Received 8 June 2014; revised 23 September 2014; accepted 3 November 2014. Date of publication 3 December, 2014;
date of current version 4 September, 2015.

Digital Object Identifier 10.1109/TETC.2014.2372151

On the Security of Compressed Sensing-Based Signal Cryptosystem

ZUYUAN YANG[1,2], (Member, IEEE), WEI YAN[1], AND YONG XIANG[2], (Senior Member, IEEE)

[1]Faculty of Automation, Guangdong University of Technology, Guangzhou 510006, China
[2]School of Information Technology, Deakin University, Melbourne, VIC 3125, Australia

CORRESPONDING AUTHOR: Z. YANG (yangzuyuan@aliyun.com)

This work was supported in part by the Program for New Century Excellent Talents in University under Grant NCET-13-0740, in part by the Natural Science Foundation of Guangdong Province under Grant S2011030002886, in part by the Guangdong Natural Science Funds for Distinguished Young Scholar under Grant 2014A030306037, in part by the National Natural Science Foundation of China under Grant 61104053, Grant 61322306, Grant 61333013, and Grant 61273192, and in part by the Australian Research Council under Grant DP 110102076.

ABSTRACT With the development of the cyber-physical systems (CPS), the security analysis of the data therein becomes more and more important. Recently, due to the advantage of joint encryption and compression for data transmission in CPS, the emerging compressed sensing (CS)-based cryptosystem has attracted much attention, where security is of extreme importance. The existing methods only analyze the security of the plaintext under the assumption that the key is absolutely safe. However, for sparse plaintext, the prior sparsity knowledge of the plaintext could be exploited to partly retrieve the key, and then the plaintext, from the ciphertext. So, the existing methods do not provide a satisfactory security analysis. In this paper, it is conducted in the information theory frame, where the plaintext sparsity feature and the mutual information of the ciphertext, key, and plaintext are involved. In addition, the perfect secrecy criteria (Shannon-sense and Wyner-sense) are extended to measure the security. While the security level is given, the illegal access risk is also discussed. It is shown that the CS-based cryptosystem achieves the extended Wyner-sense perfect secrecy, but when the key is used repeatedly, both the plaintext and the key could be conditionally accessed.

INDEX TERMS Compressed sensing, CPS, security.

I. INTRODUCTION

Mutual communication is always a natural requirement of human beings and no one wants to live in an information-isolated island. Nowadays, with the development of the network and computer technology, it becomes more and more convenient for people to communicate with each other. The fast developed network or cyber frame has greatly changed our society, and the related work will change the world more in future. As Guo *et al* mentioned in [1]: ''research advances in cyber-physical systems (CPS) promise to transform our world with systems that will far exceed those of today in terms of: effectiveness, adaptability, autonomicity, energy efficiency, precision, reliability, safety, usability, scalability, stability and user-centric applicability''.

Regarding the data process in CPS, two items should be considered particularly, one is the data size related to the efficiency, the other is the safety related to the privacy. Currently, the signals/data are often transmitted over some insecure and bandwidth-constrained communication channels, and it is appealing to encrypt and compress them. These two issues have indeed attracted much attention from researchers in different areas [2]–[4]. Traditionally, people use the compression-encryption method which means compression first and then encryption. This method brings the intuitive simplicity to encryption after the redundant data is stripped (by compression). However, it could compromise security as the source signals may have been accessed illegally before encryption (even before compression). To deal with this security concern, the encryption-compression method is developed, which first performs encryption and then compression [4]–[6]. The drawback of the encryption-compression method is its limited compression capability [3], [4].

Recently, the emerging compressed sensing (CS)-based cryptosystem has attracted much attention, due to the fact that it can address the issues of encryption and compression jointly [7]–[9]. In this cryptosystem, the key is made

of a underdetermined mixing matrix, whose row number is much less than column number [10]–[13]. Thus, when it is utilized to encrypt the plaintext, the latter is compressed into a block with smaller size at the same time. Furthermore, in both encryption and decryption processes, one only needs to carry out simple operations [14]–[16]. The computational complexities of these operations are much less than that of the current mainstream RSA encryption scheme developed by **R**ivest, **S**hamir and **A**dleman [17]. In addition, as far as the ciphertext transmission/communication is concerned, since the CS-based scheme can decrypt the whole plaintext from parts (without order) of the received ciphertext packages, it allows a high package loss rate in the corresponding communication system. As a result, the CS-based cryptosystem shows great potential to encryption and compression, plus the subsequent ciphertext transmission.

For the CS-based cryptosystem, one question arises naturally: how about its security? Just like what happened in other encryption methods [18]–[21], the security analysis of the CS-based cryptosystem has also attracted much attention [9], [10], [22]. However, the existing methods on security analysis only explore the security of the plaintext by assuming that the key is absolutely safe. Unfortunately, this assumption might not hold if the plaintext is sparse,[1] e.g., images are sparse in general [24], [25]. In this case, it is possible to exploit the prior sparsity knowledge of the plaintext to extract some information of the key from the ciphertext. This could make the cryptosystem suffer from the related-key attack [26], which has already failed the well-known wired equivalent privacy (WEP) protocol. Hence, current security analysis methods do not give a reliable solution to the CS-based cryptosystem.

In this paper, the security of the CS-based cryptosystem is analyzed, similar to [4] and [20], under the ciphertext-only attack. It is analyzed in the information theory frame, where the entropy, the conditional entropy, the mutual information (MI), and the conditional MI of the ciphertext, key, and plaintext are considered, together with the sparsity of the plaintext. Here the term ''entropy'' stands for Shannon entropy. The existing perfect secrecy criteria (Shannon-sense [27] and Wyner-sense [28]) are extended to measure the security. We also investigate the possible insecurity caused by the repeated use of the key. We show that in this case, the key and the plaintext may be partly accessed under some conditions by using the state-of-the-art information processing technology such as blind source separation (BSS) [29]–[32].

The remainder of the paper is organized as follows. Section II introduces the CS-based encryption scheme and the corresponding decryption method. The security of current CS-based cryptosystem is analyzed in Section III, together with the definitions of the extended perfect secrecy criteria. Finally, Section IV concludes the paper.

