Yu *EURASIP Journal on Advances in Signal Processing* (2017) 2017:73
DOI 10.1186/s13634-017-0508-6

EURASIP Journal on Advances in Signal Processing

RESEARCH Open Access

On the security of compressed encryption with partial unitary sensing matrices embedding a secret keystream

Nam Yul Yu

Abstract

The principle of compressed sensing (CS) can be applied in a cryptosystem by providing the notion of security. In this paper, we study the computational security of a CS-based cryptosystem that encrypts a plaintext with a partial unitary sensing matrix embedding a secret keystream. The keystream is obtained by a keystream generator of stream ciphers, where the initial seed becomes the secret key of the CS-based cryptosystem. For security analysis, the total variation distance, bounded by the relative entropy and the Hellinger distance, is examined as a security measure for the indistinguishability. By developing upper bounds on the distance measures, we show that the CS-based cryptosystem can be computationally secure in terms of the indistinguishability, as long as the keystream length for each encryption is sufficiently large with low compression and sparsity ratios. In addition, we consider a potential chosen plaintext attack (CPA) from an adversary, which attempts to recover the key of the CS-based cryptosystem. Associated with the key recovery attack, we show that the computational security of our CS-based cryptosystem is brought by the mathematical intractability of a constrained integer least-squares (ILS) problem. For a sub-optimal, but feasible key recovery attack, we consider a successive approximate maximum-likelihood detection (SAMD) and investigate the performance by developing an upper bound on the success probability. Through theoretical and numerical analyses, we demonstrate that our CS-based cryptosystem can be secure against the key recovery attack through the SAMD.

Keywords: Compressed encryption, Hellinger distance, Indistinguishability, Integer least-squares (ILS) problem, Relative entropy, Total variation distance, Stream ciphers

1 Introduction

Compressed sensing (CS) [1–4] is a novel data acquisition scheme that samples a signal at a sub-Nyquist rate, which allows simultaneous data acquisition and compression. The original signal can be faithfully recovered from the measurement samples, if it is sparse with respect to a particular basis and sampled via a random projection. With efficient measurement and stable reconstruction, the CS technique has been of interest in a variety of research fields, e.g., communications [5–7], sensor networks [8–10], image processing [11–13], and radar [14].

Recently, a great deal of attention has been paid to the CS technique for data confidentiality in information security field. A *CS-based cryptosystem* encrypts a plaintext through a CS measurement process by keeping the sensing matrix secret. Then, the ciphertext can be decrypted by a CS reconstruction process. Thus, the CS-based cryptosystem performs simultaneous data acquisition and encryption at physical layer. Such a lightweight cryptosystem is particularly attractive for secure communications in wireless sensor networks, where the resources are not sufficient for providing data confidentiality by conventional encryption.

The security potential of compressed sensing was hinted by Candes and Tao [3], where the measurement samples were referred to as a weakly encrypted ciphertext. In [15], Rachlin and Baron proved that the CS-based cryptosystem cannot be perfectly secure but might be computationally secure. Orsdemir et al. [16] showed that it is computationally secure against a key search technique via an algebraic approach. Subsequently, many researchers have studied the security of CS-based cryptosystems

Correspondence: nyyu@gist.ac.kr
School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), Gwangju, Korea

for practical applications, which will be discussed with more details in Section 2.3. For a comprehensive review of CS techniques in information security, readers are referred to [17].

In this paper, we study the computational security of a CS-based cryptosystem that encrypts a plaintext with a partial unitary sensing matrix embedding a secret keystream. The keystream to be embedded is obtained by a keystream generator of stream ciphers, which ensures fast and efficient generation of the keystream. Assuming that the keystream is part of the original one with an extremely long period, we renew it at each encryption, which leads to a *one-time sensing* (*OTS*) cryptosystem. Then, the initial seed (or state) of the original keystream generator is essentially the secret *key* of the CS-based cryptosystem. With the sensing matrix, we demonstrate that the CS-based cryptosystem theoretically guarantees a stable and robust CS decryption for a legitimate recipient.

For security analysis, we first use probability metrics to investigate the security in a statistical manner. The *total variation* (*TV*) distance [18] between probability distributions of ciphertexts conditioned on a pair of plaintexts is examined as a security measure for the indistinguishability [19] of our CS-based cryptosystem. We investigate the TV distance by developing upper bounds on the *relative entropy* [20] and the *Hellinger* distance [21], which demonstrates that our CS-based cryptosystem can be computationally secure in terms of the indistinguishability, as long as the keystream length for each encryption is sufficiently large with low compression $\left(\frac{M}{N}\right)$ and sparsity $\left(\frac{K}{N}\right)$ ratios.

Next, we analyze the security of our CS-based cryptosystem by examining the resistance against a cryptanalytic attack. We consider a potential *chosen plaintext attack* (*CPA*) from an adversary to recover the key of our CS-based cryptosystem. In the CPA, the adversary needs to restore a keystream embedded in CS encryption, which is nontrivial unlike in stream ciphers, since the keystream is not outstanding from a known plaintext-ciphertext pair. Associated with the key recovery attack, we show that the security of our CS-based cryptosystem is based on the mathematical intractability of a constrained *integer least-squares* (*ILS*) problem. For a sub-optimal, but feasible key recovery attack, we consider a *successive approximate maximum-likelihood* (*ML*) *detection* (*SAMD*) for the adversary's CPA and investigate the performance by developing an upper bound on the success probability. Finally, theoretical analysis and numerical results reveal that our CS-based cryptosystem can be secure against the key recovery attack through the SAMD.

This paper is organized as follows. Section 2 reviews the CS principle, discusses some known CS-based cryptosystems, and summarizes the contributions of this paper. In Section 3, we describe a mathematical model of the CS-based cryptosystem proposed by this paper. We discuss a theoretical guarantee of CS decryption for a legitimate recipient by the cryptosystem. In Section 4, we analyze the indistinguishability of our CS-based cryptosystem, to demonstrate the computational security. Section 5 introduces an adversary's potential CPA strategy for key recovery, where we describe the details and examine the performance of SAMD. Section 6 presents numerical results to demonstrate the reliability and the security of our CS-based cryptosystem. Finally, concluding remarks will be given in Section 7.

2 Background

2.1 Notations

A matrix (or a vector) is represented by a boldface upper (or lower) case letter. $\mathbf{U}^T$ and $|\mathbf{U}|$ denote the transpose and the determinant of a matrix $\mathbf{U}$, respectively. $\mathrm{tr}(\mathbf{U})$ denotes the trace of a matrix $\mathbf{U}$ or the sum of all diagonal entries of $\mathbf{U}$. $\mathbf{U}(k,t)$ is an entry of an $M \times N$ matrix $\mathbf{U}$ in the kth row and the tth column, where $0 \le k \le M-1$ and $0 \le t \le N-1$. $\mu(\mathbf{U})$ denotes the maximum magnitude of the entries of $\mathbf{U}$, i.e., $\mu(\mathbf{U}) = \max_{k,t} |\mathbf{U}(k,t)|$. $\mathrm{diag}(\mathbf{s})$ is a diagonal matrix whose diagonal entries are from a vector $\mathbf{s}$. An identity matrix is denoted by $\mathbf{I}$, where the dimension is determined in the context. $\mathbf{W}$ is a conventional $N \times N$ Walsh-Hadamard matrix, where $\mathbf{W}\mathbf{W}^T = \mathbf{W}^T\mathbf{W} = N\mathbf{I}$. Also, $\mathbf{D}$ denotes a discrete-cosine transform (DCT) matrix, where $\mathbf{D}\mathbf{D}^T = \mathbf{D}^T\mathbf{D} = N\mathbf{I}$. For a vector $\mathbf{x} = (x_0, \cdots, x_{N-1})^T \in \mathbb{R}^N$, the l_p-norm of $\mathbf{x}$ is denoted by $||\mathbf{x}||_p = \left(\sum_{k=0}^{N-1} |x_k|^p\right)^{\frac{1}{p}}$, where $1 \le p < \infty$. If the context is clear, $||\mathbf{x}||$ denotes the l_2-norm of $\mathbf{x}$. A vector $\mathbf{n} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ is a Gaussian random vector with mean $\mathbf{0} = (0, \cdots, 0)^T$ and covariance $\sigma^2\mathbf{I}$. Finally, $\mathbb{E}[\cdot]$ denotes the average of a random vector or a random matrix.

Table 1 summarizes the abbreviations of this paper.

2.2 Compressed sensing

Compressed sensing (CS) [1–3] is to recover a sparse signal from the measurements that are believed to be incomplete. A signal $\mathbf{x} \in \mathbb{R}^N$ is called *K-sparse* with respect to a sparsifying (orthonormal) basis $\mathbf{\Psi}$ if $\boldsymbol{\theta} = \mathbf{\Psi}\mathbf{x}$ has at most K nonzero entries, where $K \ll N$. The sparse signal $\mathbf{x}$ is linearly measured by $\mathbf{r} = \mathbf{\Phi}\mathbf{x} + \mathbf{n} = \mathbf{\Phi}\mathbf{\Psi}^T\boldsymbol{\theta} + \mathbf{n} \in \mathbb{R}^M$, where $\mathbf{\Phi}$ is an $M \times N$ measurement matrix with $M \ll N$ and $\mathbf{n} \in \mathbb{R}^M$ is a measurement noise. The CS theory states that if the *sensing* matrix $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}^T$ obeys the *restricted isometry property* (*RIP*) [2], a stable and robust reconstruction of $\boldsymbol{\theta}$ can be guaranteed from the incomplete measurement $\mathbf{r}$. The CS reconstruction is accomplished by solving the l_1-minimization problem of

Table 1 Abbreviations

CoSaMP	Compressive sampling matching pursuit
CPA	Chosen plaintext attack
CS	Compressed sensing
CVP	Closest vector problem
DCT	Discrete-cosine transform
GSD	Generalized sphere decoding
ILS	Integer least-squares
KPA	Known plaintext attack
LFSR	Linear feedback shift register
ML	Maximum-likelihood
NMSE	Normalized mean squared error
NP	Nondeterministic polynomial time
PNR	Plaintext-to-noise power ratio
RIP	Restricted isometry property
RSSI	Received signal strength indicator
SAMD	Successive approximate maximum-likelihood detection
SSG	Self-shrinking generator
TV	Total variation

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} ||\boldsymbol{\theta}||_1 \text{subject to } ||\mathbf{A}\boldsymbol{\theta} - \mathbf{r}||_2 \leq \epsilon$$

with convex optimization or greedy algorithms [4]. For simplicity, this paper assumes $\boldsymbol{\Psi} = \mathbf{I}$, or that $\mathbf{x}$ is sparse in canonical basis, which yields the sensing matrix of $\mathbf{A} = \boldsymbol{\Phi}$.

2.3 Prior works on CS-based cryptosystems

Since the foundational works of [15] and [16], there have been many research efforts on CS-based cryptosystems. Bianchi, Bioglio, and Magli [22, 23] analyzed the security of a noiseless CS-based cryptosystem utilizing random Gaussian sensing matrices in an OTS manner. In [24], a similar analysis has been made for a noiseless CS-based cryptosystem having a circulant sensing matrix for efficient CS processes. Cambareri et al. [25] proposed a CS-based cryptosystem that supports multiclass encryption using a random Bernoulli matrix and its class-dependent variations. In spite of exploiting different security measures, i.e., indistinguishability [23] and asymptotic spherical security [25], the security analyses of [23] and [25] showed that the statistical properties of ciphertexts reveal only the information about the energy of the plaintexts. The security of the multiclass encryption scheme has been further investigated in [26] against a known plaintext attack (KPA), by examining the average number of candidate solutions matching a plaintext-ciphertext pair.

