Digital Signal Processing 58 (2016) 85–92

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Semi-tensor compressed sensing

Dong Xie [a,b], Haipeng Peng [a,b,*], Lixiang Li [a,b], Yixian Yang [a,b]

[a] Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, 100876 Beijing, PR China

[b] National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, 100876 Beijing, PR China

ARTICLE INFO

Article history:
Available online 1 August 2016

Keywords:
Compressed sensing
Semi-tensor product
Restricted isometry property
Spark
Coherence

ABSTRACT

In compressed sensing (CS), sparse or compressible signals can be reconstructed with fewer samples than the Nyquist–Shannon theorem requires. Over the past ten years, CS has developed into a relatively mature theory and this brand-new technique has been widely used in many fields such as image processing, wireless communication and medical imaging. In this paper, we propose a new model for signal compression and reconstruction based on semi-tensor product, called STP-CS, which is a generalization of traditional CS. Like traditional CS, we investigate some reconstruction conditions of STP-CS in terms of the spark, the coherence and the restricted isometry property (RIP). The experimental results show that STP-CS has the flexibility to choose a lower-dimensional sensing matrix for signal compression and reconstruction.

1. Introduction

As one of emerging research fields in signal processing, compressed sensing (CS) has attracted considerable attention in recent ten years, because it can reconstruct sparse signals from very few incoherent measurements [1–3]. The standard framework of CS is a special case of underdetermined linear equations

$$\mathbf{y} = \mathbf{\Phi}\boldsymbol{\theta}, \tag{1}$$

where $\mathbf{\Phi}$ is an $m \times n$ matrix with $m < n$. Eq. (1) reflects that the original n-dimensional signal $\boldsymbol{\theta}$ is compressed into an m-dimensional vector $\mathbf{y}$. However, it is impossible to directly reconstruct the original signal $\boldsymbol{\theta}$ from $\mathbf{y}$ because there are infinitely many solutions for Eq. (1). Fortunately, some natural signals can be represented using only a few non-zero coefficients in a suitable basis or dictionary [4,5]. Namely,

$$\boldsymbol{\theta} = \mathbf{\Psi}\mathbf{x}, \tag{2}$$

where $\mathbf{\Psi}$ is a sparsifying dictionary and $\mathbf{x}$ is a sparse vector. We say a vector $\mathbf{x}$ is k-sparse, denoted by $\mathbf{x} \in \sum_k$, if it has at most $k \ll n$ nonzero entries. Thus, we have

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x} = \mathbf{A}\mathbf{x}, \tag{3}$$

where $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$ is regarded as the sensing matrix in CS. For a matrix $\mathbf{A}$, the *spark* of $\mathbf{A}$ is the smallest number of columns of $\mathbf{A}$ that are linearly dependent. Donoho and Elad showed that if $spark(\mathbf{A}) > 2k$, then for each measurement vector $\mathbf{y} \in \mathbb{R}^m$ there exists at most one signal $\mathbf{x} \in \sum_k$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$ [6]. Because finding sparse solutions to underdetermined systems of linear equations is in general NP-hard [1], the sensing matrix $\mathbf{A}$ satisfying such condition is impracticable. Thus, how to construct sensing matrices becomes one of the most important research directions in the field of CS. Fortunately, Candès and Tao proposed a typical criterion for constructing sensing matrix, called restricted isometry property (RIP) [1,7]. A matrix $\mathbf{A}$ satisfies the RIP of order k if there exists a $\delta_k^{\mathbf{A}} \in (0, 1)$ such that

$$(1 - \delta_k^{\mathbf{A}})\|\mathbf{x}\|_2^2 \le \|\mathbf{A}\mathbf{x}\|_2^2 \le (1 + \delta_k^{\mathbf{A}})\|\mathbf{x}\|_2^2 \tag{4}$$

holds for all $\mathbf{x} \in \sum_k$. Random matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ with entries drawn from a Gaussian distribution, a Bernoulli distribution or more generally a sub-Gaussian distribution have $spark(\mathbf{A}) = m + 1$ with high probability [8]. And such matrices satisfy the RIP with overwhelming probability, providing that $m = O((\delta_k^{\mathbf{A}})^{-2} k \log \frac{n}{k})$ [8,9]. However, such random constructions are often not feasible for real-word applications because some sensing devices with little storage resources are impossible to store all the entries of the sensing matrix when the size of the matrix is very large. To reduce the storage burden, some deterministic approaches for constructing sensing matrices have been proposed, such as structurally subsampled matrices [8], Toeplitz sensing matrices [10], and chaotic sensing matrices [11].

* Corresponding author at: Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, 100876 Beijing, PR China.

E-mail address: penghaipeng@bupt.edu.cn (H. Peng).

http://dx.doi.org/10.1016/j.dsp.2016.07.003
1051-2004/

While the RIP guarantees recovery of k-sparse signals, verifying that a general matrix $\mathbf{A}$ satisfies the RIP has a combinatorial computational complexity, since one must essentially consider $\binom{n}{k}$ submatrices. In many cases it is preferable to use properties of $\mathbf{A}$ that are easily computable to provide more concrete recovery guarantees. The *coherence* of a matrix is one such property. The coherence $\mu(\mathbf{A})$ of a matrix $\mathbf{A}$ is the largest absolute normalized inner product between any two columns of $\mathbf{A}$:

$$\mu(\mathbf{A}) = \max_{1\le i\ne j\le n} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j\rangle|}{\|\mathbf{a}_i\|_2\|\mathbf{a}_j\|_2}, \tag{5}$$

where $\mathbf{a}_i$ denotes the i-th column of $\mathbf{A}$. It is easy to show that $\mu(\mathbf{A}) \in [\sqrt{\frac{n-m}{m(n-1)}}, 1]$ [12,13].

A straightforward approach to obtain the original k-sparse vector $\mathbf{x}$ from Eq. (3) can be viewed as the optimization problem of

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{A}\mathbf{x} = \mathbf{y}, \tag{6}$$

which is called l_0 optimization problem. Since the convex property of l_1 norm, a classic method used in compressed sensing is to replace $\|\mathbf{x}\|_0$ with $\|\mathbf{x}\|_1$, i.e.,

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{A}\mathbf{x} = \mathbf{y}. \tag{7}$$

If $\delta_{2k}^{\mathbf{A}} < \sqrt{2} - 1$, the solution to the l_1 problem is that of the l_0 problem [7]. In addition to l_1-based algorithms, many greedy algorithms, such as orthogonal matching pursuit [14], StOMP [15] and CoSaMP [16], have been proposed for sparse signal reconstruction.

Based on traditional CS, Gan introduced the concept of block compressed sensing (BCS) for natural images [17], where image acquisition is performed in a block-by-block manner through the same sensing matrix. In BCS, an image is first divided into $B \times B$ non-overlapping blocks and then acquired using an appropriately sized sensing matrix. Namely, suppose that $\mathbf{x}_i$ is a vector representing the i-th block of the input image. The corresponding $\mathbf{y}_i$ is then $\mathbf{y}_i = \mathbf{A}_B\mathbf{x}_i$, where $\mathbf{A}_B$ is an $m_B \times B^2$ sensing matrix with $m_B = \lfloor \frac{mB^2}{n} \rfloor$. Because of lightweight reconstruction complexity and lower storage overhead, BCS has been widely used in various multiple-image scenarios, such as video and multi-view imagery [18,19].