The following notations are used throughout the paper:

$\mathbf{x}, x_i$	Column vector, the ith element of $\mathbf{x}$
$\mathbf{X}, \mathbf{x}_j, x_{ij}$	Matrix, the jth column of $\mathbf{X}$, the (i, j)th entry of $\mathbf{X}$
$\mathbf{X}^T, \mathbf{X}^{-1}$	Transpose of $\mathbf{X}$, inverse of $\mathbf{X}$
$f_{\mathbf{x}}, f_{\mathbf{x},\mathbf{y}}$	Probability density function (PDF) of $\mathbf{x}$, joint PDF of $\mathbf{x}$ and $\mathbf{y}$
$f_{\mathbf{x}=\mathbf{y}}$	Probability of $\mathbf{x} = \mathbf{y}$
$H(x_i), H(\mathbf{x})$	Entropy of x_i, joint entropy of the elements in $\mathbf{x}$
$H(x_i\|y_j)$	Conditional entropy of x_i given y_i
$I(\mathbf{x}; \mathbf{y})$	MI of $\mathbf{x}$ and $\mathbf{y}$
$I(\mathbf{x}; \mathbf{y}, \mathbf{z})$	MI of $\mathbf{x}$ and the combination of $\mathbf{y}$, $\mathbf{z}$
$I(\mathbf{x}; \mathbf{y}\|\mathbf{z})$	Conditional MI of $\mathbf{x}$ and $\mathbf{y}$ given $\mathbf{z}$
$\Re$	Real number set
$\mathbf{0}$	Zero column vector (or matrix)

II. CS-BASED ENCRYPTION

The typical CS-based encryption model is as follows [9], [22]:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where $\mathbf{x} \in \Re^m$ denotes the ciphertext or encrypted message, $\mathbf{A} \in \Re^{m\times n}$ stands for the key matrix, $\mathbf{s} \in \Re^n$ is the plaintext or source message, and m, n denote the length of the blocks of the ciphertext and the plaintext, respectively. In the above CS-based encryption model, m is often much smaller than n, and the plaintext $\mathbf{s}$ is sparse (meaning that its values are close to zero at most time instants) or has a sparse representation under some known basis matrix. Traditionally, $\mathbf{s}$ is considered to be sparse with at most k nonzero elements ($k < m$), i.e. k-sparse [15]. The key $\mathbf{A}$ is designed to satisfy a k-order restricted isometry property, which means there exists a constant $\delta_k \in (0, 1)$ such that

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

holds for all k-sparse $\mathbf{s}$ [33]. In practice, the key $\mathbf{A}$ is often constructed to be a matrix with independent entries to get an optimal restricted isometry constant [34].

In the encryption procedure, $\mathbf{s}$ is encrypted into the ciphertext $\mathbf{x}$ with the key $\mathbf{A}$, by using a mixing or matrix multiplication method. This process is quite simple and it needs little computational cost. In the decryption procedure, based on the received ciphertext $\mathbf{x}$ and the available key $\mathbf{A}$, the plaintext $\mathbf{s}$ can be recovered through searching the sparsest solution satisfying (1). This can be solved by using a number of existing methods [14], [16]. For reference, the model of the L1-norm based robust optimization method is given below:

$$\begin{cases} \text{Minimize } \|\mathbf{s}\|_1 \\ \text{subject to } \mathbf{x} = \mathbf{A}\mathbf{s} \end{cases} \tag{3}$$

where $\|\mathbf{s}\|_1 = \sum_{i=1}^{n} |s_i|$. Since $\mathbf{s}$ is known to be k-sparse in prior, the computational cost of solving (3) can approach to kmn, nearly a linear complexity as $k, m \ll n$ [16].

[1] A basic premise in the CS-based cryptosystem is that the plaintext is sparse or has a sparse representation. Otherwise, the ciphertext cannot be decrypted effectively.

This implies that the CS-based scheme is a very competitive method for encryption. Besides, data compression is also achieved during the encryption process as m is often much smaller than n. In addition, since the whole plaintext can be decrypted from parts of the ciphertext by using the CS-based scheme, it allows a high package loss rate in the corresponding communication system for ciphertext transmission.

For the CS-based cryptosystem in (1), security is a primary concern. Given that the security of the key is guaranteed, the security of the plaintext can be analyzed using the methods in [9], [10], and [22]. Yet, in the context of CS-based cryptosystems, the plaintext is required to be sparse to ensure effective decryption. Then an interesting question is whether and how the sparsity feature of the plaintext affects the security of the CS-based cryptosystem. This question will be answered in Section III under the following assumptions:

A-1) The entries of $\mathbf{A}$ are statistically independent.
A-2) The elements of $\mathbf{s}$ have the same distribution.
A-3) The entries of $\mathbf{A}$ and the elements of $\mathbf{s}$ are statistically independent.

The assumptions A-1) and A-2) are widely utilized in the CS-based cryptosystem and its security analysis [8], [22], [33], [34]. The assumption A-3) is common and practical [4], [17].

III. SECURITY OF THE CS-BASED CRYPTOSYSTEM

We start from the (Shannon) entropy and the MI of the random variables/vectors/matrices in information theory frame. The entropy reflects the uncertainty of a random quantity. It is always nonnegative and reaches zero if and only if the random quantity becomes certain. Compared with the entropy, the MI is often used to measure the extent of the correlation between two random quantities and it is provided for the first rigorous statistical treatment of secrecy by Shannon [27]. The larger the MI is, the greater the correlation is. MI is also nonnegative and becomes zero if and only if the associated random quantities are statistically independent [4]. There is intimate relationship between the entropy and MI, and this relationship will be invoked for formulae derivation in this section.

A. EXTENDED PERFECT SECRECY CRITERIA

Based on MI, the perfect secrecy criteria of Shannon-sense [27] and Wyner-sense [28] are traditionally used to measure the security. They all reflect the correlation of the ciphertext and the plaintext, assuming that the key is always absolutely secure. However, as we can see, the information of the key is contained in the public ciphertext and thus it is generally an active factor to the security of the cryptosystem. If the key is cracked, the plaintext will be broken inevitably. Even if only a little information of the key is extracted, the consequence could be devastating. An important example is that the WEP protocol has been broken by the related-key attack, without knowing much information of the key [26]. In the following, we extend the afore-mentioned two criteria to measure the security of both the plaintext and the key.

For a cryptosystem Φ with plaintext M, key T, and ciphertext E, let

$$G(\mathrm{E}, \mathrm{M}, \mathrm{T}) = I(\mathrm{E}; \mathrm{M}) + I(\mathrm{E}; \mathrm{T}). \tag{4}$$

Then, the extension of the Shannon-sense perfect secrecy is defined as follows.