In addition to the secret sensing matrix, a CS-based cryptosystem may employ an extra cryptographic primitive, which can be considered as a *product cipher*. For instance, scrambling or random permutation has been additionally accomplished, before [27] or after [28] CS encryption. In [29], nonlinear diffusion has been added to quantized ciphertexts. Zhang et al. [30] proposed a bi-level protected CS (BLP-CS), where the sparsifying basis and the sensing matrix are generated with different secret keys. In the BLP-CS, the knowledge of both the sparsifying basis and the sensing matrix is required for CS decryption.

To gain a resistance against KPA and CPA, a CS-based cryptosystem normally operates in an OTS manner, by renewing the sensing matrix at each encryption. As the renewal requires the additional complexity and can quickly waste up the cryptographic resource for generating each sensing matrix, a CS-based cryptosystem reusing the sensing matrix during multiple encryptions has also been of interest. However, it is insecure against KPA and CPA, since an adversary can easily recover the sensing matrix with N linearly independent plaintexts by solving the system of linear equations [15]. While reusing the same sensing matrix, the BLP-CS [30] attempted to overcome the weakness and to achieve a CPA-resistance by ensuring a RIPless reconstruction for an adversary.

CS-based cryptosystems can work in a framework of *physical layer security* [31]. The emerging technology of physical layer security is a promising paradigm for enhancing wireless security [32], by exploiting the randomness of wireless channel characteristics. In [33], Agrawal and Vishwanath derived sufficient conditions for secret communications via CS in a wiretap channel. Reeves at al. [34] investigated the secrecy capacity of a wiretap channel employing CS. Dautov and Tsouri [35] used the received signal strength indicator (RSSI) from wireless channels for secure key establishment in a CS-based cryptosystem, where the shared key can be used to form a common sensing matrix in a sender and a recipient. In practice, a variety of CS-based cryptosystems concerning the security and privacy of multimedia, imaging, and smart grid data have been suggested and studied in [36–39].

2.4 Summary of contributions

The main results of this paper are summarized in comparison with prior works. Our CS-based cryptosystem encrypts a plaintext with a partial unitary sensing matrix embedding a secret keystream, which is used only once for each encryption. Thus, it operates in an OTS manner, similar to those of [22–25], but different from the BLP-CS [30]. It can further reduce the consumption of the cryptographic resource by renewing only the keystream of length N, not replacing the entire $M \times N$ sensing matrix, at each encryption. Unlike the BLP-CS, our CS-based cryptosystem uses only a single cryptographic primitive, or the secret keystream, while keeping the sparsifying basis public. Furthermore, the secret keystream can be efficiently

generated by a keystream generator of stream ciphers. Based on the RIP analysis, the knowledge of the sensing matrix, or equivalently the keystream, theoretically guarantees a reliable CS decryption.

In security analysis, we obtain the result by two different approaches. On the one hand, we demonstrate the indistinguishability of our CS-based cryptosystem, by investigating the TV distance between probability distributions of a pair of ciphertexts. This statistical approach seems like the analysis of [23], but we use a new probability metric of the Hellinger distance [21] to characterize the TV distance. On the other hand, we consider a potential CPA from an adversary for key recovery of our CS-based cryptosystem. By formulating the CPA as an NP-hard problem, we show that the success of the CPA is computationally infeasible for a sufficiently large keystream length. In addition, we introduce a sub-optimal but feasible CPA strategy and investigate the performance with the highest possible success probability. Finally, the CPA performance turns out to be quite poor even under an optimistic scenario, which guarantees the security against the CPA for our CS-based cryptosystem. The second type of security analysis is new in this paper.

3 Mathematical model

3.1 CS encryption with a partial unitary sensing matrix

A CS-based cryptosystem encrypts a sparse plaintext $\mathbf{x} \in \mathbb{R}^N$ through the CS measurement process with a sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$, which produces the ciphertext $\mathbf{r} = \mathbf{\Phi x} + \mathbf{n} \in \mathbb{R}^M$, where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ is a measurement noise. This paper proposes a CS-based cryptosystem that employs a partial unitary sensing matrix $\mathbf{\Phi}$ embedding a secret keystream, as defined in Definition 1.

Definition 1 *The sensing matrix*[1] *of our CS-based cryptosystem is defined by*

$$\mathbf{\Phi} = \frac{1}{\sqrt{M}}\mathbf{R}_\Omega\mathbf{U} = \frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1 diag(\mathbf{s})\mathbf{U}_2. \quad (1)$$

In (1), $\mathbf{R}_\Omega$ *is a public random subsampling operator that selects* M *rows out of* N *ones uniformly at random, where the selected indices are specified by* $\Omega = \{\omega_0, \cdots, \omega_{M-1}\}$. *Also,* $\mathbf{U}_i \in \mathbb{R}^{N\times N}$ *is a unitary matrix, i.e.,* $\mathbf{U}_i^T\mathbf{U}_i = \mathbf{U}_i\mathbf{U}_i^T = N\mathbf{I}$ *for* $i = 1$ *and 2, respectively. In particular, each entry of* $\mathbf{U}_1$ *has unit magnitude, i.e.,* $|\mathbf{U}_1(k,t)| = 1$ *for all* $0 \le k,t \le N-1$. *Finally,* $\mathbf{U} = \frac{1}{\sqrt{N}}\mathbf{U}_1 diag(\mathbf{s})\mathbf{U}_2$ *is also unitary for* $\mathbf{s} \in \{-1,+1\}^N$, *where* $\mathbf{s}$ *is a secret keystream to be embedded in* $\mathbf{\Phi}$ *for each CS encryption.*

In this paper, we use $\mathbf{U}_1 = \mathbf{H}$, or an $N \times N$ Hadamard matrix that employs a binary m-sequence [40] of period $N - 1 = 2^n - 1$ for a positive integer n, i.e., $\mathbf{d} = (d_0, \cdots, d_{2^n-2})$, where $d_k \in \{0, 1\}$. For $0 \le k, t \le N-1$, each entry of $\mathbf{H}$ is given by

$$\mathbf{H}(k,t) = \begin{cases} 1, & \text{if } k = 0 \text{ or } t = 0, \\ (-1)^{d_{k+t-2}}, & \text{otherwise,} \end{cases}$$

where the index $k+t-2$ is computed modulo 2^n-1. From the structure, $\mathbf{H}$ is symmetric, or $\mathbf{H}^T = \mathbf{H}$. As $\mathbf{d}$ has the *ideal two-level autocorrelation* [40], i.e.,

$$\sum_{k=0}^{2^n-2}(-1)^{d_k+d_{k+\tau}} = \begin{cases} 2^n-1, & \text{if } \tau = 0, \\ -1, & \text{if } 1 \le \tau \le 2^n-2, \end{cases}$$

where $k+\tau$ is computed modulo 2^n-1, it is obvious that $\mathbf{H}\mathbf{H}^T = \mathbf{H}^T\mathbf{H} = N\mathbf{I}$. Since $\mathbf{H}$ is public, the structure and the initial state of an n-stage linear feedback shift register (LFSR) generating the binary m-sequence $\mathbf{d}$ are publicly known.

3.2 Keystream generation for CS encryption

In the sensing matrix $\mathbf{\Phi}$ of (1), we assume that $\mathbf{s}$ is a segment of length N from the original keystream of an extremely long period, which enables to renew the keystream $\mathbf{s}$ at each CS encryption. For fast and efficient keystream generation, one may employ an LFSR-based nonlinear keystream generator of stream ciphers. For example, we may consider the combinatorial sequence generator [41], the filtering sequence generator [42], the clock-controlled generator [43, 44], the shrinking generator [45], and the self-shrinking generator (SSG) [46], each of which presents a simple structure but a remarkable resistance against various attacks. For more details on keystream generators and stream ciphers, see [47] and [48]. Regarding the keystream of our CS-based cryptosystem, we make the following assumption.

Assumption 1 *An original keystream from a stream cipher is designed to have nice pseudorandomness properties [40] such as balance, large period, low autocorrelation, and large linear complexity. With the properties, we assume that each element of the keystream* $\mathbf{s}$ *takes* $+1$ *or* -1 *independently and uniformly at random, which facilitates the security analysis of our CS-based cryptosystem.*

When we employ a keystream generator to produce the keystream $\mathbf{s}$, the initial seed (or state) of the generator is essentially the *key* of our CS-based cryptosystem. The key should be kept secret between a sender and a legitimate recipient, whereas the structure of the keystream generator can be publicly known. For secure key exchange, we may establish a separate secure channel, or use the key establishment via the RSSI from wireless channels as in [35].

3.3 CS decryption

For CS decryption, a noisy ciphertext $\mathbf{r} = \mathbf{\Phi x} + \mathbf{n} \in \mathbb{R}^M$ is available for an adversary as well as a legitimate recipient, where $\mathbf{n} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$ is a measurement noise. A legitimate recipient of the ciphertext $\mathbf{r}$, who knows $\mathbf{\Phi}$, attempts to recover the plaintext $\mathbf{x}$ by conducting a CS reconstruction. Meanwhile, an adversary will make various attempts to recover the plaintext $\mathbf{x}$ or the keystream $\mathbf{s}$, with no knowledge of $\mathbf{\Phi}$.

Proposition 1 presents the reliability and the stability of our CS-based cryptosystem for a legitimate recipient, which is from the RIP result [49, 50] of a partial unitary sensing matrix.

Proposition 1 [49, 50] *For a legitimate recipient, our CS-based cryptosystem theoretically guarantees a stable decryption of a K-sparse plaintext with bounded errors, as long as* $M = \mathcal{O}\left(\mu^2(\mathbf{U}) \cdot K \log^4 N\right)$.

When $\mathbf{U}_1 = \mathbf{H}$, numerical experiments revealed that $\mu(\mathbf{U}) = \mathcal{O}\left(\sqrt{\log N}\right)$ for i) $\mathbf{U}_2 = \mathbf{W}$ or ii) $\mathbf{U}_2 = \mathbf{D}$, if each entry of the keystream $\mathbf{s}$ takes $+1$ or -1 uniformly at random. In this case, if $M = \mathcal{O}\left(K \log^5 N\right)$, Proposition 1 guarantees a stable decryption.

Table 2 summarizes a symmetric-key CS-based cryptosystem proposed in this paper.

4 Security analysis

A CS-based cryptosystem cannot be perfectly secure [15] but is believed to be *computationally secure* [15, 16]. In this section, we analyze the computational security of our CS-based cryptosystem by studying the notion of indistinguishability [19].

Assume that a cryptosystem produces a ciphertext by encrypting one of two possible plaintexts. The cryptosystem is said to have the *indistinguishability*, if no adversary can determine in polynomial time which of the two plaintexts corresponds to the ciphertext, with probability significantly better than that of a random guess [51]. In short, if a cryptosystem has the indistinguishability, an adversary is unable to learn any partial information of the plaintext in polynomial time from a given ciphertext.

In specific, let us consider an *indistinguishability experiment* [51] with a constraint of K-sparse plaintexts. First of all, an adversary creates a pair of plaintexts $\mathbf{x}_1$ and $\mathbf{x}_2$ with at most K nonzero entries per each. Then, our CS-based cryptosystem produces a ciphertext $\mathbf{r} = \mathbf{\Phi x}_h + \mathbf{n}$ by randomly selecting h, where $h = 1$ or 2. Given $\mathbf{r}$, the adversary attempts to figure out which plaintext, $\mathbf{x}_1$ or $\mathbf{x}_2$, was encrypted for the ciphertext, by carrying out a polynomial time test $\mathcal{D} : \mathbf{r} \to h \in \{1, 2\}$.