In this paper, we propose a new model for signal compression and reconstruction based on semi-tensor product (STP), called STP-CS, which can be viewed as a generalization of tradition CS. This new model breaks the dimension matching condition of the traditional CS model in Eq. (3), i.e., the number of columns of the sensing matrix $\mathbf{A}$ must be equal to the length of the signal $\mathbf{x}$. Under this brand-new model, we first analyze the uniqueness of sparse solution in terms of *spark* and *coherence* from a theoretical point of view. Subsequently, we find that the RIP constant of order k in traditional CS is equal to that in our proposed STP-CS model. It implies that some classical sensing matrices in CS, such as Gaussian, Bernoulli, and Chaotic sensing matrix, can also be used in STP-CS. In addition, we give the exact reconstruction condition on the sensing matrix, which is sufficient for a variety of algorithms to be able to successfully reconstruct the original sparse signal from measurements. At last, the experiment results prove the validity of our theory analysis. Compared to several previous methods for signal compression and reconstruction, the main advantages of our proposed STP-CS can be summarized as follows:

- *Low-storage overhead*. With the help of the semi-tensor product theory, STP-CS can compress high-dimensional signals using lower-dimensional sensing matrices. As a generalization of CS, STP-CS has the flexibility to choose a lower-dimensional sensing matrix. In addition, the experimental results show that the storage overhead of sensing matrices in STP-CS is smaller than that in BCS when the size of each block is not too small.
- *Parallel reconstruction*. The reconstruction algorithm in STP-CS can be implemented in a parallel fashion. The theoretical analysis indicates that a reconstruction instance in STP-CS can be transformed into some independent reconstruction instances in CS. Thus, it can simultaneously perform the reconstruction phase among multiple CS decoders and will lead to the reduction of the total reconstruction time.

The rest of this paper is organized as follows. Section 2 recalls some basic background knowledge of semi-tensor product. We will first introduce the STP-CS model and then give our theoretical results in Section 3. In Section 4, some experiments are carried out to simulate the performance of STP-CS. Section 5 provides a comparison among traditional CS, BCS, and our proposed STP-CS. Last we conclude this paper in Section 6.

2. Semi-tensor product

The concept of STP of matrices was proposed by Cheng et al., which is a generalization of conventional matrix product [20–23]. This novel theory is able to perform matrix multiplication when two matrices do not meet the dimension matching condition. STP has received great attention in a variety of areas, including multilinear algebra [26], game theory [25], and boolean networks [24].

Definition 1. [22] Let $\mathbf{x}$ be a row vector of dimension np, and $\mathbf{y}$ be a column vector with dimension p. Split $\mathbf{x}$ into p equal blocks, named $\mathbf{x}^1, \cdots, \mathbf{x}^p$, which are $1 \times n$ vectors. Define the STP, denoted by $\ltimes$, as

$$\begin{cases} \mathbf{x} \ltimes \mathbf{y} = \sum_{i=1}^{p} \mathbf{x}^i y_i \in \mathbb{R}^{1\times n}; \\ \mathbf{y}^T \ltimes \mathbf{x}^T = \sum_{i=1}^{p} y_i (\mathbf{x}^i)^T \in \mathbb{R}^{n\times 1}. \end{cases} \tag{8}$$

Definition 2. [22] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{p\times q}$. If either n is a factor of p or p is a factor of n, then we define the STP of $\mathbf{A}$ and $\mathbf{B}$ as the following: $\mathbf{C}$ consists of $m \times q$ blocks as $\mathbf{C} = (\mathbf{c}_{ij})$ and each block is

$$\mathbf{c}_{ij} = \mathbf{a}_i \ltimes \mathbf{b}^j, \ i = 1, 2, \cdots, m, \ j = 1, 2, \cdots, q, \tag{9}$$

where $\mathbf{a}_i$ is the i-th row of $\mathbf{A}$ and $\mathbf{b}^j$ is the j-th column of $\mathbf{B}$.

Given two matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{p\times q}$, the *Kronecker product* between them is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix} \in \mathbb{R}^{mp\times nq}. \tag{10}$$

Equivalently, we can also define the STP using Kronecker product.

Definition 3. [23] The STP of two matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{p\times q}$ is defined as

$$\mathbf{A} \ltimes \mathbf{B} = (\mathbf{A} \otimes \mathbf{I}_{t/n})(\mathbf{B} \otimes \mathbf{I}_{t/p}), \tag{11}$$

where t is the least common multiple of n and p, i.e., $t = lcm(n, p)$.

Remark 1. Note that $(\mathbf{A} \otimes \mathbf{I}_{t/n}) \in \mathbb{R}^{mt/n\times t}$ and $(\mathbf{B} \otimes \mathbf{I}_{t/p}) \in \mathbb{R}^{t\times qt/p}$, so $\mathbf{A} \ltimes \mathbf{B} \in \mathbb{R}^{mt/n\times qt/p}$.

Remark 2. If $p = n$, then $\mathbf{A} \ltimes \mathbf{B} = (\mathbf{A} \otimes \mathbf{I}_1)(\mathbf{B} \otimes \mathbf{I}_1) = \mathbf{AB}$. It is the standard matrix product.

3. Semi-tensor compressed sensing

3.1. Model

For a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ with $m < n$ and a signal $\mathbf{x} \in \mathbb{R}^p$, the standard model of STP-CS is defined as

$$\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}, \tag{12}$$

where $\mathbf{y} \in \mathbb{R}^{mt/n\times t/p}$ and $t = lcm(n, p)$. Evidently $t \geq p$. In order to reduce the number of measurements, we can choose the first case in Definition 2, i.e., $n \mid p$. Thus, $t = p$ and $\mathbf{y} \in \mathbb{R}^{mp/n}$.

For a p-dimensional signal $\mathbf{x}$, the number of columns of sensing matrices in traditional CS must be equal to p. However, the STP-CS model enables sensing matrices with n columns to compress the signals successfully, where n is a factor of p. If $n \neq p$, the operation of Eq. (3) can not be performed because it does not meet the dimension matching condition. When $p = n$, the STP-CS degenerates into the traditional CS. Therefore STP-CS can be viewed as a generalization of tradition CS.

Note that natural signals can be represented using only a few non-zero coefficients in a suitable basis or dictionary. For convenience, we consider $\mathbf{x}$ to be k-sparse in what follows. Next, we give our related theoretical results about such novel data compression and reconstruction model.