Definition 1: The cryptosystem Φ achieves the extended Shannon-sense (ESS) perfect secrecy if

$$G(\mathrm{E}, \mathrm{M}, \mathrm{T}) = 0. \tag{5}$$

Notably, both the correlation of ciphertext-plaintext and that of ciphertext-key are involved in the function G, denoted by $I(\mathrm{E}; \mathrm{M})$ and $I(\mathrm{E}; \mathrm{T})$, respectively. Here, an example is given to show what the cryptosystem Φ could be if it satisfies the ESS perfect secrecy.

Example 1: For the Latin Square system in [27], let the plaintext M take values $\mathrm{M}_i, i = 1, 2, \ldots, n$ with equal probability, and it is encrypted by the key T which takes values $\mathrm{T}_i, i = 1, 2, \ldots, n$ with equal probability too. Denote the corresponding ciphertext E to be $\mathrm{E}_i, i = 1, 2, \ldots, n$, which are calculated by the following rule:

$$\mathrm{E}_k = \mathrm{T}_i \bigoplus \mathrm{M}_j \tag{6}$$

where $k = \mathrm{mod}(i + j - 1, n)$ and $\bigoplus$ denotes the overlay of two quantities. The system (6) forms a Latin Square and Fig. 1 gives an illustration for the case of $n = 5$ [27].

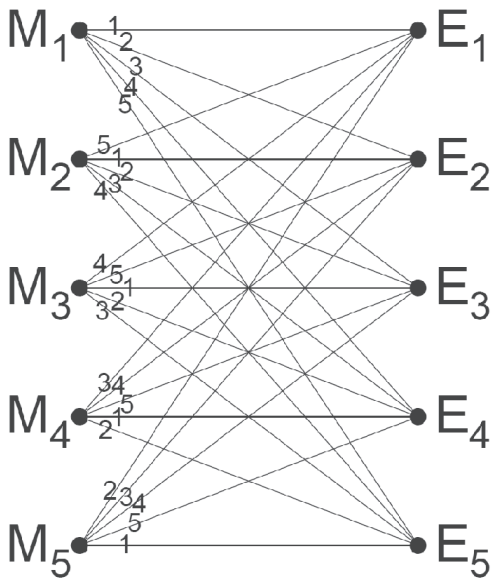


FIGURE 1. Latin Square cryptosystem.

In this system, for all $i \in 1, 2, \ldots, n$, we have the following probabilities:

$$\begin{cases} f_{\mathrm{M}=\mathrm{M}_i} = \frac{1}{n} \\ f_{\mathrm{T}=\mathrm{T}_i} = \frac{1}{n} \\ f_{\mathrm{E}=\mathrm{E}_i} = \frac{1}{n} \end{cases}. \tag{7}$$

Then, the entropy of E is

$$H(\mathrm{E}) = \log n \tag{8}$$

and the conditional entropy of E given T is

$$\begin{aligned} H(\mathrm{E}|\mathrm{T}) &= \sum_{i=1}^{n} f_{\mathrm{T}=\mathrm{T}_i} H(\mathrm{E}|\mathrm{T}_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} H(\mathrm{E}|\mathrm{T}_i) \\ &= \frac{1}{n} n H(\mathrm{M}) \\ &= \log n. \end{aligned} \tag{9}$$

Similarly, the conditional entropy of E given M is

$$H(\mathrm{E}|\mathrm{M}) = \log n. \tag{10}$$

From (8)-(10), it yields

$$\begin{cases} I(\mathrm{E};\mathrm{M}) = H(\mathrm{E}) - H(\mathrm{E}|M) = 0 \\ I(\mathrm{E};\mathrm{T}) = H(\mathrm{E}) - H(\mathrm{E}|T) = 0. \end{cases} \tag{11}$$

Therefore, $G(\mathrm{E}, \mathrm{M}, \mathrm{T})$ equals zero and thus this Latin Square system achieves ESS perfect secrecy.

It is worth noting that if M takes values without equal probability, the conditional entropy of E given T will be less than $\log n$ and thus $I(\mathrm{E};\mathrm{T}) > 0$, although $I(\mathrm{E};\mathrm{M})$ still equals zero. Alternatively, even if the ciphertext E and the plaintext M have already been mutually independent, E could be still correlated with the key T. Regarding the CS-based cryptosystem, the plaintext is sparse (for the sake of effective decryption), implying that the probability of taking value zero is greater than that of taking other values. Since the plaintext does not take values with equal probability, the ciphertext could be correlated with the key. As a result, one may extract some information of the key from the ciphertext, and then further access the plaintext. This leads to the motivation that the key security should also be considered in the analysis of the CS-based cryptosystem.

While MI provides a theoretical criterion for security evaluation of cryptosystems, its asymptotic notion has also been explored for some particular applications. Here, we refer to the Wyner-sense perfect secrecy [28] developed by Wyner for wireless channel transmission. This criterion has been used to measure the security of a compression-involved cryptosystem [4], which is similar to the CS-based system concerned in this paper. Since Wyner's criterion considers only the MI of the ciphertext and the plaintext, we extend it as follows.

Definition 2: The cryptosystem Φ has extended Wyner-sense (EWS) perfect secrecy if

$$\lim_{n\to\infty} \frac{G(\mathrm{E}, \mathrm{M}, \mathrm{T})}{n} = 0 \tag{12}$$

where n denotes the length of the plaintext.

In the remainder of this section, the security of the CS-based cryptosystem (1) will be analyzed under the ESS and the EWS perfect secrecy criteria.

B. SECURITY ANALYSIS

Regarding the cryptosystem (1), the function G can be rewritten as

$$G(\mathbf{x}, \mathbf{s}, \mathbf{A}) = I(\mathbf{x}; \mathbf{s}) + I(\mathbf{x}; \mathbf{A}) \tag{13}$$

Based on G, the security will be analyzed in both one-time padding and multi-time padding cases.

1) SECURITY IN THE CASE OF ONE-TIME PADDING

We first conduct security analysis in the scenario that the key is used for only one time, i.e., one-time padding. We start from the ESS criterion and have the following theorem.

Theorem 1: The cryptosystem (1) does not achieve ESS perfect secrecy, i.e.,

$$G(\mathbf{x}, \mathbf{s}, \mathbf{A}) > 0. \tag{14}$$

As shown in [22], $I(\mathbf{x}; \mathbf{s}) > 0$. Considering that MI is always nonnegative, we have $I(\mathbf{x}; \mathbf{A}) \geq 0$. Then, it follows from (13) that $G(\mathbf{x}, \mathbf{s}, \mathbf{A}) > 0$, i.e., (14) holds.