In this paper, we make use of the total variation (TV) distance [18] to evaluate the performance of the indistinguishability experiment. Let $d_{\mathrm{TV}}(p_1, p_2)$ be the TV distance between the probability distributions $p_1 = \Pr(\mathbf{r}|\mathbf{x}_1)$ and $p_2 = \Pr(\mathbf{r}|\mathbf{x}_2)$. Then, it is readily checked from [52] that the probability that an adversary can successfully distinguish the plaintexts by some kind of the binary hypothesis test $\mathcal{D}$ is bounded by

$$p_d \leq \frac{1}{2} + \frac{d_{\mathrm{TV}}(p_1, p_2)}{2}. \tag{2}$$

Therefore, if $d_{\mathrm{TV}}(p_1, p_2)$ approaches to zero, the probability of success will be at most that of a random guess, which leads to the indistinguishability of a cryptosystem. Consequently, one can argue that a cryptosystem with $d_{\mathrm{TV}}(p_1, p_2)$ closer to zero would be more secure in terms of the indistinguishability. Since computing $d_{\mathrm{TV}}(p_1, p_2)$ directly is difficult [53], we compute two probability metrics instead to bound the TV distance, which ultimately examines the indistinguishability of our CS-based cryptosystem.

Table 2 Symmetric-key CS-based cryptosystem

Public	Subsampling operator $\mathbf{R}_\Omega$
	Unitary matrices $\mathbf{U}_1$ and $\mathbf{U}_2$
	Structure of a keystream generator
Secret	Initial seed (or state) $\mathbf{k} \in \{0, 1\}^L$ of a keystream generator
Keystream generation	With the initial seed $\mathbf{k}$, a keystream $\mathbf{s} \in \{-1, +1\}^N$ is generated. The keystream $\mathbf{s}$ is renewed at each encryption.
CS encryption	With the keystream $\mathbf{s}$ and a plaintext $\mathbf{x} \in \mathbb{R}^N$, a ciphertext is generated by $\mathbf{r} = \mathbf{\Phi x} + \mathbf{n} \in \mathbb{R}^M$, where $\mathbf{\Phi} = \frac{1}{\sqrt{MN}} \mathbf{R}_\Omega \mathbf{U}_1 \mathrm{diag}(\mathbf{s}) \mathbf{U}_2$ and $\mathbf{n}$ is a measurement noise.
CS decryption	The plaintext $\mathbf{x}$ is reconstructed by a CS recovery algorithm with the knowledge of $\mathbf{s}$.

4.1 Relative entropy

In [23] and [24], the *relative entropy* (or the Kullback-Leibler divergence [20]) has been used to quantify the indistinguishability. Precisely, the relative entropy of two probability distributions gives an upper bound on the TV distance by Pinsker's inequality [54] or the refinements [55], which ultimately bounds the success probability of the indistinguishability experiment by (2).

In (1), one may assume that the entries of $\mathbf{\Phi}$ are asymptotically Gaussian for a sufficiently large N, since each one can be seen as the sum of independent random variables weighted by each entry of $\mathbf{s}$. Along with the Gaussian noise $\mathbf{n}$, we assume that $\mathbf{r}$, conditioned on $\mathbf{x}_1$ (or $\mathbf{x}_2$), is a jointly Gaussian random vector. Also, $\mathbb{E}[\mathbf{\Phi}] =$

$\frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1 \cdot \mathbb{E}[\mathrm{diag}(\mathbf{s})] \cdot \mathbf{U}_2 = \mathbf{0}$ for a given $\mathbf{R}_\Omega$, as each entry of $\mathbf{s}$ takes ± 1 with probability 1/2 under Assumption 1. Thus, $\mathbb{E}[\mathbf{r}|\mathbf{x}_h] = \mathbb{E}[\mathbf{\Phi}] \cdot \mathbf{x}_h + \mathbb{E}[\mathbf{n}] = \mathbf{0}$. With the Gaussian random vector $\mathbf{r}$, the relative entropy between $p_1 = \Pr(\mathbf{r}|\mathbf{x}_1)$ and $p_2 = \Pr(\mathbf{r}|\mathbf{x}_2)$ has the following closed-form expression [56]

$$D(p_1||p_2) = \frac{1}{2}\left[\log\frac{|\mathbf{C}_2|}{|\mathbf{C}_1|} + \mathrm{tr}\left(\mathbf{C}_2^{-1}\mathbf{C}_1\right) - M\right], \quad (3)$$

where $\mathbf{C}_1$ and $\mathbf{C}_2$ are the covariance matrices of $\mathbf{r}$ conditioned on $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. By measuring the relative entropy by (3), we obtain an upper bound on the TV distance, i.e.,

$$d_{\mathrm{TV}}(p_1, p_2) \le \min\left(\sqrt{\frac{D(p_1||p_2)}{2}},\ 1\right) \quad (4)$$

by Pinsker's inequality. In (4), the upper bound is set to be at most 1, since $d_{\mathrm{TV}}(p_1, p_2) \in [0, 1]$.

In what follows, we present an upper bound on the relative entropy with some constraints on plaintexts, which subsequently yields an analytic upper bound on the maximum TV distance by (4).

Theorem 1 *In our CS-based cryptosystem, assume that each plaintext* $\mathbf{x}$ *has at most K nonzero entries with the constant energy* $\mathcal{E}_x = ||\mathbf{x}||^2$*. Then, the relative entropy of (3) is bounded by*

$$D(p_1||p_2) \le \frac{M}{2}\left(K\mu^2(\mathbf{U}_2)\cdot \mathrm{PNR} - \log\left(K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+1\right)\right), \quad (5)$$

where $\mathrm{PNR} = \frac{\mathcal{E}_x}{M\sigma^2}$ *is the plaintext-to-noise power ratio (PNR).*

Proof See the Appendix. □

In Theorem 1, $\mu(\mathbf{U}_2) = 1$ if $\mathbf{U}_2 = \mathbf{W}$, while $\mu(\mathbf{U}_2) = \sqrt{2}$ if $\mathbf{U}_2 = \mathbf{D}$. However, if $\mathbf{U}_2 = \sqrt{N}\mathbf{I}$, the upper bound increases as N for $\mu(\mathbf{U}_2) = \sqrt{N}$. Thus, Theorem 1 implies that one must not use $\mathbf{U}_2 = \sqrt{N}\mathbf{I}$, to achieve the indistinguishability of our CS-based cryptosystem.

To ensure a reliable CS decryption for a legitimate recipient, our CS-based cryptosystem can set $K = \mathcal{O}\left(\frac{M}{\mu^2(\mathbf{U})\log N}\right)$ for nonuniform CS recovery [57], which yields the following corollary.

Corollary 1 *In our CS-based cryptosystem with* $\mathbf{U}_1 = \mathbf{H}$ *and* $N = 2^n$*, assume* $\mathbf{U}_2 = \mathbf{W}$ *or* $\mathbf{D}$*, where* $\mu(\mathbf{U}) = \mathcal{O}(\sqrt{\log N})$*. In Theorem 1, if* $K \le \frac{cM}{n^2}$ *with a constant c, then*

$$D(p_1||p_2) \le \frac{M}{2}\left(\frac{cM\mu^2(\mathbf{U}_2)}{n^2}\cdot \mathrm{PNR} - \log\left(\frac{cM\mu^2(\mathbf{U}_2)}{n^2}\cdot\mathrm{PNR}+1\right)\right).$$

Thus, if the keystream length N is sufficiently large with given M and PNR*, our CS-based cryptosystem will have low relative entropy, which contributes to the indistinguishability against an adversary, while guaranteeing the reliability for a legitimate recipient.*

4.2 Hellinger distance

To bound the TV distance, we may use another probability metric, the *Hellinger* distance [21]. In our CS-based cryptosystem, recall that the ciphertext $\mathbf{r}$, conditioned on $\mathbf{x}_h$, is assumed to be a jointly Gaussian random vector with zero mean and the covariance matrix $\mathbf{C}_h$, where $h = 1$ or 2. Then, the Hellinger distance for the multivariate Gaussian distributions p_1 and p_2 is given by [58, 59]

$$d_{\mathrm{H}}(p_1, p_2) = \sqrt{1 - \frac{|\mathbf{C}_1|^{\frac{1}{4}}|\mathbf{C}_2|^{\frac{1}{4}}}{|\mathbf{C}_3|^{\frac{1}{2}}}}, \quad (6)$$

where $\mathbf{C}_3 = \frac{\mathbf{C}_1+\mathbf{C}_2}{2}$. The Hellinger distance is particularly useful by giving both upper and lower bounds on the TV distance [60], i.e.,

$$d_{\mathrm{H}}^2(p_1, p_2) \le d_{\mathrm{TV}}(p_1, p_2) \le d_{\mathrm{H}}(p_1, p_2)\sqrt{2 - d_{\mathrm{H}}^2(p_1, p_2)}. \quad (7)$$

In what follows, we present an upper bound on the Hellinger distance of (6), which leads to an analytic upper bound on the maximum TV distance by (7).

Theorem 2 *Recall the assumptions and definitions of Theorem 1. In our CS-based cryptosystem, the Hellinger distance of (6) is bounded by*

$$d_{\mathrm{H}}(p_1, p_2) \le \sqrt{1 - \left(\frac{2\sqrt{K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+1}}{K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+2}\right)^{\frac{M}{4}}}, \quad (8)$$

where $\mathrm{PNR} = \frac{\mathcal{E}_x}{M\sigma^2}$.

Proof See the Appendix. □

Corollary 2 *In our CS-based cryptosystem with* $\mathbf{U}_1 = \mathbf{H}$ *and* $N = 2^n$*, assume* $\mathbf{U}_2 = \mathbf{W}$ *or* $\mathbf{D}$*, where* $\mu(\mathbf{U}) = \mathcal{O}\left(\sqrt{\log N}\right)$*. In Theorem 2, if* $K \le \frac{cM}{n^2}$ *with a constant c, then*

$$d_{\mathrm{H}}(p_1, p_2) \le \sqrt{1 - \left(\frac{2n\sqrt{cM\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+n^2}}{cM\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+2n^2}\right)^{\frac{M}{4}}}.$$

Thus, if the keystream length N is sufficiently large with given M and PNR*, our CS-based cryptosystem will have low Hellinger distance, which contributes to the indistinguishability against an adversary, while guaranteeing the reliability for a legitimate recipient.*

Remark 1 *Theorems* 1 *and* 2 *suggest that the relative entropy and the Hellinger distance will approach to zero as PNR decreases. Accordingly, our CS-based cryptosystem will have low TV distance by* (4) *and* (7) *at low PNR. Similarly, the TV distance will be low when M and K are small, respectively. Consequently, our CS-based cryptosystem can be indistinguishable at low PNR for small M and K.*

Remark 2 *When* $N = 2^n$ *increases, Corollaries* 1 *and* 2 *suggest that if M is fixed, the relative entropy and the Hellinger distance will decrease at a given* PNR *by reducing* $K = \mathcal{O}\left(\frac{M}{n^2}\right)$*, which will be confirmed by numerical results of Section 5. On the other hand, if M increases with* $M = \mathcal{O}\left(Kn^2\right)$ *for a given K, numerical results reveal that they also decrease over N at a given* PNR*, which contradicts Theorems* 1 *and* 2*. This observation implies that there is a room to improve the bounds of the theorems. Combined with Remark* 1*, the TV distance will be low if the keystream length N is sufficiently large with low compression* $\left(\frac{M}{N}\right)$ *and sparsity* $\left(\frac{K}{N}\right)$ *ratios, which leads to the asymptotic indistinguishability of our CS-based cryptosystem.*

5 Potential key recovery attack

In this section, we consider a potential key recovery attack in which an adversary attempts to recover the key of our CS-based cryptosystem. In the CPA, the adversary tries to restore a keystream from a ciphertext (stage 1) and then to recover the original key from the restored keystream via algebraic cryptanalysis (stage 2). With a sufficiently long key, we assume that the number of keystream bits required for the algebraic cryptanalysis, denoted by D, is much larger than the ciphertext length M. For a convenience of analysis, we assume $D = N$, which means that the adversary needs to restore a keystream of full length N from stage 1. Figure 1 illustrates the potential CPA from an adversary for key recovery. This section discusses the adversary's strategy for keystream recovery in stage 1. Once a keystream is successfully restored through stage 1, a known cryptanalysis [47, 48] can be carried out in stage 2 for key recovery, which will not be discussed in this paper.

