in: Proceeding of the 43rd Asilomar Conference on Conference on Signals, Systems and. Computers (IEEE, Pacific Grove, 2008), pp. 16–21
- L. Jacques, J.N. Laska, P.T. Boufounos, R.G. Baraniuk, Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. IEEE Trans. Inf. Theory 59(4), 2082–2102 (2013)
- Y. Plan, R. Vershynin, Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach. IEEE Trans. Inf. Theory 59(1), 482–494 (2013)
- 7. L. J. Zhang, J. F. Yi, R. Jin, Efficient algorithms for robust one-bit compressive sensing, in: *Proceedings of the International Conference on International Conference on Machine Learning* (IEEE, Beijing, 2014), pp. 820–828
- 8. J.Y. Hou, J.J. Wang, F. Zhang, J.W. Huang, One-bit compressed sensing via  $\ell_p(0 -minimization method.$ Inverse Probl.**36**(5), 055005 (2020)
- 9. Z. Wang, F.L. Liu, Y.H. Jia, H.Y. Yang, Y.Y. Guo, One bit compressive sensing with off-grid targets. Digit. Signal Prog. 115, 103008 (2021)
- E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory 52(2), 489–509 (2006)

- R.E. Carrillo, A.B. Ramirez, G.R. Arce, K.E. Barner, B.M. Sadler, Robust compressive sensing of sparse signals: a review. EURASIP J. Adv. Signal Process. 108, 1–17 (2016)
- J. Mo, P. Schniter, N.G. Prelcic, R.W. Heath, Channel estimation in millimeter wave MIMO systems with one-bit quantization, in: *Proceedings of the 48th Asilomar Conference on Signals, System and Computer* (IEEE, Pacific Grove, 2014), pp. 957–961
- Y. Noh, S. Hong, Compressed sensing based active user detection in MIMO systems with one-bit ADC. IEEE Trans. Veh. Technol. 72(1), 1313–1317 (2022)
- 14. J.P. Xiong, Q.H. Tang, 1-bit compressive data gathering for wireless sensor networks. J. Sens. 2014(7), 177–183 (2014)
- X. Dong, Y. Zhang, A MAP approach for 1-bit compressive sensing in synthetic aperture radar imaging. IEEE Geosci. Remote Sens. Lett. 12(6), 1237–1241 (2015)
- J.Y. Hou, F. Zhang, H.Q. Qiu, J.J. Wang, Y. Wang, Y.D. Meng, Robust low-tubal-rank tensor recovery from binary measurements. IEEE Trans. Pattern Anal. Mach. Intell. 44(8), 4355–4373 (2022)
- E.J. Candes, M.B. Wakin, S.P. Boyd, Enhancing sparsity by reweighted ℓ<sub>1</sub>minimization. J. Fourier Anal. Appl. 14, 877–905 (2008)
- 18. A. Ai, A. Lapanowski, Y. Plan, R. Vershynin, One-bit compressed sensing with non-Gaussian measurements. Linear Algebra Appl. 441, 222–239 (2014)
- S. Dirksen, S. Mendelson, Non-Gaussian hyperplane tessellations and robust one-bit compressed sensing. J. Eur. Math. Soc. 23(9), 2913–2947 (2021)
- S. Dirksen, H.C. Jung, H. Rauhut, One-bit compressed sensing with partial Gaussian circulant matrices. Inf. Inference 9(3), 601–626 (2020)
- Z.Q. Liu, S. Ghosh, J. Scarlett, Robust 1-bit compressive sensing with partial Gaussian circulant matrices and generative priors, in: 2021 IEEE Information Theory Workshop (ITW) (IEEE, Kanazawa, 2021), pp. 1–6
- S. Dirksen, S. Mendelson, Robust one-bit compressed sensing with partial circulant matrices. Ann. Appl. Probab. 33(3), 1874–1903 (2023)
- 23. D.Z. Cheng, Semi-tensor product of matrices and its application to Morgen's problem. Sci. China Ser. 44(3), 195–212 (2001)
- 24. D.Z. Cheng, L.J. Zhang, On semi-tensor product of matrices and its applications. Acta Math. Appl. Sin. 19(2), 219–228 (2003)
- 25. D.Z. Cheng, H.S. Qi, A.C. Xue, A survey on semi-tensor product of matrices. J. Syst. Sci. Complex 20(2), 304–322 (2007)
- D.W. Zhao, H.P. Peng, L.X. Li, S.L. Hui, Y.X. Yang, Novel way to research nonlinear feedback shift register. Sci. China Ser. 57(9), 1–14 (2014)
- Y.H. Wu, T.L. Shen, Policy iteration approach to control residual gas fraction in IC engines under the framework of stochastic logical dynamics. IEEE Trans. Control Syst. Technol. 25, 1100–1107 (2017)
- K.Z. Zhang, L.J. Zhang, S.S. Mou, An application of invertibility of Boolean control networks to the control of the mammalian cell cycle. IEEE ACM Trans. Comput. Biol. Bioinform. 14, 225–229 (2017)
- 29. D. Xie, H.P. Peng, L.X. Li, Y.X. Yang, Semi-tensor compressed sensing. Digit. Signal Prog. 58, 85–92 (2016)
- Y. Plan, R. Vershynin, Dimension reduction by random hyperplane tessellations. Discrete Comput. Geom. 51(2), 438–461 (2014)
- J.R. Chen, C.L. Wang, M.K. Ng, D. Wang, High dimensional statistical estimation under uniformly dithered one-bit quantization. IEEE Trans. Inf. Theory 69(8), 5151–5187 (2023)
- B. Le, T.W. Rondeau, J.H. Reed, C.W. Bostian, Analog-to-digital converters. IEEE Signal Process. Mag. 22(6), 69–77 (2005)
- 33. G.B. Moody, R.G. Mark, The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag. 20(3), 45–50 (2001)
- J. Shen, One-bit compressed sensing via one-shot hard thresholding, in: 36-th Conference on Uncertainty in Artificial Intelligence (PMLR, Toronto, 2020), pp. 510–519
- P. Xiao, B. Zhao, Robust one-bit compressive sensing with weighted ℓ<sub>T</sub>norm minimization. Signal Process. 164, 380–385 (2019)
- 36. S.G. Lingala, M. Jacob, Blind compressive sensing dynamic MRI. IEEE Trans. Med. Imaging 32(6), 1132–1145 (2013)
- L. Gan, Block compressed sensing of natural images, in: 2007 15th International Conference on Digital Signal Processing (IEEE, Cardiff, 2007), pp. 403–406
- J.E. Fowler, S. Mun, E.W. Tramel, Block-based compressed sensing of images and video. Found. Trends Signal Process. 4(4), 297–416 (2012)
- Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13(4), 600–612 (2004)

#### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.