The ESS criterion is very strong and it is generally not easy for a cryptosystem to reach that security level. So, it is not surprising that the CS-based cryptosystem in (1) does not satisfy the ESS criterion.[2] It is interesting to investigate whether the CS-based cryptosystem (1) fulfills the EWS criterion which is weaker than the ESS criterion. The result is given by the following Theorem 2.

Theorem 2: The cryptosystem (1) achieves the EWS perfect secrecy.

Proof: Recall that the plaintext is $k(k < m)$-sparse and m is fixed in advance in the standard CS frame. This means that there are at most k nonzero elements in $\mathbf{s}$. Without loss of generality, let $\mathbf{s}^a = [s_1, s_2, \ldots, s_{n-k}]^T$ be a zero vector. Also, we denote

$$\begin{cases} \mathbf{s}^b = [s_{n-k+1}, s_{n-k+2}, \ldots, s_n]^T \\ \mathbf{A}^a = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-k}]^T \\ \mathbf{A}^b = [\mathbf{a}_{n-k+1}, \mathbf{a}_{n-k+2}, \ldots, \mathbf{a}_n]^T. \end{cases} \tag{15}$$

Then, based on the matrix decomposition operation, (1) can be rewritten as

$$\begin{aligned} \mathbf{x} &= \mathbf{A}\mathbf{s} \\ &= \mathbf{A}^a\mathbf{s}^a + \mathbf{A}^b\mathbf{s}^b. \end{aligned} \tag{16}$$

Since $k < m$ and the entries of $\mathbf{A}$ are independent, $\mathbf{A}^b$ is full column rank. Thus, based on the Moore-Penrose matrix inverse, $\mathbf{s}^b$ can be calculated by

$$\mathbf{s}^b = \left(\left(\mathbf{A}^b\right)^T\left(\mathbf{A}^b\right)\right)^{-1}\left(\mathbf{A}^b\right)^T(\mathbf{x} - \mathbf{A}^a\mathbf{s}^a). \tag{17}$$

Furthermore, based on the features of entropy function [35], it holds that

$$\begin{cases} H\left(\mathbf{s}^a|\mathbf{s}\right) = 0 \\ H(\mathbf{s}) \leq H\left(\mathbf{s}^a\right) + H\left(\mathbf{s}^b\right) \end{cases} \tag{18}$$

[2] An modified system with enhanced security is shown in [23].

and

$$\begin{aligned} H\left(\mathbf{s}|\mathbf{s}^a\right) &= H\left(\mathbf{s},\mathbf{s}^a\right) - H\left(\mathbf{s}^a\right) \\ &= H\left(\mathbf{s}^a|\mathbf{s}\right) + H(\mathbf{s}) - H\left(\mathbf{s}^a\right) \\ &= H(\mathbf{s}) - H\left(\mathbf{s}^a\right) \\ &\le H\left(\mathbf{s}^a\right) + H\left(\mathbf{s}^b\right) - H\left(\mathbf{s}^a\right) \\ &= H\left(\mathbf{s}^b\right). \end{aligned} \tag{19}$$

Let us consider the first item $I(\mathbf{x};\mathbf{s})$ of $G(\mathbf{x},\mathbf{s},\mathbf{A})$ given $\mathbf{s}^a$. As shown in the assumption A-2), the elements of $\mathbf{s}^b$ are with the same distribution. Thus, they have the same entropy. Notably, their joint entropy is no greater than the summation of each entropy. Denoting the size of the space in which each element takes values to be M (M is often finite, e.g., it equals 256 in image source messages), their joint entropy has the following relationship with M based on the well-known maximum entropy property[3]:

$$\begin{aligned} H\left(\mathbf{s}^b\right) &\le \underbrace{H(s_{n-k+1}) + H(s_{n-k+2}) + \cdots + H(s_n)}_{k \text{ terms}} \\ &\le \log M + \log M + \cdots + \log M \\ &= k \log M. \end{aligned} \tag{20}$$

Then, we have

$$\begin{aligned} I(\mathbf{x};\mathbf{s}|\mathbf{s}^a) &= H(\mathbf{s}|\mathbf{s}^a) - H(\mathbf{s}|\mathbf{x},\mathbf{s}^a) \\ &\le H(\mathbf{s}|\mathbf{s}^a) \\ &\le H\left(\mathbf{s}^b\right) \\ &\le k \log M. \end{aligned} \tag{21}$$

Now, we consider the second item $I(\mathbf{x};\mathbf{A})$ of $G(\mathbf{x},\mathbf{s},\mathbf{A})$ given $\mathbf{s}^a$. Since $\mathbf{A} = \left[\mathbf{A}^a,\mathbf{A}^b\right]$, then based on the chain rule of MI [35], [36], $I(\mathbf{x};\mathbf{A}|\mathbf{s}^a)$ can be written as

$$\begin{aligned} I(\mathbf{x};\mathbf{A}|\mathbf{s}^a) &= I\left(\mathbf{x};\mathbf{A}^a,\mathbf{A}^b|\mathbf{s}^a\right) \\ &= I\left(\mathbf{x};\mathbf{A}^b|\mathbf{s}^a\right) + I\left(\mathbf{x};\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}^a\right) \\ &= I\left(\mathbf{x};\mathbf{A}^b|\mathbf{s}^a\right) + H\left(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}^a\right) \\ &\quad - H\left(\mathbf{A}^a|\mathbf{x},\mathbf{A}^b,\mathbf{s}^a\right). \end{aligned} \tag{22}$$

Due to the fact that $\mathbf{A}^a$ is independent of $\mathbf{A}^b$ and $\mathbf{s}$ is independent of $\mathbf{A}$ from the assumptions A-1) and A-3), it yields

$$\begin{aligned} H\left(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}\right) &= H\left(\mathbf{A}^a|\mathbf{s}\right) - I\left(\mathbf{A}^a;\mathbf{A}^b|\mathbf{s}\right) \\ &= H(\mathbf{A}^a) - (H(\mathbf{A}^a|\mathbf{s}) + H(\mathbf{A}^b|\mathbf{s}) \\ &\quad - H(\mathbf{A}^a,\mathbf{A}^b|\mathbf{s})) \\ &= H(\mathbf{A}^a) - H(\mathbf{A}^a) - H(\mathbf{A}^b) + H(\mathbf{A}|\mathbf{s}) \\ &= -H(\mathbf{A}^b) + H(\mathbf{A}) \\ &= -H(\mathbf{A}^b) + H(\mathbf{A}^a) + H(\mathbf{A}^b) \\ &= H(\mathbf{A}^a). \end{aligned} \tag{23}$$

[3]The entropy of a random variable x reaches the maximum $\log|x|$ if and only if x is uniformly distributed, where $|x|$ denotes the size of the space in which x takes values [35].