5.1 Mathematical intractability of keystream recovery

In stage 1 of the CPA, an adversary needs to observe a correct N-bit keystream from a ciphertext that has been encrypted by a chosen plaintext. We assume that the adversary will choose a plaintext $\mathbf{x}$ such that each entry of $\widehat{\mathbf{x}} = \mathbf{U}_2\mathbf{x}$ is nonzero for a unitary matrix $\mathbf{U}_2$. Then, the corresponding ciphertext is given by

$$\begin{aligned}\mathbf{r} = \mathbf{\Phi x} + \mathbf{n} &= \frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1\mathrm{diag}(\mathbf{s})\mathbf{U}_2\mathbf{x} + \mathbf{n} \\ &= \frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1\mathrm{diag}(\widehat{\mathbf{x}})\mathbf{s} + \mathbf{n} \\ &= \mathbf{As} + \mathbf{n},\end{aligned} \tag{9}$$

where $\mathbf{A} = \frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1\mathrm{diag}(\widehat{\mathbf{x}})$. Unlike in stream ciphers, restoring the keystream $\mathbf{s}$ from the known plaintext-ciphertext pair is not a trivial task, since $\mathbf{s}$ is hidden under compression in $\mathbf{r}$.

From the ciphertext $\mathbf{r}$ of (9), an adversary needs to find a most likely keystream, which is equivalent to a *maximum-likelihood* (*ML*) estimate of

$$\widehat{\mathbf{s}} = \underset{\mathbf{s}\in\{-1,+1\}^N}{\mathrm{argmin}} \; ||\mathbf{r} - \mathbf{As}||^2. \tag{10}$$

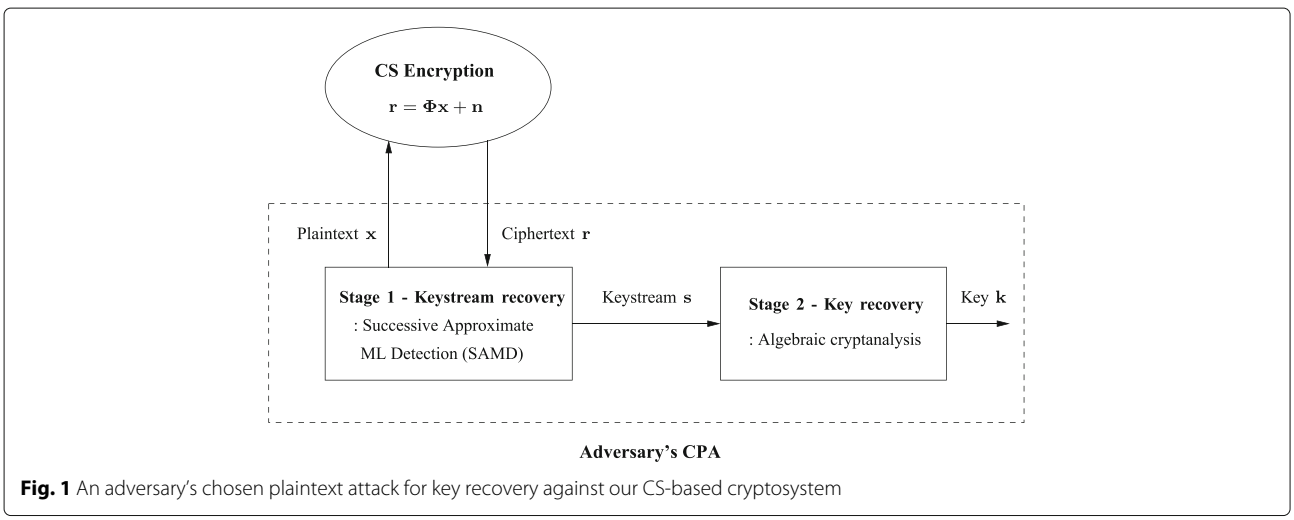

Fig. 1 An adversary's chosen plaintext attack for key recovery against our CS-based cryptosystem

Finding the ML solution of (10) is known as a constrained *integer least-squares (ILS)* problem, which is also called a *closest vector problem* (*CVP*) [61] in lattices. For a general $\mathbf{A}$, the constrained ILS problem is proven to be NP hard [62].

To find a most likely keystream of (10), an exhaustive ML search requires the complexity of $\mathcal{O}\left(2^N\right)$, which would be computationally infeasible if the keystream length N is sufficiently large. Alternatively, the generalized sphere decoding (GSD) algorithms [63–65] can find an ML solution to the ILS problem of the *underdetermined* system with $M < N$. However, as it has the complexity exponential in $N - M$ [63–65], the GSD cannot be applicable to the ILS problem with $M \ll N$. To the best of our knowledge, there is no polynomial-time algorithm to find an ML solution of (10) with $M \ll N$ for a sufficiently large N.

In summary, the computational security of our CS-based cryptosystem against the key recovery attack is brought by the mathematical hardness that no polynomial-time algorithm is known to find an ML solution to the underdetermined ILS problem. In fact, the mathematical intractability of the ILS problem has been exploited by public-key cryptosystems [66–68]. In our symmetric-key CS-based cryptosystem, it also ensures that if the keystream length N is sufficiently large with $M \ll N$, no adversary will be able to find a most likely keystream of length N in polynomial time, which demonstrates the computational security of our CS-based cryptosystem against the key recovery attack.

5.2 Successive approximate maximum-likelihood detection (SAMD)

In Section 5.1, we demonstrated that the ML detection would be infeasible for keystream recovery, as long as the keystream length is sufficiently large. As an alternative, we consider a sub-optimal, but feasible keystream recovery process for the CPA. Instead of restoring an N-bit keystream at once, we assume that an adversary attempts to restore a disjoint J-bit segment[2] of the keystream from each detection, where $J \ll N$, and repeats the detection $\lceil \frac{N}{J} \rceil$ times successively to restore the keystream of full length N. In this subsection, we describe the details of the successive detection process for keystream recovery.

For a convenience of analysis, we assume a chosen plaintext such that $\widehat{\mathbf{x}} = \left(\sqrt{MN}, \cdots, \sqrt{MN}\right)^T$ in (9), which yields $\mathbf{A} = \mathbf{R}_\Omega \mathbf{U}_1$ for our analysis[3]. In the keystream recovery, an adversary has a freedom to choose the value of J and the J-bit positions of a keystream to be restored at the ith detection. Let $\Theta_i \subset \{0, \cdots, N-1\}$ be a set of indices, where $|\Theta_i| = J$ if $1 \le i \le n_s - 1$ and $|\Theta_i| = N - (n_s - 1)J$ if $i = n_s$, respectively, for $n_s = \lceil \frac{N}{J} \rceil$. Also, $\Theta_a \cap \Theta_b = \phi$ for $a \ne b$, where ϕ is an empty set, and $\Theta_1 + \cdots + \Theta_{n_s} = \{0, \cdots, N-1\}$.

Let $\mathbf{s}_{\Theta_i} \in \{-1, +1\}^{|\Theta_i|}$ be a $|\Theta_i|$-bit vector, where the entries are taken from the indices of Θ_i in the keystream $\mathbf{s}$. At the ith detection, an adversary attempts to find $\mathbf{s}_{\Theta_i}$ from the ciphertext $\mathbf{r}$ of (9). With $\mathbf{s}_{\Theta_1}, \cdots, \mathbf{s}_{\Theta_{i-1}}$ that have been detected from the previous detections, the ith detection should use a new ciphertext $\mathbf{r}_i$ by subtracting their contribution from $\mathbf{r}$, i.e.,

$$\mathbf{r}_i = \mathbf{r} - \sum_{h=1}^{i-1} \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Theta_h}^T \widehat{\mathbf{s}}_{\Theta_h}, \tag{11}$$

where $\widehat{\mathbf{s}}_{\Theta_h}$ is an estimate from the hth detection. In (11), $\mathbf{R}_{\Theta_h}^T$ is an $N \times J$ column selection operator that selects J columns of $\mathbf{U}_1$ whose indices are specified by Θ_h. Let $\Delta_i = \{0, \cdots, N-1\} \setminus (\Theta_1 + \cdots + \Theta_i)$, where $\Delta_{n_s} = \phi$, and $\mathbf{R}_{\Delta_i}^T$ be an $N \times (N - iJ)$ column selection operator whose indices are specified by Δ_i. By assuming $\widehat{\mathbf{s}}_{\Theta_h} = \mathbf{s}_{\Theta_h}$ for $1 \le h \le i-1$, we have from (11)

$$\begin{aligned} \mathbf{r}_i &= \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Theta_i}^T \mathbf{s}_{\Theta_i} + \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Delta_i}^T \mathbf{s}_{\Delta_i} + \mathbf{n} \\ &= \mathbf{m}_i + \mathbf{w}_i + \mathbf{n}, \end{aligned} \tag{12}$$

where $\mathbf{m}_i = \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Theta_i}^T \mathbf{s}_{\Theta_i}$ corresponds to a desired component to be detected at the ith detection, $\mathbf{w}_i = \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Delta_i}^T \mathbf{s}_{\Delta_i}$ is an interfering component from the keystream segments that have not been detected yet, and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ is a Gaussian random noise.