3.2. Spark

In traditional CS, the *spark* provides a complete characterization of when sparse recovery is possible. Donoho and Elad introduced a straightforward guarantee:

Lemma 1. *[6] If spark*$(\mathbf{A}) > 2k$, *then for each measurement vector* $\mathbf{y} \in \mathbb{R}^m$ *there exists at most one signal* $\mathbf{x} \in \sum_k$ *such that* $\mathbf{y} = \mathbf{A}\mathbf{x}$.

Theorem 1. *If* $lcm(n, p) = p$, *then spark*$(\mathbf{A})$ =*spark*$(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}})$.

Proof. If *spark*$(\mathbf{A}) = h$, then there exists a vector $\mathbf{x}$ with h nonzero entries such that $\mathbf{A}\mathbf{x} = \mathbf{0}$. Let $\triangle$ be the subset of indices of nonzero entries in $\mathbf{x}$. Evidently, $|\triangle| = h$. We have $\sum_{i\in\triangle} \mathbf{a}_i x_i = \mathbf{0}$, where $\mathbf{a}_i$ is the i-th column of $\mathbf{A}$ and x_i is the i-th entry of $\mathbf{x}$. From the construction of $\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}$, we naturally get $\sum_{i\in\triangle} \mathbf{c}_{i\cdot\frac{p}{n}} x_i = \mathbf{0}$, where $\mathbf{c}_i$ is the i-th column of $\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}$. Thus,

$$spark(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) \leq |\triangle| = spark(\mathbf{A}). \tag{13}$$

If $spark(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) = h$, then there also exists a vector $\mathbf{x}$ with h nonzero entries such that $(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}})\mathbf{x} = \mathbf{0}$. Similarly, let $\triangle$ be the subset of indices of nonzero entries in $\mathbf{x}$ and $|\triangle| = h$. We easily obtain

$$\sum_{i\in\triangle} \mathbf{c}_i x_i = \mathbf{0}, \tag{14}$$

where $\mathbf{c}_i$ is the i-th column of $\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}$. For $\forall i \in \triangle$, there exists at least one index $j \in \triangle$ such that

$$\frac{p}{n} \mid i - j. \tag{15}$$

If not, then $\sum_{i\in\triangle} \mathbf{c}_i x_i$ must be not equal to 0. From (15), we can see that the set $\triangle$ has an equivalence relation defined on its elements. Thus, the set $\triangle$ can be partitioned into a finite number of classes $\triangle_s$ such that for any $i, j \in \triangle_s$, Eq. (15) holds. Namely,

$$\triangle = \sum_{k=1}^{r} \triangle_s, \tag{16}$$

where r $(1 \leq r \leq [\frac{h}{2}])$ is a positive integer. Note that $h \geq |\triangle_s| \geq 2$. From (14) and (16), we have

$$\sum_{i\in\triangle_1} \mathbf{c}_i x_i + \sum_{i\in\triangle_2} \mathbf{c}_i x_i + \cdots, + \sum_{i\in\triangle_r} \mathbf{c}_i x_i = \mathbf{0}. \tag{17}$$

For any two different equivalence classes $\triangle_{s_1}$ and $\triangle_{s_2}$, $\sum_{i\in\triangle_{s_1}} \mathbf{c}_i x_i$ and $\sum_{i\in\triangle_{s_2}} \mathbf{c}_i x_i$ must be linearly independent. From (17) we have $\sum_{i\in\triangle_s} \mathbf{c}_i x_i = \mathbf{0}$ for any s. That is to say, $\sum_{i\in\triangle_s} \mathbf{a}_{\lceil \frac{in}{p} \rceil} x_i = \mathbf{0}$. Thus,

$$spark(\mathbf{A}) \leq \min_{1\leq s\leq r} |\triangle_s| \leq h = spark(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}). \tag{18}$$

From (13) and (18), we complete our proof. □

Note that $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x} = (\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}})\mathbf{x}$. From Theorem 1 and Lemma 1, we obtain the following straightforward corollary.

Corollary 1. *If spark*$(\mathbf{A}) > 2k$, *then for each measurement vector* $\mathbf{y} \in \mathbb{R}^{m\cdot\frac{p}{n}}$ *there exists at most one* k*-sparse signal* $\mathbf{x} \in \mathbb{R}^p$ *such that* $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}$.

Theorem 2. *For each measurement* $\mathbf{y} \in \mathbb{R}^{m\cdot\frac{p}{n}}$, *if there exists at most one signal* $\mathbf{x} \in \sum_k$ *such that* $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}$, *then spark*$(\mathbf{A}) > 2[\frac{kn}{p}]$, *where* $[\cdot]$ *is the floor function.*

Proof. We prove this theorem by contradiction. Let $\gamma = 2[kn/p]$ and we assume that $spark(\mathbf{A}) \leq \gamma$. Then there exists a nonzero vector $\mathbf{x} \in \sum_\gamma$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$. Without loss of generality, we suppose that $\sum_{i=1}^{\gamma} \mathbf{a}_i x_i = 0$. We extend each m-dimensional vector $\mathbf{a}_i$ into an mp/n-dimensional vector $\mathbf{b}_i$ via inserting many 0's after each element of $\mathbf{a}_i$. Namely,

$$\mathbf{b}_i = (a_{i1}, \overbrace{0, \cdots, 0}^{\frac{p}{n}-1}, a_{i2}, \underbrace{0, \cdots, 0}_{\frac{p}{n}-1}, \cdots, a_{im}, \overbrace{0, \cdots, 0}^{\frac{p}{n}-1})^T. \tag{19}$$

We obtain $\mathbf{b}_i^j$ through moving the last j element of $\mathbf{b}_i$ to the beginning, where $j = 0, 1, 2, \cdots, \frac{p}{n} - 1$. Obviously $\mathbf{b}_i = \mathbf{b}_i^0$. For any j, we have

$$\sum_{i=1}^{\gamma} \mathbf{b}_i^j x_i = \mathbf{0}. \tag{20}$$

Note that

$$(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) = (\mathbf{b}_1^0, \mathbf{b}_1^1, \cdots, \mathbf{b}_1^{\frac{p}{n}-1}, \cdots, \mathbf{b}_n^0, \cdots, \mathbf{b}_n^{\frac{p}{n}-1}). \tag{21}$$

From (20) and (21), we obtain

$$(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) \cdot (\overbrace{x_1, \cdots, x_1}^{\frac{p}{n}}, \cdots, \underbrace{x_\gamma, \cdots, x_\gamma}_{\frac{p}{n}}, 0, \cdots, 0)^T = \mathbf{0}. \tag{22}$$

Let $\mathbf{X} = (x_1, \cdots, x_1, \cdots, x_\gamma, \cdots, x_\gamma, 0, \cdots, 0)^T$. Evidently $\mathbf{X} \in \sum_{\frac{\gamma p}{n}}$. Because γ is an even integer, we can write $\mathbf{X} = \mathbf{X}_1 - \mathbf{X}_2$, where $\mathbf{X}_1, \mathbf{X}_2 \in \sum_{\frac{\gamma p}{2n}}$. Thus, $(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) \cdot (\mathbf{X}_1 - \mathbf{X}_2) = \mathbf{0}$, i.e., $\mathbf{A} \ltimes \mathbf{X}_1 = \mathbf{A} \ltimes \mathbf{X}_2$. Because $\frac{\gamma p}{2n} = [\frac{kn}{p}]\frac{p}{n} \leq k$, $\mathbf{X}_1, \mathbf{X}_2 \in \sum_k$. This contradicts our assumption that there exists at most one signal $\mathbf{x} \in \sum_k$ such that $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}$. Therefore, we complete our proof. □

Remark 3. In fact, previous studies showed that *spark*$(\mathbf{A}) > 2k$ is also a necessary condition for the uniqueness of the sparse solution [29]. Theorem 2 contains this result for the case $p = n$.