Specifically, we have

$$H\left(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}^a\right) = H(\mathbf{A}^a). \tag{24}$$

In addition, in the case of $\mathbf{s}^a = \mathbf{0}$, it holds that $\mathbf{x}$ is equal to $\mathbf{A}^b\mathbf{s}^b$. Clearly, $\mathbf{x}$ is certain if $\mathbf{A}^b$ and $\mathbf{s}^b$ are given, but it is generally not reversible without additional condition. Thus we obtain

$$\begin{aligned} H\left(\mathbf{A}^a|\mathbf{x},\mathbf{A}^b,\mathbf{s}^a\right) &= H\left(\mathbf{A}^a|\mathbf{A}^b\mathbf{s}^b,\mathbf{A}^b,\mathbf{s}^a\right) \\ &\ge H\left(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}^b,\mathbf{A}^b,\mathbf{s}^a\right) \\ &= H(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}^b,\mathbf{s}^a) \\ &= H(\mathbf{A}^a|\mathbf{A}^b,\mathbf{s}) \\ &= H(\mathbf{A}^a). \end{aligned} \tag{25}$$

Substituting (24) and (25) into (22), it follows

$$\begin{aligned} I(\mathbf{x};\mathbf{A}|\mathbf{s}^a) &\le I\left(\mathbf{x};\mathbf{A}^b|\mathbf{s}^a\right) + H(\mathbf{A}^a) - H(\mathbf{A}^a) \\ &= I\left(\mathbf{x};\mathbf{A}^b|\mathbf{s}^a\right) \\ &= H\left(\mathbf{A}^b|\mathbf{s}^a\right) - H\left(\mathbf{A}^b|\mathbf{x},\mathbf{s}^a\right). \end{aligned} \tag{26}$$

As we know, there are km entries in $\mathbf{A}^b$, and they are not only independent mutually but also independent of $\mathbf{s}^a$. Let C be the maximum size of the spaces in which each entry takes values,[4] then based on the maximum entropy property, we have

$$\begin{aligned} H\left(\mathbf{A}^b|\mathbf{s}^a\right) &= H\left(\mathbf{A}^b\right) \\ &= \sum_{i=1}^{m}\sum_{j=n-k+1}^{n} H\left(a_{ij}\right) \\ &\le \sum_{i=1}^{m}\left(\underbrace{\log C + \log C + \cdots + \log C}_{k \text{ terms}}\right) \\ &= km \log C. \end{aligned} \tag{27}$$

Since

$$H(\mathbf{A}^b|\mathbf{x},\mathbf{s}^a) \ge 0 \tag{28}$$

substituting (27) and (28) into (26) gives

$$\begin{aligned} I(\mathbf{x};\mathbf{A}|\mathbf{s}^a) &\le km \log C - H(\mathbf{A}^b|\mathbf{x},\mathbf{s}^a) \\ &\le km \log C. \end{aligned} \tag{29}$$

From (21) and (29), it holds that

$$I(\mathbf{x};\mathbf{s}|\mathbf{s}^a) + I(\mathbf{x};\mathbf{A}|\mathbf{s}^a) \le k \log M + km \log C. \tag{30}$$

Considering the prior knowledge that $\mathbf{s}^a = \mathbf{0}$, then one can conclude that

$$\lim_{n\to\infty} \frac{G(\mathbf{x},\mathbf{s},\mathbf{A})}{n} \le \lim_{n\to\infty} \frac{k \log M + km \log C}{n} = 0. \tag{31}$$

[4]Similar to [4], C is considered to be finite here.

Note that MI is always nonnegative. Therefore, $G(\mathbf{x}, \mathbf{s}, \mathbf{A})/n$ is always nonnegative, and thus it approaches zero with the increase of n. This completes the proof. □

From Theorems 1 and Theorem 2, we can see that if the key is used for only one time (i.e., one-time padding), the cryptosystem (1) achieves the EWS perfect secrecy, although fails to reach the stronger ESS perfect secrecy.

2) SECURITY IN THE CASE OF MULTI-TIME PADDING

It is known that using the key for many times may lead to insecurity [20]. So, it is important to analyze the security of the cryptosystem (1) in the situation of multi-time padding, where the key is used repeatedly. Regarding also the cipher-only attack, we have the following theorem.

Theorem 3: If the key is used repeatedly for N times, where $N \geq n$ and the plaintexts are independent, then the cryptosystem (1) does not achieve EWS perfect secrecy.

Proof: In the case that the key is used repeatedly for N times, the ciphertexts form a matrix with N columns, as well as the plaintexts. Denote them to be $\mathbf{X}$ and $\mathbf{S}$, respectively. For the convenience of derivation, $\forall i \in 1, 2, \ldots, N$, let $\mathbf{X}_i$ and $\mathbf{S}_i$ be the matrices composed of the first i columns of $\mathbf{X}$ and $\mathbf{S}$, respectively. That is

$$\begin{cases} \mathbf{X}_i = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i] \\ \mathbf{S}_i = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_i] \end{cases} \tag{32}$$

where $\mathbf{x}_i$ and $\mathbf{s}_i$ denote the ith ciphertext and plaintext, respectively. An equivalent form of (32) is

$$\begin{cases} \mathbf{X}_i = [\mathbf{X}_{i-1}, \mathbf{x}_i] \\ \mathbf{S}_i = [\mathbf{S}_{i-1}, \mathbf{s}_i]. \end{cases} \tag{33}$$

Then, considering the MI of $\mathbf{X}_N$ and $\mathbf{S}_N$, we have

$$\begin{aligned} I(\mathbf{X}_N; \mathbf{S}_N) &= I(\mathbf{X}_N; \mathbf{S}_{N-1}, \mathbf{s}_N) \\ &= I(\mathbf{X}_N; \mathbf{s}_N) + I(\mathbf{X}_N; \mathbf{S}_{N-1}|\mathbf{s}_N) \\ &= I(\mathbf{s}_N; \mathbf{X}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_N|\mathbf{s}_N) \\ &= I(\mathbf{s}_N; \mathbf{X}_{N-1}, \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_N|\mathbf{s}_N) \\ &= I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{s}_N; \mathbf{X}_{N-1}|\mathbf{x}_N) \\ &\quad + I(\mathbf{S}_{N-1}; \mathbf{X}_N|\mathbf{s}_N). \end{aligned} \tag{34}$$