In (12), $\mathbf{w}_{n_s} = \mathbf{0}$ since $\Delta_{n_s} = \phi$. On the other hand, if $1 \le i \le n_s - 1$, each entry of $\mathbf{w}_i$ is taken from the sum of $N - iJ$ column vectors of $\mathbf{R}_\Omega \mathbf{U}_1$, each of which is weighted by the entry of $\mathbf{s}_{\Delta_i}$. Since each entry of $\mathbf{s}_{\Delta_i}$ takes $+1$ or -1 randomly and independently under Assumption 1, $\mathbf{w}_i$ will follow the jointly Gaussian distribution by the central limit theorem [69]. By noting that $\mathbf{w}_i + \mathbf{n}$ can be modeled as a Gaussian random vector for $1 \le i \le n_s$, $\mathbf{r}_i$ is also Gaussian for a given $\mathbf{s}_{\Theta_i}$. Then,

$$\begin{aligned} \mathbb{E}\left[\mathbf{r}_i | \mathbf{s}_{\Theta_i}\right] &= \mathbb{E}\left[\mathbf{m}_i | \mathbf{s}_{\Theta_i}\right] + \mathbb{E}\left[\mathbf{w}_i | \mathbf{s}_{\Theta_i}\right] + \mathbb{E}\left[\mathbf{n} | \mathbf{s}_{\Theta_i}\right] \\ &= \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Theta_i}^T \mathbf{s}_{\Theta_i} = \mathbf{m}_i, \end{aligned} \tag{13}$$

where $\mathbb{E}\left[\mathbf{w}_i | \mathbf{s}_{\Theta_i}\right] = \mathbb{E}\left[\mathbf{n} | \mathbf{s}_{\Theta_i}\right] = \mathbf{0}$, since $\mathbf{s}_{\Theta_i}$ is independent of $\mathbf{w}_i$ and $\mathbf{n}$, respectively. Also, the covariance of $\mathbf{r}_i$ is given by

$$\mathbb{E}\Big[(\mathbf{r}_i - \mathbf{m}_i)(\mathbf{r}_i - \mathbf{m}_i)^T | \mathbf{s}_{\Theta_i}\Big] = \mathbb{E}\Big[(\mathbf{w}_i + \mathbf{n})(\mathbf{w}_i + \mathbf{n})^T\Big] = \mathbf{K}_i + \sigma^2 \mathbf{I}, \tag{14}$$

where $\mathbf{w}_i$ and $\mathbf{n}$ are independent. In (14),

$$\begin{aligned} \mathbf{K}_i = \mathbb{E}\left[\mathbf{w}_i \mathbf{w}_i^T\right] &= \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Delta_i}^T \cdot \mathbb{E}\left[\mathbf{s}_{\Delta_i} \mathbf{s}_{\Delta_i}^T\right] \cdot \mathbf{R}_{\Delta_i} \mathbf{U}_1^T \mathbf{R}_\Omega^T \\ &= \mathbf{R}_\Omega \mathbf{U}_1 \mathbf{R}_{\Delta_i}^T \cdot \mathbf{R}_{\Delta_i} \mathbf{U}_1^T \mathbf{R}_\Omega^T, \end{aligned} \tag{15}$$

where $\mathbb{E}\left[\mathbf{s}_{\Delta_i}\mathbf{s}_{\Delta_i}^T\right] = \mathbf{I}$. Since $\mathbf{K}_i$ does not depend on $\mathbf{s}_{\Theta_i}$, the covariance of $\mathbf{r}_i$ in (14) is equal for all possible $\mathbf{s}_{\Theta_i} \in \{-1,+1\}^{|\Theta_i|}$ at each ith detection. Under the Gaussian model of $\mathbf{r}_i$ with equal covariance, we can apply the ML decision rule [70] at the ith detection, which yields

$$\widehat{\mathbf{s}}_{\Theta_i} = \underset{\mathbf{s}_{\Theta_i}\in\{-1,+1\}^{|\Theta_i|}}{\operatorname{argmin}} (\mathbf{r}_i - \mathbf{m}_i)^T \left(\mathbf{K}_i + \sigma^2\mathbf{I}\right)^{-1} (\mathbf{r}_i - \mathbf{m}_i). \tag{16}$$

In (11) and (12), we assumed that all the estimates $\widehat{\mathbf{s}}_{\Theta_h}$, $1 \le h \le i-1$, from the previous detections are correct, and then ignored the estimation errors $\mathbf{s}_{\Theta_h} - \widehat{\mathbf{s}}_{\Theta_h}$ while subtracting the contribution from $\mathbf{r}$. Therefore, (16) cannot be a true ML detection, but an optimistic approximation to the adversary.

Finally, the adversary carries out the approximate ML detection of (16) n_s times successively for $1 \le i \le n_s$ and restores the full N-bit keystream by combining the disjoint $|\Theta_i|$-bit estimates of $\widehat{\mathbf{s}}_{\Theta_i}$. Throughout this paper, the detection process is called a *successive approximate ML detection* (*SAMD*). In what follows, we present an upper bound on the success probability of the SAMD.

Theorem 3 *In the SAMD, recall the approximate ML decision rule of* (16) *applied at each ith detection for* $1 \le i \le n_s$, *where* $n_s = \lceil \frac{N}{J} \rceil$. *Let* $\lambda_{min}(\mathbf{K}_i)$ *be the minimum eigenvalue of the covariance matrix* $\mathbf{K}_i$ *in* (15). *Let* P_{succ} *be the probability that an N-bit keystream can be successfully restored by the SAMD. Then,*

$$P_{succ} \le \prod_{i=1}^{n_s} \left(1 - Q\left(\sqrt{\frac{M\mu^2(\mathbf{U}_1)}{\lambda_{min}(\mathbf{K}_i) + \sigma^2}}\right)\right) \triangleq P_{succ,UB}, \tag{17}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$.

Proof See the Appendix. □

Theorem 3 shows the result for a general unitary matrix $\mathbf{U}_1$, which suggests that our CS-based cryptosystem should choose an $N \times N$ unitary matrix $\mathbf{U}_1$ such that $\mu(\mathbf{U}_1)$ is as small as possible, regardless of N, in order to degrade the performance of the SAMD. In this paper, $\mu(\mathbf{U}_1) = 1$ from $\mathbf{U}_1 = \mathbf{H}$.

The upper bound on the success probability of Theorem 3 represents the highest possible performance that the SAMD can achieve with no estimation errors at each detection, which is an optimistic scenario for an adversary. In reality, the actual probability of success will be much lower than the upper bound, due to estimation errors and error propagation through detections. If an adversary finds a solution of (16) via an exhaustive search, the complexity of each detection of the SAMD will be $\mathcal{O}\left(2^J\right)$ with $J \ll N$.

5.3 Minimum eigenvalues of $\mathbf{K}_i$

Theorem 3 implies that minimizing $\lambda_{\min}(\mathbf{K}_i)$ can improve the performance of the SAMD. At the ith detection of the SAMD, it is an adversary that determines the selection operator $\mathbf{R}_{\Theta_i}$. Therefore, if the adversary appropriately chooses Θ_i (or equivalently Δ_i) to minimize $\lambda_{\min}(\mathbf{K}_i)$, the success probability of the SAMD can be improved. In this paper, we consider three possible selections for Θ_i that the adversary may choose reasonably.

1) Uniform selection: $\Theta_i = \left\{i-1,\ \lfloor\frac{N}{J}\rfloor + i - 1,\ \cdots,\ (J-1)\lfloor\frac{N}{J}\rfloor + i - 1\right\}$.
2) Consecutive selection: $\Theta_i = \{(i-1)J,\ (i-1)J+1,\ \cdots,\ iJ-1\}$.
3) Random selection: Θ_i selects the J indices from $\{0,\cdots,N-1\} \setminus (\Theta_1 + \cdots + \Theta_{i-1})$ uniformly at random.

Each selection is valid for $1 \le i \le n_s - 1$, and $\Theta_{n_s} = \{0,\cdots,N-1\} \setminus (\Theta_1 + \cdots + \Theta_{n_s-1})$, where $n_s = \lceil\frac{N}{J}\rceil$. To further minimize $\lambda_{\min}(\mathbf{K}_i)$, the adversary might be able to develop a more sophisticated selection of Θ_i by exploiting the structure of $\mathbf{R}_\Omega$ and $\mathbf{U}_1$. However, we leave this issue open for future research. Regarding the selection operator, we have the following assumption.

Assumption 2 *Once an adversary chooses a value of J and a type of selection, we assume that they will be fixed through the entire detections of the SAMD.*

Intuitively, the larger J will ensure better detection performance for the SAMD, since a longer keystream segment that can be subtracted from each detection may contribute less interference. The intuition will be justified by the numerical results of Section 6. In this regard, Assumption 2 is valid, since the adversary's reasonable option is to fix the value of J to the largest possible one allowed by the computing power. In addition, the numerical results of Section 6 show that $\lambda_{\min}(\mathbf{K}_i)$ is not so affected by the type of selections, which also supports Assumption 2.

In what follows, we present a theoretical lower bound on $\lambda_{\min}(\mathbf{K}_i)$ for $1 \le i \le n_s$, if Θ_i is a random selection.

Theorem 4 *In our CS-based cryptosystem with* $\mathbf{U}_1 = \mathbf{H}$, *assume that an adversary chooses a random selection for*

Θ_i in the ith detection of the SAMD, where $1 \le i \le n_s = \lceil \frac{N}{J} \rceil$. Let $I_T = \lceil \frac{N - c_2 M \log M}{J} \rceil$ for a constant $c_2 > 0$. Then,

$$\lambda_{\min}(\mathbf{K}_i) \ge \begin{cases} \left(\sqrt{N - iJ} - \sqrt{c_1 M \log M}\right)^2, & \text{if } i < I_T, \\ 0, & \text{if } i \ge I_T \end{cases} \tag{18}$$

with high probability, where c_1 is a constant with $0 < c_1 < c_2$.

Proof See the Appendix. □

The numerical results of Section 6 show that the lower bound also holds for uniform and consecutive selections. Using the bound, Corollary 3 presents a further upper bound on the success probability of the SAMD, which is straightforward from Theorems 3 and 4 with $\mu(\mathbf{H}) = 1$.

Corollary 3 *In our CS-based cryptosystem with $\mathbf{U}_1 = \mathbf{H}$, if an adversary chooses a random selection for Θ_i, $1 \le i \le n_s$ during the SAMD, $P_{succ,\,UB}$ in Theorem 3 is bounded by*

$$P_{\text{succ,UB}} \le \left(1 - Q\left(\sqrt{\frac{M}{\sigma^2}}\right)\right)^{n_s - I_T + 1} \cdot \prod_{i=1}^{I_T - 1} \left(1 - Q\left(\sqrt{\frac{M}{(\sqrt{N - iJ} - \sqrt{c_1 M \log M})^2 + \sigma^2}}\right)\right) \triangleq P_{\text{succ,U}^2\text{B}},$$

where $I_T = \lceil \frac{N - c_2 M \log M}{J} \rceil$ for constants c_1 and c_2 with $0 < c_1 < c_2$.

6 Numerical results

This section presents numerical results to demonstrate the reliability and the security of our CS-based cryptosystem. In numerical experiments, each plaintext $\mathbf{x}$ has at most K nonzero entries, where the positions are chosen uniformly at random and the coefficients are taken from the Gaussian distribution. In CS encryption, $\mathbf{\Phi} = \frac{1}{\sqrt{MN}} \mathbf{R}_\Omega \mathbf{U}_1 \text{diag}(\mathbf{s}) \mathbf{U}_2$, where $\mathbf{U}_1 = \mathbf{H}$, and $\mathbf{U}_2 = \mathbf{W}$ or $\mathbf{D}$. Also, the secret keystream $\mathbf{s}$ is generated by the *self-shrinking generator* [46] of a 128-stage LFSR. For CS decryption, the CoSaMP recovery algorithm [71] has been employed for a legitimate recipient to decrypt each ciphertext with the knowledge of $\mathbf{\Phi}$.

6.1 CS decryption of a legitimate recipient

Figure 2 demonstrates the performance of CS decryption of a legitimate recipient, where the plaintext length is $N = 1024$ and the ciphertext length is $M = 48$. The figure sketches the normalized mean squared error (NMSE), defined by $\text{NMSE} = \mathbb{E}\left[\frac{||\mathbf{x} - \widehat{\mathbf{x}}||^2}{||\mathbf{x}||^2}\right]$, where $\mathbf{x}$ and $\widehat{\mathbf{x}}$ are original and decrypted plaintexts, respectively. We examine the performance with total 10000 plaintexts at a given PNR , where each one has at most $K = 4$ nonzero entries. For comparison, we sketch the performance of CS reconstruction with a random Gaussian sensing matrix for $\mathbf{\Phi}$. The figure shows that the performance of our CS decryption is as good as that of CS recovery with a random Gaussian sensing matrix. As a consequence, it demonstrates that our CS-based cryptosystem guarantees a reliable CS decryption for a legitimate recipient.