3.3. Coherence

Theorem 3. $\mu(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) = \mu(\mathbf{A})$.

Proof. Let $\mathbf{c}_i$ be the i column of $\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}$. From (21), we can obtain $\langle \mathbf{c}_i, \mathbf{c}_j \rangle = 0$ if $\frac{p}{n} \nmid i-j$. If $\frac{p}{n} \mid i-j$, we have $\langle \mathbf{c}_i, \mathbf{c}_j \rangle = \langle \mathbf{b}^s_{\lceil \frac{in}{p} \rceil}, \mathbf{b}^s_{\lceil \frac{jn}{p} \rceil} \rangle = \langle \mathbf{a}_{\lceil \frac{in}{p} \rceil}, \mathbf{a}_{\lceil \frac{jn}{p} \rceil} \rangle$, where $s = (i \mod \frac{p}{n}) - 1$ if $i \mod \frac{p}{n} \neq 0$, else $s = \frac{p}{n} - 1$.

Namely,

$$\langle \mathbf{c}_i, \mathbf{c}_j \rangle = \begin{cases} \langle \mathbf{a}_{\lceil \frac{in}{p} \rceil}, \mathbf{a}_{\lceil \frac{jn}{p} \rceil} \rangle & \text{if } \frac{p}{n} \mid i-j, \\ 0 & \text{else}, \end{cases} \tag{23}$$

because $\|\mathbf{c}_i\|_2 = \|\mathbf{b}_{\lceil \frac{in}{p} \rceil}\|_2 = \|\mathbf{a}_{\lceil \frac{in}{p} \rceil}\|_2$. Thus,

$$\begin{aligned} \mu(\mathbf{A} \otimes \mathbf{I}_{\frac{p}{n}}) &= \max_{1 \le i \ne j \le p} \frac{|\langle \mathbf{c}_i, \mathbf{c}_j \rangle|}{\|\mathbf{c}_i\|_2 \|\mathbf{c}_j\|_2} = \max_{1 \le i \ne j \le p} \frac{|\langle \mathbf{a}_{\lceil \frac{in}{p} \rceil}, \mathbf{a}_{\lceil \frac{jn}{p} \rceil} \rangle|}{\|\mathbf{a}_{\lceil \frac{in}{p} \rceil}\|_2 \|\mathbf{a}_{\lceil \frac{jn}{p} \rceil}\|_2} \\ &= \mu(\mathbf{A}). \qquad \square \end{aligned}$$

For any matrix $\mathbf{A}$, $spark(\mathbf{A}) \ge 1 + \frac{1}{\mu(\mathbf{A})}$ [6]. Note that $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x} = (\mathbf{A} \otimes \mathbf{I}_{p/n})\mathbf{x}$. From Corollary 1 and Theorem 3, we have the following corollary.

Corollary 2. *If $k < \frac{1}{2}(1 + \frac{1}{\mu(\mathbf{A})})$, then for each measurement vector $\mathbf{y} \in \mathbb{R}^{mp/n}$ there exists at most one signal $x \in \sum_k$ such that $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}$.*

3.4. Restricted isometry property

Theorem 4. *Suppose that a matrix $\mathbf{A}$ satisfies RIP of order k with constant $\delta^{\mathbf{A}}_k$ $(0 < \delta^{\mathbf{A}}_k < 1)$, then $\mathbf{A} \otimes \mathbf{I}_{p/n}$ satisfies RIP of order k with the same constant, i.e., $\delta^{\mathbf{A} \otimes \mathbf{I}_{p/n}}_k = \delta^{\mathbf{A}}_k$.*

Proof. If $\mathbf{A}$ satisfies RIP of order k with constant $\delta^{\mathbf{A}}_k$, then

$$(1 - \delta^{\mathbf{A}}_k)\|\mathbf{x}\|_2^2 \le \|\mathbf{A}\mathbf{x}\|_2^2 \le (1 + \delta^{\mathbf{A}}_k)\|\mathbf{x}\|_2^2 \tag{24}$$

holds for all $\mathbf{x} \in \sum_k$. For any n-dimensional vector $\mathbf{x} \in \sum_k$ satisfying the above inequality, we construct a p-dimensional vector $\mathbf{x}'$ via adding many 0's after $\mathbf{x}$. Namely, $\mathbf{x}' = (\mathbf{x}, 0, \cdots, 0)^T$. Obviously, $\|\mathbf{x}\|_2 = \|\mathbf{x}'\|_2$ and $\mathbf{x}' \in \sum_k$. So,

$$\|(\mathbf{I}_{p/n} \otimes \mathbf{A})\mathbf{x}'\|_2^2 = \left\| \begin{pmatrix} \mathbf{A} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\|_2^2 = \|\mathbf{A}\mathbf{x}\|_2^2. \tag{25}$$

Substituting (25) into (24), we obtain

$$(1 - \delta^{\mathbf{A}}_k)\|\mathbf{x}'\|_2^2 \le \|(\mathbf{I}_{p/n} \otimes \mathbf{A})\mathbf{x}'\|_2^2 \le (1 + \delta^{\mathbf{A}}_k)\|\mathbf{x}'\|_2^2. \tag{26}$$

From the definition of RIP, we know that $\delta^{\mathbf{I}_{p/n} \otimes \mathbf{A}}_k$ satisfies (26) for all p-dimensional vectors $\mathbf{x}' \in \sum_k$. Thus,

$$\delta^{\mathbf{I}_{p/n} \otimes \mathbf{A}}_k \ge \delta^{\mathbf{A}}_k. \tag{27}$$

From the theory of Kronecker product, we know that there exist two permutation matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ such that $\mathbf{A} \otimes \mathbf{I}_{p/n} = \mathbf{P}_1(\mathbf{I}_{p/n} \otimes \mathbf{A})\mathbf{P}_2$ [27]. For any k-sparse vector $\mathbf{x}' \in \mathbb{R}^p$,

$$\|(\mathbf{A} \otimes \mathbf{I}_{p/n})\mathbf{x}'\|_2^2 = \|\mathbf{P}_1(\mathbf{I}_{p/n} \otimes \mathbf{A})\mathbf{P}_2\mathbf{x}'\|_2^2 = \|(\mathbf{I}_{p/n} \otimes \mathbf{A})\mathbf{P}_2\mathbf{x}'\|_2^2.$$