Since the conditional MI is always nonnegative, it follows

$$I(\mathbf{s}_N; \mathbf{X}_{N-1}|\mathbf{x}_N) \geq 0. \tag{35}$$

From (34) and (35), it results in

$$\begin{aligned} I(\mathbf{X}_N; \mathbf{S}_N) &\geq I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_N|\mathbf{s}_N) \\ &= I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}, \mathbf{x}_N|\mathbf{s}_N) \\ &= I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}|\mathbf{s}_N) \\ &\quad + I(\mathbf{S}_{N-1}; \mathbf{x}_N|\mathbf{X}_{N-1}, \mathbf{s}_N). \end{aligned} \tag{36}$$

With the same reason that the conditional MI is nonnegative, we have

$$I(\mathbf{S}_{N-1}; \mathbf{x}_N|\mathbf{X}_{N-1}, \mathbf{s}_N) \geq 0. \tag{37}$$

It yields from (37) and (36) that

$$I(\mathbf{X}_N; \mathbf{S}_N) \geq I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}|\mathbf{s}_N). \tag{38}$$

Considering the second item of the right side in (38), three random quantities are involved. Regarding the "MI" of three random quantities $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, one has the following formulae[5] [36]:

$$\begin{cases} I(\mathbf{x}; \mathbf{y}; \mathbf{z}) = I(\mathbf{x}; \mathbf{y}) - I(\mathbf{x}; \mathbf{y}|\mathbf{z}) \\ I(\mathbf{x}; \mathbf{y}; \mathbf{z}) = I(\mathbf{x}; \mathbf{z}) - I(\mathbf{x}; \mathbf{z}|\mathbf{y}) \\ I(\mathbf{x}; \mathbf{y}; \mathbf{z}) = I(\mathbf{y}; \mathbf{z}) - I(\mathbf{y}; \mathbf{z}|\mathbf{x}). \end{cases} \tag{39}$$

Then it follows from the formulae in (39) that

$$\begin{aligned} &I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}; \mathbf{s}_N) \\ &\quad = I(\mathbf{S}_{N-1}; \mathbf{s}_N) - I(\mathbf{S}_{N-1}; \mathbf{s}_N|\mathbf{X}_{N-1}) \end{aligned} \tag{40}$$

and

$$\begin{aligned} &I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}; \mathbf{s}_N) \\ &\quad = I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}) - I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}|\mathbf{s}_N). \end{aligned} \tag{41}$$

Besides, since the plaintexts are mutually independent, then $\mathbf{s}_N$ is independent of $\mathbf{S}_{N-1}$. This means that

$$I(\mathbf{S}_{N-1}; \mathbf{s}_N) = 0. \tag{42}$$

Meanwhile, due to the nonnegativity of the conditional MI, we have

$$I(\mathbf{S}_{N-1}; \mathbf{s}_N|\mathbf{X}_{N-1}) \geq 0. \tag{43}$$

As a result, one can conclude from (40)-(43) that

$$I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}|\mathbf{s}_N) \geq I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}). \tag{44}$$

Substituting (44) into (38), we obtain

$$\begin{aligned} I(\mathbf{X}_N; \mathbf{S}_N) &\geq I(\mathbf{s}_N; \mathbf{x}_N) + I(\mathbf{S}_{N-1}; \mathbf{X}_{N-1}) \\ &= I(\mathbf{x}_N; \mathbf{s}_N) + I(\mathbf{X}_{N-1}; \mathbf{S}_{N-1}). \end{aligned} \tag{45}$$

From the viewpoint of recursion, it holds from (45) that

$$I(\mathbf{X}_{N-1}; \mathbf{S}_{N-1}) \geq I(\mathbf{x}_{N-1}; \mathbf{s}_{N-1}) + I(\mathbf{X}_{N-2}; \mathbf{S}_{N-2}). \tag{46}$$

Moreover, $\forall i \in \{1, 2, \ldots, N-2\}$, we have

$$I(\mathbf{X}_{N-i}; \mathbf{S}_{N-i}) \geq I(\mathbf{x}_{N-i}; \mathbf{s}_{N-i}) + I(\mathbf{X}_{N-i-1}; \mathbf{S}_{N-i-1}). \tag{47}$$

Furthermore, it is easy to see from (32) that there is only one column in $\mathbf{X}_1$, as well as in $\mathbf{S}_1$. That is to say, $\mathbf{X}_1$ and $\mathbf{S}_1$ degenerate to $\mathbf{x}_1$ and $\mathbf{s}_1$, respectively, which leads to

$$I(\mathbf{X}_1; \mathbf{S}_1) = I(\mathbf{x}_1; \mathbf{s}_1). \tag{48}$$

Similarly, we have

$$I(\mathbf{X}; \mathbf{S}) = I(\mathbf{X}_N; \mathbf{S}_N). \tag{49}$$

[5]It is worth noting that the "MI" of three random quantities fails to satisfy the nonnegativity. However, the MI viewpoint is still utilized here for the convenience of formula derivation.

Based on (45)-(49), one can obtain

$$\begin{aligned} I(\mathbf{X};\mathbf{S}) &= I(\mathbf{X}_N;\mathbf{S}_N) \\ &\geq I(\mathbf{x}_N;\mathbf{s}_N) + I(\mathbf{X}_{N-1};\mathbf{S}_{N-1}) \\ &\geq I(\mathbf{x}_N;\mathbf{s}_N) + I(\mathbf{x}_{N-1};\mathbf{s}_{N-1}) \\ &\quad + I(\mathbf{X}_{N-2};\mathbf{S}_{N-2}) \\ &\vdots \\ &\geq I(\mathbf{x}_N;\mathbf{s}_N) + I(\mathbf{x}_{N-1};\mathbf{s}_{N-1}) \\ &\quad + \cdots + I(\mathbf{x}_2;\mathbf{s}_2) + I(\mathbf{X}_1;\mathbf{S}_1) \\ &= I(\mathbf{x}_N;\mathbf{s}_N) + I(\mathbf{x}_{N-1};\mathbf{s}_{N-1}) \\ &\quad + \cdots + I(\mathbf{x}_2;\mathbf{s}_2) + I(\mathbf{x}_1;\mathbf{s}_1) \\ &= \sum_{i=1}^{N} I(\mathbf{x}_i;\mathbf{s}_i). \end{aligned} \tag{50}$$