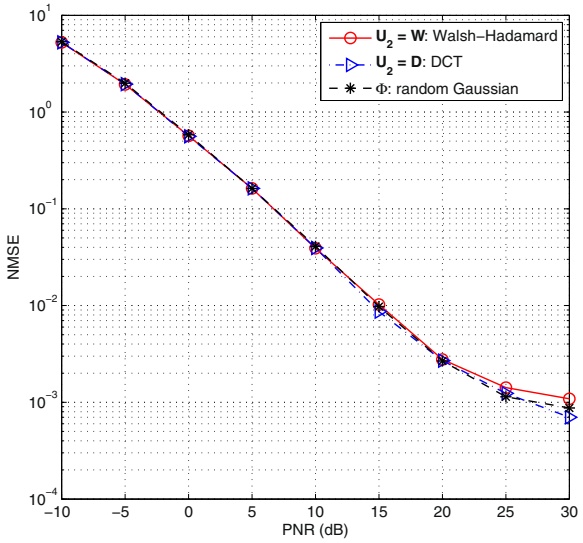

Fig. 2 The normalized mean squared error (NMSE) of CS decryption for a legitimate recipient

6.2 Indistinguishability

Figure 3 displays the upper and lower bounds of TV distance over PNR with $\mathbf{U}_2 = \mathbf{W}$, where $N = 1024$, $M = 48$, and $K = 4$. In the figure, the relative entropy of (3) and the Hellinger distance of (6) were computed using the covariance matrix of (19). Averaged over 10,000 pairs of randomly generated plaintexts $(\mathbf{x}_1, \mathbf{x}_2)$ with at most K nonzero entries per each, the relative entropy and the Hellinger distance yield the bounds of (4) and (7) on the TV distance, respectively. For comparison, we also sketch the theoretical upper bounds on the TV distance, which are obtained by the maximum relative entropy of (5) and the maximum Hellinger distance of (8), respectively. The figure shows that the TV distance approaches to zero as noise level grows, which implies that our CS-based cryptosystem can be indistinguishable at low PNR. As PNR increases, however, we observe that the upper and lower bounds increase and finally converge to certain levels, respectively. More extensive simulations agreed with the implication of Remark 1 that the CS-based cryptosystem will have lower TV distances with less PNR, M, and K. We made similar observations of the TV distance when $\mathbf{U}_2 = \mathbf{D}$ and/or each plaintext has bipolar nonzero entries.

Figure 4 depicts the upper bounds on the success probability of an adversary in the indistinguishability experiment, where the best- and worst-case upper bounds of (2) are from the minimum and maximum achievable TV distances of (7), respectively, obtained by the Hellinger distance (6). In the figure, $\mathbf{U}_2 = \mathbf{W}$ and PNR $= 25$ dB. With a given ciphertext length $M = 48$, the maximum sparsity is set as $K = \lfloor cM/\log_2^2 N \rfloor$ for each $N = 2^n$, to ensure a reliable nonuniform CS decryption for a legitimate recipient, where $c = 8.5$. For comparison, we sketch the empirical success probability of CS decryption by a legitimate recipient, where a decrypted plaintext has been declared as a success if $\frac{||\mathbf{x}-\widehat{\mathbf{x}}||^2}{||\mathbf{x}||^2} < 10^{-2}$. The figure reveals that the adversary's success probability approaches to that of a random guess as the keystream length N increases, while a legitimate recipient maintains its reliability.

Figure 5 also displays the upper bounds on the success probability of an adversary in the indistinguishability experiment. At this time, the ciphertext length is kept as $M = \lceil cK \log_2^2 N \rceil$ for each $N = 2^n$ with a given $K = 4$, where $c = 0.12$. As in Fig. 4, it also reveals that the adversary's success probability approaches to 0.5 as the keystream length N increases, while a legitimate recipient maintains its reliability. In conclusion, the empirical results of Figs. 4 and 5 show that if the keystream length N is sufficiently large with low compression $\left(\frac{M}{N}\right)$ and sparsity $\left(\frac{K}{N}\right)$ ratios, our CS-based cryptosystem can be computationally secure in terms of the indistinguishability, while guaranteeing a reliable CS decryption for a legitimate recipient.

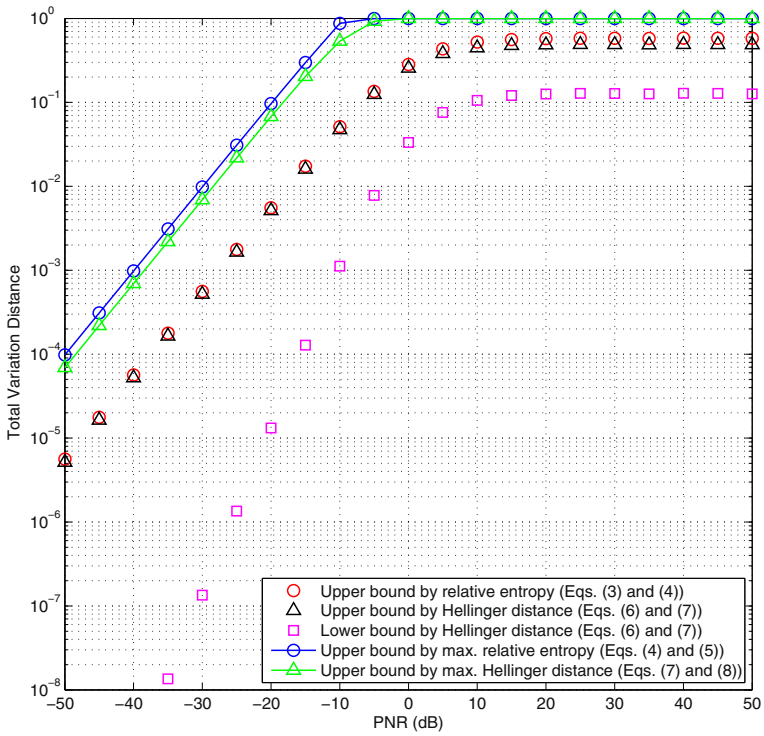

Fig. 3 The upper and lower bounds of total variation distance over PNR

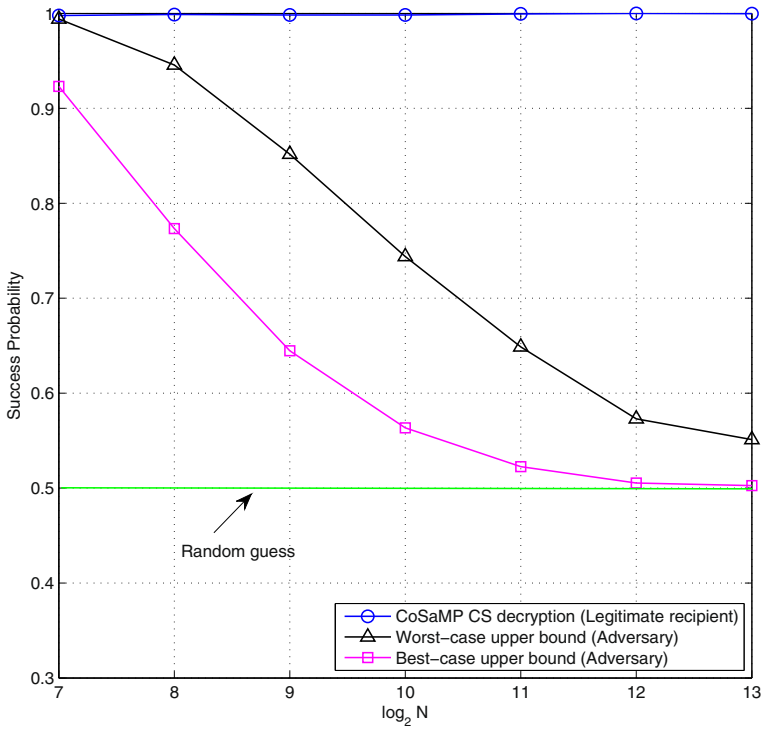

Fig. 4 The success probability of legitimate recipient and adversary for a given *M*

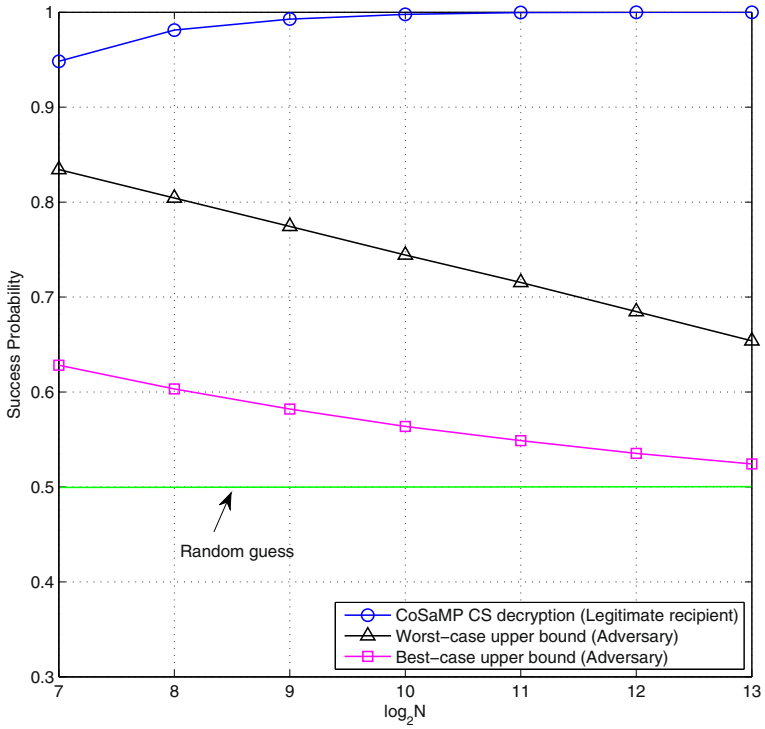

Fig. 5 The success probability of legitimate recipient and adversary for a given *K*

6.3 Performance of SAMD

Figure 6 sketches the minimum eigenvalues of the covariance matrix $\mathbf{K}_i$ of (15) at the ith detection for various $J \in \{32, 48, 64, 80\}$, where $N = 1024$ and $M = 48$. For comparison, it also sketches the lower bound of Theorem 4, where $c_1 = 0.5$ and $c_2 = 1$. For each i, we tested with 100,000 pairs of (Ω, Θ_i) for random subsampling and selection operators $\mathbf{R}_\Omega$ and $\mathbf{R}_{\Theta_i}$, where Θ_i had been fixed through the tested pairs in case of uniform and consecutive selections. In each subfigure, $\lambda_{\min}(\mathbf{K}_i)$ is sketched over $1 \le i \le n_s - 1$, where $n_s = \lceil \frac{N}{J} \rceil$. Figure 6 shows that if J increases, $\lambda_{\min}(\mathbf{K}_i)$ decreases faster over i, which suggests that the detection performance will be improved as J increases. It is plausible because if more keystream bits are detected from the ith detection with no estimation errors, more interfering components can be subtracted from the $(i+1)$th detection. In addition, it appears that the minimum eigenvalues are irrelevant to the types of Θ_i, which means that an adversary may expect no benefits from a particular selection of Θ_i. Finally, Fig. 6 demonstrates that the lower bound of Theorem 4 is valid, not only for random selection but also for uniform and consecutive selections.