Therefore,

$$\delta^{\mathbf{A} \otimes \mathbf{I}_{p/n}}_k = \delta^{\mathbf{I}_{p/n} \otimes \mathbf{A}}_k. \tag{28}$$

From (27) and (28), we have

$$\delta^{\mathbf{A} \otimes \mathbf{I}_{p/n}}_k \ge \delta^{\mathbf{A}}_k. \tag{29}$$

On the other hand, previous studies [8,28] have shown that $\delta^{\mathbf{A} \otimes \mathbf{I}_{p/n}}_k \le (1 + \delta^{\mathbf{A}}_k)(1 + \delta^{\mathbf{I}_{p/n}}_k) - 1$. Because $\mathbf{I}_{p/n}$ is an orthonormal basis, $\delta^{\mathbf{I}_{p/n}}_k = 0$ [8]. Therefore, we have

$$\delta^{\mathbf{A} \otimes \mathbf{I}_{p/n}}_k \le \delta^{\mathbf{A}}_k. \tag{30}$$

From (29) and (30), we complete our proof. $\square$

By Theorem 4 and [7], we have the following reconstruction condition.

Corollary 3. *For any p-dimensional vector $\mathbf{x} \in \sum_k$, if the sensing matrix $\mathbf{A}$ satisfies the RIP of order $2k$ with the constant $\delta^{\mathbf{A}}_{2k} < \sqrt{2} - 1$, then the exact recovery in STP-CS model is possible by the l_1 optimization.*

Remark 4. Theorem 4 and Corollary 3 indicate that if $\mathbf{A}$ satisfies the RIP, then any k-sparse vector $\mathbf{x} \in \mathbb{R}^p$ can be reconstructed from the measurements $\mathbf{y} = \mathbf{A} \ltimes \mathbf{x}$. The advantage of the STP-CS model is that for the same size of signal, the number of columns of sensing matrix can be set into a factor of that in traditional CS.

3.5. Parallel reconstruction algorithms

Some traditional reconstruction algorithms in CS can be used for our proposed STP-CS model because $\mathbf{A} \ltimes \mathbf{x} = (\mathbf{A} \otimes \mathbf{I}_{p/n})\mathbf{x}$. However, according to the structure of $\mathbf{A} \otimes \mathbf{I}_{p/n}$, a reconstruction instance in STP-CS model can be transformed into p/n independent reconstruction instances in traditional CS model. Algorithm 1 shows the specific process and the correctness is explained as follows.

Algorithm 1 Reconstruction algorithm of STP-CS.

```
Require: y ∈ ℝ^{mp/n} and A ∈ ℝ^{m×n}.
Ensure: x ∈ ℝ^p.
for i = 1 to p/n do
    y^i = (y_i, y_{p/n+i}, ··· , y_{(m−1)p/n+i})^T ∈ ℝ^m.
end for
for i = 1 to p/n do
    x^i <--(CS reconstruction algorithm)-- (y^i, A).
end for
X = [x^1, x^2, ··· , x^{p/n}] ∈ ℝ^{n×p/n}.
return x := vec(X^T).   {vec(X) denotes the vectorization of X}
```

Evidently,

$$\mathbf{A} \otimes \mathbf{I}_{p/n} = \begin{pmatrix} a_{11} & \cdots & 0 & a_{12} & \cdots & 0 & \cdots & \cdots & a_{1n} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{11} & 0 & \cdots & a_{12} & \cdots & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & a_{22} & \cdots & 0 & \cdots & \cdots & a_{2n} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{21} & 0 & \cdots & a_{22} & \cdots & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & 0 & a_{m2} & \cdots & 0 & \cdots & \cdots & a_{mn} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{m1} & 0 & \cdots & a_{m2} & \cdots & \cdots & 0 & \cdots & a_{mn} \end{pmatrix}. \tag{31}$$

For $i = 1, \cdots, p/n$, we have

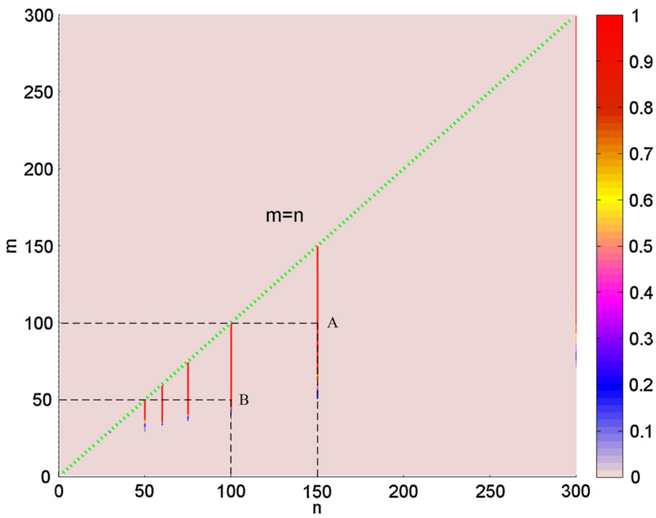


Fig. 1. The effects of m and n on the percentage of recovery. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$
\begin{aligned}
y_i &= \sum_{j=1}^{n} a_{1j} x_{(j-1)\frac{p}{n}+i} \\
&= (a_{11}, a_{12}, \cdots, a_{1n}) \cdot (x_i, x_{\frac{p}{n}+i}, \cdots, x_{(n-1)\frac{p}{n}+i})^T, \\
y_{i+\frac{p}{n}} &= \sum_{j=1}^{n} a_{2j} x_{(j-1)\frac{p}{n}+i} \\
&= (a_{21}, a_{22}, \cdots, a_{2n}) \cdot (x_i, x_{\frac{p}{n}+i}, \cdots, x_{(n-1)\frac{p}{n}+i})^T, \\
&\vdots \\
y_{i+(m-1)\frac{p}{n}} &= \sum_{j=1}^{n} a_{mj} x_{(j-1)\frac{p}{n}+i} \\
&= (a_{m1}, a_{m2}, \cdots, a_{mn}) \cdot (x_i, x_{\frac{p}{n}+i}, \cdots, x_{(n-1)\frac{p}{n}+i})^T.
\end{aligned}
\tag{32}
$$

From Eq. (32) we obtain

$$
\begin{pmatrix} y_i \\ y_{\frac{p}{n}+i} \\ \cdots \\ y_{(m-1)\frac{p}{n}+i} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_i \\ x_{\frac{p}{n}+i} \\ \cdots \\ x_{(n-1)\frac{p}{n}+i} \end{pmatrix}. \tag{33}
$$

Let $\mathbf{x}^i = (x_i, x_{\frac{p}{n}+i}, \cdots, x_{(n-1)\frac{p}{n}+i})^T$ and $\mathbf{y}^i = (y_i, y_{\frac{p}{n}+i}, \cdots, y_{(m-1)\frac{p}{n}+i})^T$. Therefore, for each i we can use some classical CS reconstruction algorithms to obtain $\mathbf{x}^i$. If all reconstruction instances complete, let $\mathbf{X} = [\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^{p/n}]$. Last the original signal $\mathbf{x}$ can be reconstructed through performing the vectorization on $\mathbf{X}^T$.