From the Lemma in [22], one can see that each ciphertext is not independent of the corresponding plaintext, i.e., $\forall i, \exists \delta > 0$, which results in

$$I(\mathbf{x}_i;\mathbf{s}_i) \geq \delta. \tag{51}$$

Then, if $N \geq n$, it follows from (50) and (51) that

$$\begin{aligned} \lim_{n\to\infty} \frac{G(\mathbf{X},\mathbf{S},\mathbf{A})}{n} &= \lim_{n\to\infty} \frac{I(\mathbf{X};\mathbf{S}) + I(\mathbf{X};\mathbf{A})}{n} \\ &\geq \lim_{n\to\infty} \frac{I(\mathbf{X};\mathbf{S})}{n} \\ &\geq \lim_{n\to\infty} \frac{\sum_{i=1}^{N} I(\mathbf{x}_i;\mathbf{s}_i)}{n} \\ &\geq \lim_{n\to\infty} \frac{N\delta}{n} \\ &\geq \delta \\ &> 0. \end{aligned} \tag{52}$$

Eq. (52) means that the cryptosystem (1) does not achieve the EWS perfect secrecy any more if the key is used no less than n times. This completes the proof. □

Comparing Theorem 2 and Theorem 3, it is clear that the security level of the cryptosystem (1) is lowered after the key is used repeatedly. This is consistent with the existing results [20]. However, it does not mean that the plaintext or key can be necessarily wiretapped or cracked in the multi-time padding case. This raises an interesting question: how does the multi-time padding scheme affect the security of the cryptosystem (1) (against ciphertext-only attack)? We will answer this question from the viewpoint of the BSS problem.

BSS is a technology which can recover the mixing matrix and the sources to some extent (depending on the properties of the mixing matrix and the sources) from the measured mixtures only [37], [38]. As we know, in the multi-time padding case, the cryptosystem (1) will become a linear mixing model of BSS, where the ciphertexts, the key, and the plaintexts correspond to the mixtures, the mixing matrix, and the sources, respectively. Then, some methods in BSS may be invoked to crack the key and the plaintexts, such as the sparse BSS methods [30], [39]. Regarding this, we have the following theorem.

Theorem 4: Given that the key is used repeatedly for N times in the cryptosystem (1). If the plaintexts are mutually independent, then with the increase of N, parts of the key and the plaintexts can be illegally accessed with probability 1.

Proof: Based on (1), the multi-time padding model is constructed as following:

$$\mathbf{X} = \mathbf{AS} \tag{53}$$

where $\mathbf{X}$ is an $m \times N$ matrix composed of the ciphertexts column by column, $\mathbf{A}$ denotes the same key in (1), and $\mathbf{S}$ is a $n \times N$ matrix composed of N plaintexts column by column. We can show that $\mathbf{A}$ and $\mathbf{S}$ have the following properties:

c1) any square $m \times m$ submatrix of $\mathbf{A}$ is nonsingular;
c2) each column of $\mathbf{S}$ has at most $m-1$ nonzero elements;
c3) $\mathbf{S}$ is sufficiently rich which is defined in [30].

Firstly, from the assumption A-1), it is obvious that the entries of $\mathbf{A}$ are independent. Then, any m columns of $\mathbf{A}$ are linearly independent. This implies that any $m \times m$ submatrix in $\mathbf{A}$ is nonsingular.

Secondly, in the CS-based cryptosystem, the plaintext should be k-sparse with $k < m$ for effective decryption. Therefore, in order to correctly decrypt the whole $\mathbf{S}$ in (53), each column of $\mathbf{S}$ should be k-sparse. That is, each column of $\mathbf{S}$ has at most $m-1$ nonzero elements.

Thirdly, based on the assumption A-2) and the condition that plaintexts are mutually independent, one can see that with the increase of N, it is with probability 1 that there exists in $\mathbf{S}$ a $n \times m$ submatrix whose first $n-m+1$ rows are zero and the remaining $m-1$ rows satisfy the Haar condition.[6] Similarly, for any index set ϕ with elements number $n-m+1$, where $\phi \subset \{1, 2, \ldots, n\}$, it is with probability 1 that there exists a $n \times m$ submatrix whose rows within index set ϕ are zero and the remaining $m-1$ rows satisfy the Haar condition. Note that for any $n \times m$ matrix $\mathbf{B}$, if there exists an $(m-1) \times m$ matrix satisfying the Haar condition, then any $m-1$ columns of $\mathbf{B}$ are linearly independent. Thus, based on the definition in [30], $\mathbf{S}$ is sufficiently rich.

Given the above three properties of $\mathbf{A}$ and $\mathbf{S}$, it results from [30, Th. 1], in conjunction with its proof, that $\mathbf{A}$ and $\mathbf{S}$ can be recovered (up to permutation and scaling) from the ciphertexts $\mathbf{X}$ by using a BSS technology. Since the columns of $\mathbf{A}$ are often normalized in advance in CS, there remains only permutation. Thus, with the increase of N, it is with probability 1 that all parts of the key can be accessed/obtained illegally (as well as the parts of the plaintexts), although their orders are unknown. This completes the proof. □

Compared with Theorem 3 showing that the CS-based encryption scheme with multi-time padding is insecure in theory, Theorem 4 and its proof further show some practical risks arising from the newly developed BSS technology.

IV. CONCLUSION

In this paper, the security of the standard CS-based cryptosystem is analyzed in the information theory frame,

[6] An $m \times n$ matrix satisfies the Haar condition if any m columns of the matrix are linearly independent [38].

which is important for data processing in the fast developed CPS. Specifically, the features of the MI and the entropy of random quantities are involved. From the proposed extended perfect secrecy criteria based on MI, it is proved that the EWS criterion is achieved by the CS-based encryption scheme, but the stronger ESS criterion cannot be achieved. It is also shown that in the multi-time padding case, where the key is used repeatedly, the security of the cryptosystem will fall below the EWS perfect secrecy level if the quantity of the used time is no less than that of the block length of the plaintext. Regarding the multi-time padding case, it is further proved that under some conditions, the key and the plaintext can be cracked partly by using modern information processing technology based on BSS (knowing only the ciphertexts). This means that there may exist some kind of insecurity when the usage of the key turns from one-time padding to multi-time padding, although permutation remains in the recovered parts. A possible solution could be utilizing the random key series, instead of the repeated usage of one fixed key.

REFERENCES

[1] S. Guo, H. Frey, N. Kato, and Y. Liu, "Special issue on cyber-physical systems (CPS)—Part I," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 1, pp. 6–9, Jun. 2013.