Figure 7 displays the upper bounds on the success probability of the SAMD for keystream recovery. For comparison, it also sketches the theoretical upper bound of Corollary 3 for random selection Θ_i. In view of the adversary's bounded computing power, we set $J \le 128$, where the complexity of each detection in the SAMD will be $\mathcal{O}\left(2^J\right)$ by an exhaustive search. Since $\lambda_{\min}(\mathbf{K}_i)$ has similar values for different types of Θ_i's in Fig. 6, the upper bounds of Fig. 7 are also similar for every selection types. Moreover, the upper bounds increase over J, which is obvious from the sharp decline of $\lambda_{\min}(\mathbf{K}_i)$ over J, observed from Fig. 6. However, even if an adversary chooses a large value of J, the upper bounds on the success probability are still

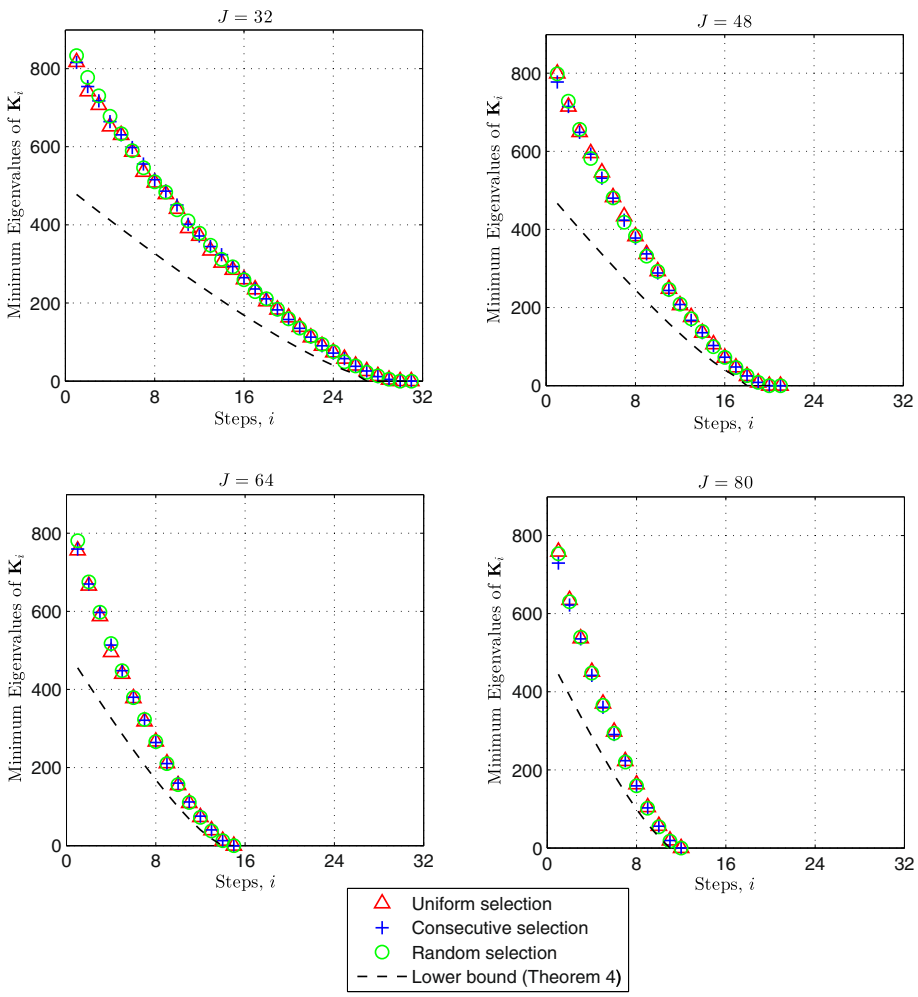

Fig. 6 The minimum eigenvalues of $\mathbf{K}_i$ at the ith detection of the SAMD

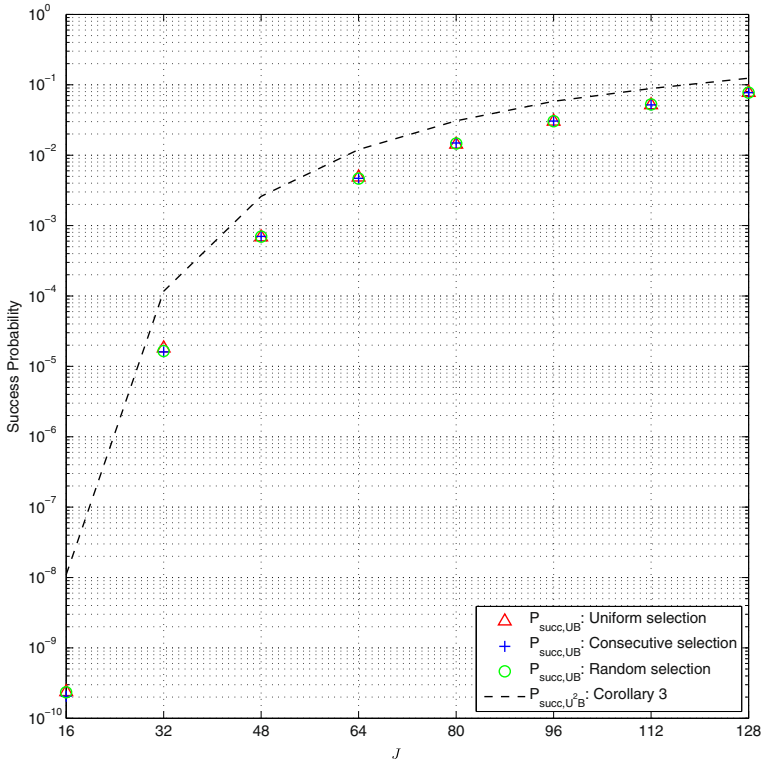

Fig. 7 The upper bounds on the success probability of the SAMD

significantly low, which implies that the potential of the SAMD to restore a correct N-bit keystream is pessimistic. Note that this is the result of an optimistic scenario, and in reality, the actual probability of success of the SAMD will be much lower than the upper bounds, due to estimation errors and their propagation through the SAMD.

7 Conclusions

This paper has proposed a CS-based cryptosystem that encrypts a plaintext with a partial unitary sensing matrix embedding a secret keystream. We demonstrated that our CS-based cryptosystem can offer a theoretically and empirically reliable decryption performance for a legitimate recipient, which is the first contribution of this paper. Then, we examined the indistinguishability of our CS-based cryptosystem by studying the TV distance as a security measure. To investigate the TV distance, we developed upper bounds on the relative entropy and the Hellinger distance, respectively. From the second contribution, we showed that our CS-based cryptosystem can be computationally secure in terms of the indistinguishability, as long as the keystream length for each encryption is sufficiently large with low compression and sparsity ratios.

In addition, we considered a potential CPA from an adversary to recover the key of our CS-based cryptosystem. The computational security of our CS-based cryptosystem against the CPA is based on the mathematical hardness that no polynomial-time algorithm is known to find an ML solution to the underdetermined ILS problem for keystream recovery. As a sub-optimal approach, we introduced the SAMD for an adversary to restore a secret keystream in polynomial time. In the third contribution, we developed an upper bound on the success probability of the SAMD and demonstrated that the performance of the keystream recovery through the SAMD is very pessimistic. In conclusion, our CS-based cryptosystem with a partial unitary sensing matrix embedding a secret keystream can be secure against the CPA, while guaranteeing a stable and robust decryption for a legitimate recipient.

Endnotes

[1] This paper assumes that a plaintext $\mathbf{x}$ is sparse in canonical basis, or $\mathbf{\Psi} = \mathbf{I}$. In general, if a plaintext $\mathbf{x}$ is sparse with respect to an arbitrary orthonormal basis $\mathbf{\Psi}$, i.e., $\mathbf{x} = \mathbf{\Psi}^T\boldsymbol{\theta}$, the sensing matrix $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}^T$ maintains the form of (1) by considering $\mathbf{U}_2\mathbf{\Psi}^T$ as a new unitary matrix $\mathbf{U}_2$.

[2] In the last detection, $(N - (\lceil \frac{N}{J} \rceil - 1)J)$-bit segment will be restored, where $\lceil \frac{N}{J} \rceil$ denotes the nearest integer greater than or equal to $\frac{N}{J}$.

[3] Under this assumption, numerical results showed that the upper bound on the success probability of the successive detection is more favorable for an adversary than that of $\widehat{\mathbf{x}}$ with arbitrary nonzero entries.

Appendices

Proof of Theorem 1

We give a brief sketch for the proof of Theorem 1, as the underlying technique is similar to that of Theorem 1 in [72]. Similar to Lemma 1 of [72], the covariance matrix of $\mathbf{r}$ is given by

$$\mathbf{C}_h = \mathbb{E}\left[\mathbf{r}\mathbf{r}^T|\mathbf{x}_h\right] = \mathbf{R}_\Omega \widetilde{\mathbf{C}}_h \mathbf{R}_\Omega^T + \sigma^2\mathbf{I}, \tag{19}$$

where $\widetilde{\mathbf{C}}_h = \frac{1}{N}\mathbf{U}_1^T \mathrm{diag}\left(\frac{|\widehat{\mathbf{x}}_h|^2}{M}\right)\mathbf{U}_1$ for $\widehat{\mathbf{x}}_h = \mathbf{U}_2\mathbf{x}_h$. Let $\lambda_1(\mathbf{C}_h) \geq \cdots \geq \lambda_M(\mathbf{C}_h)$ be the eigenvalues of $\mathbf{C}_h$, while $\lambda_1(\widetilde{\mathbf{C}}_h) \geq \cdots \geq \lambda_N\left(\widetilde{\mathbf{C}}_h\right)$ be the eigenvalues of $\widetilde{\mathbf{C}}_h$. With $\widehat{\mathbf{x}}_h = \mathbf{U}_2\mathbf{x}_h = \left(\widehat{x}_{h,0}, \cdots, \widehat{x}_{h,N-1}\right)^T$, let $\mathbf{v}_h = \left(v_{h,0}, \cdots, v_{h,N-1}\right)^T$, where $v_{h,k} = |\widehat{x}_{h,\pi(k)}|^2$ for $k = 0, \cdots, N-1$, and $\pi(k)$ is a permutation for $v_{h,0} \geq \cdots \geq v_{h,N-1}$. From the definition of $\widetilde{\mathbf{C}}_h$, it is clear that $\lambda_t\left(\widetilde{\mathbf{C}}_h\right) = \frac{v_{h,t-1}}{M} \geq 0$ for $t = 1, \cdots, N$.