4. Experiments

In this section, we will study our proposed STP-CS model from an experimental point of view. Given an original signal $\mathbf{x}$ and the reconstructed signal $\mathbf{x}'$, we define the reconstruction error as $\|\mathbf{x}' - \mathbf{x}\|$. We believe that the original vector is reconstructed if the reconstruction error is smaller than 10^{-10}. For a randomly selected 30-sparse vector $\mathbf{x} \in \mathbb{R}^{300}$, Fig. 1 shows the effects of m and n on the percentage of recovery. The x and y axes, respectively, represent the number of columns and rows of the sensing matrix in our proposed STP-CS model, i.e., n and m. The green line plots the function $m = n$ and the color bar on the right represents the percentage of successful reconstructions. Red, yellow and blue mean that the percentage of successful reconstructions are equal to 1, 0.6 and 0.2, respectively. For example, the coordinate of point A in Fig. 1 is $(150, 100)$ and the color of point A is red. It means that when $m = 100$ and $n = 150$, the percentage of successful reconstructions is equal to 1. Similarly, the coordinate of point B is $(100, 50)$ and the color of B is also red. It means that when $m = 50$ and $n = 100$, the percentage of successful reconstructions is also equal to 1. Of course, sometimes the percentage is not equal to 1. The result shows that the percentage of the point $(100, 40)$ is smaller than 1. The percentage of the point $(150, 20)$ equals to 0, i.e., the reconstruction process completely failed. That is, the value of m can not be arbitrary small.

We should note that in order to achieve data compression, it requires that $n \mid p$ and $m < n$. The sensing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ in this experiment is the standard Gaussian sensing matrix. Namely, the elements of $\mathbf{A}$ are identically and independently sampled from a normal distribution with mean 0 and variance 1. We can see that n can be set equal to 50, 60, 75, 100, 150, 300. When $n = 300$, STP-CS degenerates into traditional CS. To reduce the storage space, we can set n to be relatively small size (e.g. $n = 50, 60, 75, 150$) and the percentage of recovery does not receive much influence.

For a randomly selected 12-sparse vector $\mathbf{x} \in \mathbb{R}^{300}$, Table 1 shows the percentage of recovery for different sensing matrices and different p/n, where $m = 59$. For the sensing matrices in this experiment, we choose three typical random constructions and two deterministic constructions, i.e., Gaussian matrices, Bernoulli matrices, Random matrices, Toeplitz matrices [10] and Chaotic matrices [11]. This experiment was performed 2000 times and we took the average value. Table 1 indicates that different n ($n > m$ and $n \mid p$) almost has no effect on the percentage of recovery.

In Fig. 2, we investigate the effects of the number of measurements on the percentage of recovery. We choose a 25-sparse vector

Table 1
The percentage of recovery for different p/n.

p/n	Matrix				
	Gaussian	Bernoulli	Random	Toeplitz	Chaotic
1	0.9535	0.978	0.971	0.964	0.95
2	0.963	0.982	0.97	0.962	0.9555
3	0.962	0.978	0.9695	0.963	0.9535
4	0.9575	0.9795	0.97	0.9715	0.947
5	0.958	0.9825	0.971	0.964	0.943

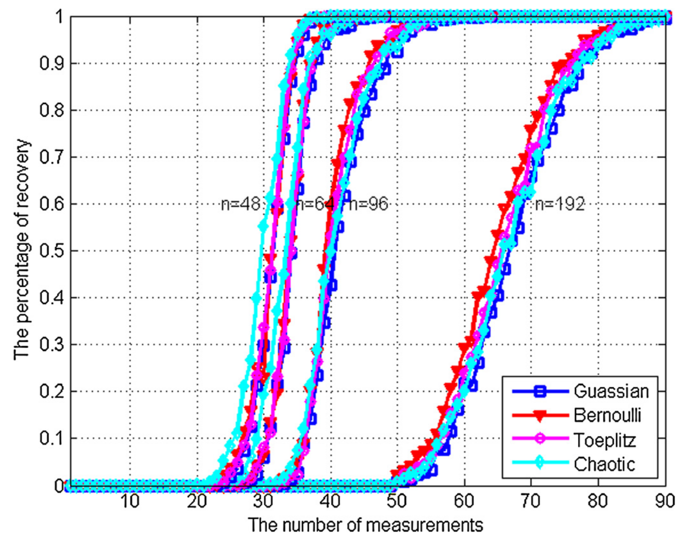


Fig. 2. The effects of m on the percentage of recovery.

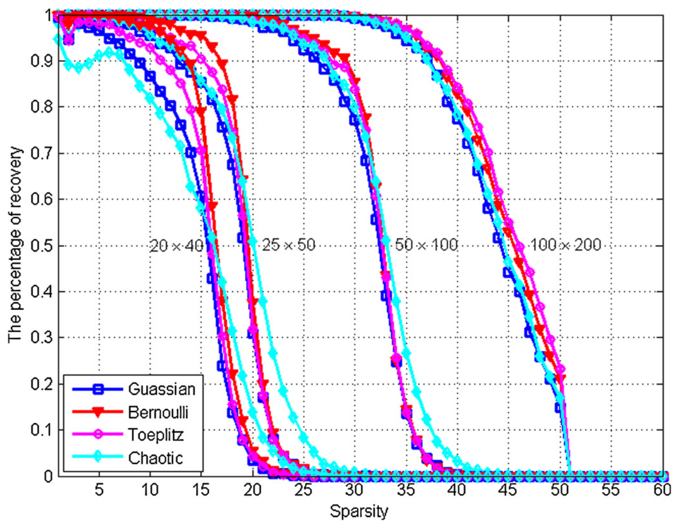


Fig. 3. The effects of the sparsity on the percentage of recovery.

$\mathbf{x} \in \mathbb{R}^{192}$ and the maximum column of sensing matrices $m_{\max}$ is 90. In the STP-CS model, the parameter n (the number of column of the sensing matrix $\mathbf{A}$) must be a factor of p (the dimension of the original vector $\mathbf{x}$). Here we consider $n = 192, 96, 64$ and 48. Note that if $n = 192$, then the STP-CS model degenerates into the traditional CS model. Fig. 2 indicates that CS must require at least 90 measurements to exactly reconstruct $\mathbf{x}$ and the size of sensing matrix is 90×192. However, we can set $m = 60, 47$, and 38 when $n = 96, 64$ and 48 in STP-CS, respectively. Thus, the STP-CS model has the advantage of reducing the size of sensing matrix compared to traditional CS. Fig. 3 investigates the effects of sparsity on the percentage of recovery for different sensing matrices. We choose a k-sparse vector $\mathbf{x} \in \mathbb{R}^{200}$ at random and the maximum sparsity $k_{\max}$ in this experiment is 60. In order to achieve the same compression ratio, we can flexibility choose different sizes of sensing matrices, i.e., 20×40, 25×50, 50×100 and 100×200. The experimental result shows that for low-sparsity signals, sensing matrix with relatively large size in STP-CS (smaller than that in CS) can exactly reconstruct the original signal. For instance, if the original signal is 18-sparse, we can use a 50×100 sensing matrix in STP-CS to reconstruct the original signal, rather than using a 100×200 sensing matrix in CS.