[2] N. Merhav, "Perfectly secure encryption of individual sequences," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1302–1310, Mar. 2013.

[3] D. Klinc, C. Hazay, A. Jagmohan, H. Krawczyk, and T. Rabin, "On compression of data encrypted with block ciphers," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6989–7001, Nov. 2012.

[4] M. Johnson, P. Ishwar, V. Prabhakaran, D. Schonberg, and K. Ramchandran, "On compressing encrypted data," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2992–3006, Oct. 2004.

[5] W. Liu, W. Zeng, L. Dong, and Q. Yao, "Efficient compression of encrypted grayscale images," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 1097–1102, Apr. 2010.

[6] X. Zhang, "Lossy compression and iterative reconstruction for encrypted image," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 53–58, Mar. 2011.

[7] J. Wu, W. Wang, Q. Liang, X. Wu, and B. Zhang, "Compressive sensing-based data encryption system with application to sense-through-wall UWB noise radar," *Secur. Commun. Netw.*, 2013, doi: 10.1002/sec.670.

[8] L. Zeng, X. Zhang, L. Chen, Z. Fan, and Y. Wang, "Scrambling-based speech encryption via compressed sensing," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 257, pp. 1–12, 2012.

[9] A. V. Sreedhanya and K. P. Soman, "Secrecy of cryptography with compressed sensing," in *Proc. IEEE Int. Conf. Adv. Comput. Commun. (ICACC)*, Coimbatore, India, Aug. 2012, pp. 207–210.

[10] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing," in *Proc. IEEE Int. Conf. Military Commun. (MILCOM)*, San Diego, CA, USA, Nov. 2008, pp. 1–7.

[11] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[12] Z. He, A. Cichocki, S. Xie, and K. Choi, "Detecting the number of clusters in n-way probabilistic clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2006–2021, Nov. 2010.

[13] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1135–1146, Feb. 2012.

[14] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4789–4812, Nov. 2008.

[15] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications.* London, U.K.: Cambridge Univ. Press, 2012.

[16] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[17] (Jul. 12, 2012). *What is the RSA Cryptosystem?* [Online]. Available: http://www.rsa.com/rsalabs

[18] K. Nuida and G. Hanaoka, "On the security of pseudorandomized information-theoretically secure schemes," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 635–652, Jan. 2013.

[19] S. H. Dau, V. Skachek, and Y. M. Chee, "On the security of index coding with side information," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3975–3988, Jun. 2012.

[20] R. S. Katti, S. K. Srinivasan, and A. Vosoughi, "On the security of randomized arithmetic codes against ciphertext-only attacks," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 19–27, Mar. 2011.

[21] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

[22] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Urbana-Champaign, IL, USA, Sep. 2008, pp. 813–817.

[23] M. Ramezani Mayiami, B. Seyfe, and H. G. Bafghi, "Perfect secrecy via compressed sensing," in *Proc. IEEE Workshop Commun. Inf. Theory (IWCIT)*, May 2013, pp. 1–5.

[24] G. Zhou, A. Cichocki, and S. Xie, "Fast nonnegative matrix/tensor factorization based on low-rank approximation," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2928–2940, Jun. 2012.

[25] Z. He, S. Xie, R. Zdunek, G. Zhou, and A. Cichocki, "Symmetric nonnegative matrix factorization: Algorithms and applications to probabilistic clustering," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2117–2131, Dec. 2011.

[26] E. Biham, "New types of cryptanalytic attacks using related keys," *J. Cryptol.*, vol. 7, no. 4, pp. 229–246, Apr. 1994.

[27] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[28] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[29] Z. Yang, Y. Xiang, Y. Rong, and S. Xie, "Projection-pursuit-based method for blind separation of nonnegative sources," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 1, pp. 47–57, Jan. 2013.

[30] P. Georgiev, F. Theis, and A. Cichocki, "Sparse component analysis and blind source separation of underdetermined mixtures," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 992–996, Jul. 2005.

[31] Z. He, S. Xie, S. Ding, and A. Cichocki, "Convolutive blind source separation in the frequency domain based on sparse representation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 5, pp. 1551–1563, Jul. 2007.

[32] G. Zhou, Z. Yang, S. Xie, and J.-M. Yang, "Mixing matrix estimation from sparse mixtures with unknown number of sources," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 211–221, Feb. 2011.

[33] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[34] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4388–4394, Sep. 2010.

[35] R. M. Gray, *Entropy and Information Theory*. New York, NY, USA: Springer-Verlag, 2013.

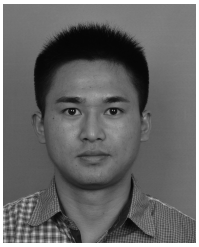
[36] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

[37] Z. Yang, G. Zhou, S. Xie, S. Ding, J.-M. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1112–1125, Apr. 2011.

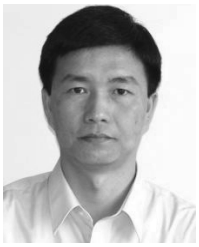
[38] A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY, USA: Wiley, 2002.

[39] Z. Yang, Y. Xiang, S. Xie, S. Ding, and Y. Rong, "Nonnegative blind source separation by sparse component analysis based on determinant measure," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 10, pp. 1601–1610, Oct. 2012.

ZUYUAN YANG (M'14) received the B.E. degree from the Hunan University of Science and Technology, Xiangtan, China, in 2003, and the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 2010. He received the Excellent Ph.D. Thesis Award of Guangdong Province. He is currently a Researcher with the Faculty of Automation, Guangdong University of Technology, Guangzhou, and his research interests include blind source separation, compressed sensing, nonnegative matrix factorization, and image processing.



WEI YAN received the B.E. degree from the Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan, China, in 2013. He is currently pursuing the master's degree with the Faculty of Automation, Guangdong University of Technology, Guangzhou, China, and his research interests include compressed sensing and nonnegative matrix factorization.



YONG XIANG (SM'12) received the Ph.D. degree in electrical and electronic engineering from The University of Melbourne, Melbourne, VIC, Australia. He is currently an Associate Professor and the Director of the Artificial Intelligence and Image Processing Research Cluster with the School of Information Technology, Deakin University, Melbourne. His research interests include signal and system estimation, information and network security, multimedia (speech/image/video) processing, and wireless sensor networks. He has served as the Program Chair, TPC Chair, Symposium Chair, and Session Chair for a number of international conferences. He serves as an Associate Editor of the IEEE ACCESS.