5. Comparison

To reduce the storage space of the sensing matrix, Gan presented a block-based model for high-dimensional signal compression and reconstruction, called BCS [17]. In fact, the main idea behind this method is that the image is first divided into small blocks with the same size and then all blocks are compressed using CS with the same sensing matrix $\mathbf{A}_B$, where the size of each block is $B \times B$. The sparsification process of BCS can be performed over each sub-block. However, Mun and Fowler [18] have pointed out that this straightforward "block-independent" method will produce severe blocking artifacts and is thus not usually a reasonable solution. To ameliorate such blocking artifacts, the sparsity transformation must take the form of a full-image transform (e.g., discrete wavelet transform), rather than being constrained to each sub-block [18]. Equivalently, the sensing matrix $\mathbf{A}_{BCS}$ of the whole image in Eq. (3) is a block diagonal matrix described as

$$\mathbf{A}_{BCS} = \mathbf{I}_N \otimes \mathbf{A}_B = \begin{pmatrix} \mathbf{A}_B & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_B & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_B \end{pmatrix}, \tag{34}$$

where N is the number of blocks. In our proposed STP-CS model, the sparsity transformation is also performed over the full-image. Previous theoretical analysis indicates that the sensing matrix $\mathbf{A}_{STP-CS}$ of the whole image can be equivalently described as

$$\mathbf{A}_{STP-CS} = \mathbf{A} \otimes \mathbf{I}_{p/n} = \begin{pmatrix} a_{11}\mathbf{I}_{p/n} & \cdots & a_{1n}\mathbf{I}_{p/n} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{I}_{p/n} & \cdots & a_{mn}\mathbf{I}_{p/n} \end{pmatrix}, \tag{35}$$

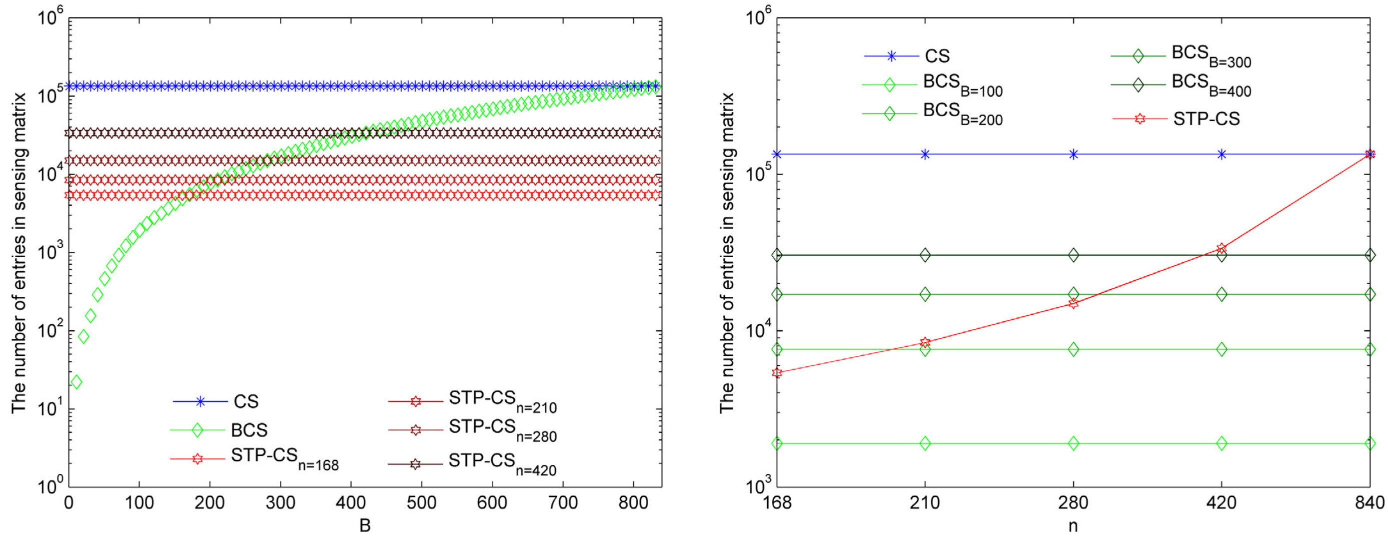
where a_{ij} is the (i, j)-th entry of $\mathbf{A}$. From Eq. (34) and (35), we can see that STP-CS is somewhat similar to BCS in structure from a high-level perspective. But they are essentially different because the tensor product operation $\otimes$ is not commutative. From a storage overhead point of view, we will compare the numbers of entries in sensing matrices between CS, BCS, and STP-CS. Suppose that the original signal $\mathbf{x}$ is a $p \times p$ image. Note that in [17], $\mathbf{x}$ is divided into small blocks with the same size B and the i-th block of the original image $\mathbf{x}$ is represented by a vector $\mathbf{x}_i$. If we take this vectorization approach, the experimental results show that the number of entries of the sensing matrix in STP-CS is much smaller than those in CS and BCS. Next, we assume that each block in BCS is not vectorized and thus all blocks are compressed using a $\lfloor mB/p \rfloor \times B$ sensing matrix. Evidently, the size of the sensing matrix in CS is $m \times p$. According to previous theoretical results, the size of the sensing matrix in STP-CS is $m \cdot n/p \times p \cdot n/p$, i.e., $mn/p \times n$, where n is a factor of p. Thus, the numbers of entries in sensing matrices between CS, BCS, and STP-CS are mp, $\lfloor mB/p \rfloor B$ and mn^2/p, respectively. Fig. 4(a) and (b), respectively, show the effects of B and n on the numbers of entries in various sensing matrices. In Fig. 4(a), we select four different n in STP-CS, i.e., $n = 168, 210, 280$ and 420, which satisfy that n is a factor of p and $m < n$. In Fig. 4(b), we chose four different B, i.e., $B = 100, 200, 300$ and 400. The experimental results indicate that the number of entries of sensing matrix in STP-CS is smaller than that in CS. Unfortunately, all experimental results show that the number of entries of sensing matrix in STP-CS is not smaller than that in BCS unless the size of each block B is very large. However, we should note that in practical applications, B can not be too small because the sparsity transformation in BCS is performed over the full-image and B should be much larger than the sparsity of the image. Take the discrete wavelet transform for example, the most of the image information are concentrated on the top-left corner after sparsification. If B is too small, the reconstruction of the image is evidently impossible. Comprehensively, it seems that STP-CS is more competitive than CS from a storage overhead point of view.

6. Conclusion

In this paper, we present a new model for data compression and reconstruction based on semi-tensor product, called STP-CS. The traditional CS model can be viewed as a special case of our proposed STP-CS model. From a theoretical point of view, we study the reconstruction condition of STP-CS in terms of *spark*, *coherence*, and RIP. In addition, a parallel method for signal reconstruction in STP-CS is proposed. Subsequently, a number of experiments show that even the sensing matrices in STP-CS are lower-dimensional, the percentage of recovery is almost the same as that in traditional CS. Last, we give a specific comparison between STP-CS and BCS. Comprehensively, the proposed STP-CS model for signal compression and reconstruction is competitive.

Acknowledgments

This paper is supported by the National Key Research and Development Program of China (Grant no. 2016YFB0800602),



(a) $p = 840$, $m = 160$. (b) $p = 840$, $m = 160$.

Fig. 4. The effects of B and n on the numbers of entries in various sensing matrices.

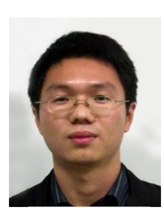
the National Natural Science Foundation of China (Grant nos. 61472045 and 61573067), the Beijing City Board of Education Science and Technology Key Project (Grant nos. KZ201510015015 and KM201510015009), and the Beijing University of Posts and Telecommunications Youth Scientific Research Innovation Program.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.dsp.2016.07.003.

References

[1] E.J. Candès, T. Tao, Decoding by linear programming, IEEE Trans. Inf. Theory 51 (12) (Dec. 2005) 4203–4215.
[2] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2) (Feb. 2006) 489–509.
[3] E.J. Candès, Compressive sampling, in: Proceedings of the International Congress of Mathematicians, vol. 3, Aug. 2006, pp. 1433–1452.
[4] S. Mallat, Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (Dec. 1993) 3397–3415.
[5] S. Foucart, A note on guaranteed sparse recovery via l_1-minimization, Appl. Comput. Harmon. Anal. 29 (1) (2010) 97–103.
[6] D.L. Donoho, M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization, Proc. Natl. Acad. Sci. 100 (5) (Mar. 2003) 2197–2202.
[7] E.J. Candès, The restricted isometry property and its implications for compressed sensing, C. R. Math. 346 (9) (May 2008) 589–592.
[8] M.F. Duarte, Y.C. Eldar, Structured compressed sensing: from theory to applications, IEEE Trans. Signal Process. 59 (9) (Sep. 2011) 4053–4085.
[9] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, Constr. Approx. 28 (3) (Dec. 2008) 253–263.
[10] J. Haupt, W.U. Bajwa, G. Raz, R. Nowak, Toeplitz compressed sensing matrices with applications to sparse channel estimation, IEEE Trans. Inf. Theory 56 (11) (Nov. 2010) 5862–5875.
[11] L. Yu, J.P. Barbot, G. Zhang, H. Sun, Compressive sensing with chaotic sequence, IEEE Signal Process. Lett. 17 (8) (Aug. 2010) 731–734.
[12] L.R. Welch, Lower bounds on the maximum cross correlation of signals, IEEE Trans. Inf. Theory 20 (3) (May 1974) 397–399.
[13] T. Strohmer, R. Heath, Grassmanian frames with applications to coding and communication, Appl. Comput. Harmon. Anal. 14 (3) (Nov. 2003) 257–275.
[14] S. Mallat, G. Davis, Z. Zhang, Adaptive time-frequency decompositions, Opt. Eng. 33 (7) (July 1994) 2183–2191.
[15] D.L. Donoho, Y. Tsaig, I. Drori, J.L. Starck, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit, IEEE Trans. Inf. Theory 58 (2) (Feb. 2012) 1094–1121.
[16] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmon. Anal. 26 (3) (May 2009) 301–321.
[17] L. Gan, Block compressed sensing of natural images, in: 15th International Conference on Digital Signal Processing, July 2007, pp. 403–406.
[18] S. Mun, J.E. Fowler, Residual reconstruction for block-based compressed sensing of video, in: Data Compression Conference, DCC, March 2011, pp. 183–192.
[19] J.E. Fowler, S. Mun, E.W. Tramel, Block-based compressed sensing of images and video, Found. Trends Signal Process. 4 (4) (2012) 297–416.
[20] D. Cheng, Semi-tensor product of matrices and its application to Morgen's problem, Sci. China Ser. Inf. Sci. 44 (3) (June 2001) 195–212.
[21] D. Cheng, L. Zhang, On semi-tensor product of matrices and its applications, Acta Math. Appl. Sin. 18 (4) (June 2003) 219–228.
[22] D. Cheng, H. Qi, A. Xue, A survey on semi-tensor product of matrices, J. Syst. Sci. Complex. 20 (2) (June 2007) 304–322.
[23] D. Cheng, Y. Dong, Semi-tensor product of matrices and its some applications to physics, Methods Appl. Anal. 10 (4) (Dec. 2003) 565–588.
[24] D. Laschov, M. Margaliot, A maximum principle for single-input Boolean control networks, IEEE Trans. Autom. Control 56 (4) (2011) 913–917.
[25] D. Cheng, Y. Zhao, Y. Mu, Strategy optimization with its application to dynamic games, in: 49th IEEE Conference on Decision and Control, Dec. 2010, pp. 5822–5827.
[26] D. Cheng, Some applications of semitensor product of matrices in algebra, Comput. Math. Appl. 52 (6) (Sep. 2006) 1045–1066.
[27] H.V. Henderson, S.R. Searle, The vec-permutation matrix, the vec operator and Kronecker products: a review, Linear Multilinear Algebra 9 (4) (1981) 271–288.
[28] M.F. Duarte, R.G. Baraniuk, Kronecker compressive sensing, IEEE Trans. Image Process. 21 (2) (Feb. 2012) 494–504.
[29] Y.C. Eldar, G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge University Press, 2012.



Dong Xie is a Ph.D. Candidate in Computer Science at the Beijing University of Posts and Telecommunications, Beijing, China. He received his M.S. degree from Hangzhou Normal University, Hangzhou, China, in 2013. His currently research interests include compressed sensing, network security, lattice-based cryptography, and homomorphic signature.

Haipeng Peng received his M.S. degree in system engineering from Shenyang University of Technology, Shenyang, China, in 2006, and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2010. He is currently an associate professor at the School of Computer Science and Technology, Beijing University of Posts and Telecommunications. His research interests include information security, network security, complex networks and control of dynamical systems. Dr. H. Peng is the co-author of 50 scientific papers.

Lixiang Li received the M.S. degree in circuit and system from Yanshan University, Qinhuangdao, China, in 2003, and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2006. She is currently a professor at the School of Computer Science and Technology, Beijing University of Posts and Telecommunications. Her research interests include swarm intelligence, information security and network security. Dr. L. Li is the co-author of 70 scientific papers.

Yixian Yang received the M.S. degree in applied mathematics in 1986 and the Ph.D. degree in electronics and communication systems in 1988 from Beijing University of Posts and Telecommunications, Beijing, China. He is the Managing Director of information security center, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include network security, information security and coding theory. Dr. Y. Yang is the co-author of 300 scientific articles and 50 patents.