In (19), $\widehat{\mathbf{C}}_h = \mathbf{R}_\Omega \widetilde{\mathbf{C}}_h \mathbf{R}_\Omega^T$ is an $M \times M$ principal submatrix of $\widetilde{\mathbf{C}}_h$, where successive application of the interlacing inequality [73] leads to $\lambda_{t+N-M}\left(\widetilde{\mathbf{C}}_h\right) \leq \lambda_t\left(\widehat{\mathbf{C}}_h\right) \leq \lambda_t\left(\widetilde{\mathbf{C}}_h\right)$ for $1 \leq t \leq M$. Thus, $\min\limits_h \min\limits_{\mathbf{x}_h} \lambda_M\left(\widehat{\mathbf{C}}_h\right) = \min\limits_h \min\limits_{\mathbf{x}_h} \lambda_N\left(\widetilde{\mathbf{C}}_h\right) = 0$ from $v_{h,N-1} \geq 0$. On the other hand, $\max\limits_h \max\limits_{\mathbf{x}_h} \lambda_1\left(\widehat{\mathbf{C}}_h\right) = \max\limits_h \max\limits_{\mathbf{x}_h} \lambda_1\left(\widetilde{\mathbf{C}}_h\right) = \max\limits_h \max\limits_{\mathbf{x}_h} \frac{v_{h,0}}{M}$. By the Cauchy-Schwarz inequality, we obtain $\frac{v_{h,0}}{M} = \frac{|\widehat{x}_{h,\pi(0)}|^2}{M} = \frac{1}{M}\left|\sum_{k\in\mathcal{S}} x_{h,k}\mathbf{U}_2(\pi(0),k)\right|^2 \leq \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M}$, where $\mathcal{S}$ is the set of nonzero entries of $\mathbf{x}_h$ with $|\mathcal{S}| \leq K$. As $\lambda_t(\mathbf{C}_h) = \lambda_t\left(\widehat{\mathbf{C}}_h\right) + \sigma^2$ from $\mathbf{C}_h = \widehat{\mathbf{C}}_h + \sigma^2\mathbf{I}$, we have

$$\begin{aligned} \lambda_{\min} &= \min_h \min_{\mathbf{x}_h} \lambda_M(\mathbf{C}_h) = \sigma^2, \\ \lambda_{\max} &= \max_h \max_{\mathbf{x}_h} \lambda_1(\mathbf{C}_h) = \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M} + \sigma^2, \end{aligned} \tag{20}$$

where $h = 1$ or 2.

Meanwhile, the upper bound on $\mathrm{tr}\left(\mathbf{C}_2^{-1}\mathbf{C}_1\right)$ in Lemma 3 of [72] yields

$$\begin{aligned} D(p_1||p_2) &\leq \frac{1}{2}\sum_{t=1}^{M}\left(\log\frac{\lambda_{M+1-t}(\mathbf{C}_2)}{\lambda_t(\mathbf{C}_1)} + \frac{\lambda_t(\mathbf{C}_1)}{\lambda_{M+1-t}(\mathbf{C}_2)} - 1\right) \\ &= \frac{1}{2}\sum_{t=1}^{M} f(z_t), \end{aligned}$$

where $f(z) = z - \log z - 1$ and $z_t = \frac{\lambda_t(\mathbf{C}_1)}{\lambda_{M+1-t}(\mathbf{C}_2)} > 0$. With $\lambda_{\min}$ and $\lambda_{\max}$ in (20), define $\tau = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{K\mu^2(\mathbf{U}_2)\mathcal{E}_x}{M\sigma^2} + 1 > 1$. Similar to the proof of Theorem 1 in [72], $D(p_1||p_2) \leq \frac{M}{2}f(\tau)$, which yields (5).

Proof of Theorem 2

We use definitions and notations in the proof of Theorem 1. Let $\lambda_1(\mathbf{C}_3) \geq \cdots \geq \lambda_M(\mathbf{C}_3)$ be the eigenvalues of $\mathbf{C}_3 = \frac{\mathbf{C}_1+\mathbf{C}_2}{2}$. Clearly, the eigenvalues of $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$ are positive by (20) and the Weyl inequality [73]. In (6), let $\Gamma = \frac{|\mathbf{C}_1|^{\frac{1}{2}}|\mathbf{C}_2|^{\frac{1}{2}}}{|\mathbf{C}_3|} \triangleq \frac{\Gamma_n}{\Gamma_d}$. Then,

$$\begin{aligned} \Gamma_d = \prod_{t=1}^{M}\lambda_t(\mathbf{C}_3) &\leq \left(\frac{\sum_{t=1}^{M}\lambda_t(\mathbf{C}_3)}{M}\right)^M = \left(\frac{\mathrm{tr}(\mathbf{C}_3)}{M}\right)^M \\ &= \left(\frac{\mathrm{tr}(\mathbf{C}_1) + \mathrm{tr}(\mathbf{C}_2)}{2M}\right)^M, \end{aligned} \tag{21}$$

where the inequality is from the arithmetic mean-geometric mean inequality. For $h = 1$ or 2, the tth diagonal entry of $\widetilde{\mathbf{C}}_h = \frac{1}{N}\mathbf{U}_1^T\mathrm{diag}\left(\frac{|\widehat{\mathbf{x}}_h|^2}{M}\right)\mathbf{U}_1$ is given by $\frac{1}{MN}\sum_{k=0}^{N-1}|\widehat{x}_{h,k}|^2\mathbf{U}_1^2(k,t) = \frac{1}{MN}||\widehat{\mathbf{x}}_h||^2 = \frac{1}{M}||\mathbf{x}_h||^2 = \frac{\mathcal{E}_x}{M}$, where $\mathbf{U}_1^2(k,t) = 1$ for $0 \leq t \leq N-1$. Note that $\widehat{\mathbf{C}}_h = \mathbf{R}_\Omega\widetilde{\mathbf{C}}_h\mathbf{R}_\Omega^T$ has the same diagonal entry of $\widetilde{\mathbf{C}}_h$. Thus, from $\mathbf{C}_h = \widehat{\mathbf{C}}_h + \sigma^2\mathbf{I}$, we have

$$\mathrm{tr}(\mathbf{C}_h) = \mathrm{tr}(\widehat{\mathbf{C}}_h) + M\sigma^2 = \mathcal{E}_x + M\sigma^2, \tag{22}$$

where (21) becomes

$$\Gamma_d \leq \left(\frac{\mathcal{E}_x}{M} + \sigma^2\right)^M. \tag{23}$$

In Γ_n, the geometric mean-harmonic mean inequality yields

$$|\mathbf{C}_h|^{\frac{1}{2}} = \left(\prod_{t=1}^{M}\lambda_t(\mathbf{C}_h)\right)^{\frac{1}{2}} \geq \left(\frac{1}{\frac{1}{M}\sum_{t=1}^{M}\lambda_t^{-1}(\mathbf{C}_h)}\right)^{\frac{M}{2}}, \tag{24}$$

where $h = 1$ or 2. By the Kantorovich inequality [74],

$$\begin{aligned} \frac{1}{M}\sum_{t=1}^{M}\lambda_t^{-1}(\mathbf{C}_h) &\leq \frac{M}{4\,\mathrm{tr}(\mathbf{C}_h)}\left(\frac{\lambda_1(\mathbf{C}_h)}{\lambda_M(\mathbf{C}_h)} + \frac{\lambda_M(\mathbf{C}_h)}{\lambda_1(\mathbf{C}_h)} + 2\right) \\ &= \frac{M}{4\,\mathrm{tr}(\mathbf{C}_h)}\left(\frac{\lambda_{\max}}{\lambda_{\min}} + \frac{\lambda_{\min}}{\lambda_{\max}} + 2\right) \\ &= \frac{M}{4\,\mathrm{tr}(\mathbf{C}_h)}\left(\tau + \frac{1}{\tau} + 2\right), \end{aligned} \tag{25}$$

where $\lambda_1(\mathbf{C}_h)$ and $\lambda_M(\mathbf{C}_h)$ have been replaced by $\lambda_{\max}$ and $\lambda_{\min}$ of (20), respectively. In (25), $\tau = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M\sigma^2} + 1 = K\mu^2(\mathbf{U}_2)\cdot \mathrm{PNR} + 1$. By (22), (24), and (25),

$$\Gamma_n \geq \left(\frac{4\sqrt{\mathrm{tr}(\mathbf{C}_1)\cdot\mathrm{tr}(\mathbf{C}_2)}}{M(\tau+\frac{1}{\tau}+2)}\right)^M = \left(\frac{4\left(\frac{\mathcal{E}_x}{M}+\sigma^2\right)}{\tau+\frac{1}{\tau}+2}\right)^M. \tag{26}$$

By combining Γ_d and Γ_n, (23) and (26) yield

$$\Gamma = \frac{\Gamma_n}{\Gamma_d} \geq \frac{\left(\frac{4\left(\frac{\mathcal{E}_x}{M}+\sigma^2\right)}{\tau+\frac{1}{\tau}+2}\right)^M}{\left(\frac{\mathcal{E}_x}{M}+\sigma^2\right)^M} = \left(\frac{2\sqrt{\tau}}{\tau+1}\right)^{\frac{M}{2}} = \left(\frac{2\sqrt{K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+1}}{K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR}+2}\right)^{\frac{M}{2}}.$$

Finally, the proof is completed by $d_{\mathrm{H}}(p_1,p_2) = \sqrt{1-\Gamma^{\frac{1}{2}}}$.

Proof of Theorem 3

In (15), $\mathbf{K}_i$ is the Gram matrix, or $\mathbf{K}_i = \mathbf{A}_i^T\mathbf{A}_i$ with $\mathbf{A}_i = \mathbf{R}_{\Delta_i}\mathbf{U}_1^T\mathbf{R}_\Omega^T$ for $1 \leq i \leq n_s - 1$, where $\lambda_{\min}(\mathbf{K}_i) \geq 0$, since $\mathbf{K}_i$ is positive semi-definite [73]. Let $\mathbf{s}_{\Theta_i}$ and $\mathbf{s}'_{\Theta_i}$ be a pair of correct and wrong J-bit segments from a keystream $\mathbf{s}$ at the index set Θ_i, respectively. From (13), $\mathbb{E}\left[\mathbf{r}_i|\mathbf{s}_{\Theta_i}\right] = \mathbf{m}_i = \mathbf{R}_\Omega\mathbf{U}_1\mathbf{R}_{\Theta_i}^T\mathbf{s}_{\Theta_i}$ and $\mathbb{E}\left[\mathbf{r}_i|\mathbf{s}'_{\Theta_i}\right] = \mathbf{m}'_i = \mathbf{R}_\Omega\mathbf{U}_1\mathbf{R}_{\Theta_i}^T\mathbf{s}'_{\Theta_i}$, respectively. Also, (14) yields $\mathbb{E}\left[(\mathbf{r}_i-\mathbf{m}_i)(\mathbf{r}_i-\mathbf{m}_i)^T|\mathbf{s}_{\Theta_i}\right] = \mathbb{E}\left[\left(\mathbf{r}_i-\mathbf{m}'_i\right)\left(\mathbf{r}_i-\mathbf{m}'_i\right)^T|\mathbf{s}'_{\Theta_i}\right] = \mathbf{K}_i + \sigma^2\mathbf{I}$. Assuming that $\mathbf{r}_i$ is a Gaussian random vector, the binary hypothesis detection of Section 3.2 in [70] reveals that the pairwise error probability that $\mathbf{s}'_{\Theta_i}$ is incorrectly detected by the ith detection is

$$\Pr\left[\mathbf{s}_{\Theta_i}\to\mathbf{s}'_{\Theta_i}\middle|\,\mathbf{s}_{\Theta_i},\mathbf{s}'_{\Theta_i}\right] \geq Q\left(\frac{||\mathbf{m}_i-\mathbf{m}'_i||}{2\sqrt{\lambda_{\min}(\mathbf{K}_i)+\sigma^2}}\right) = Q\left(\frac{||\mathbf{R}_\Omega\mathbf{U}_1\mathbf{R}_{\Theta_i}^T\left(\mathbf{s}_{\Theta_i}-\mathbf{s}'_{\Theta_i}\right)||}{2\sqrt{\lambda_{\min}(\mathbf{K}_i)+\sigma^2}}\right). \tag{27}$$

We assume that the pairwise error event occurs only for a specific $\mathbf{s}'_{\Theta_i}$, which is closest to $\mathbf{s}_{\Theta_i}$, and ignore all the other $\mathbf{s}'_{\Theta_i}$. In other words, we take into account only a single $\mathbf{s}'_{\Theta_i}$, where $\mathbf{s}_{\Theta_i}-\mathbf{s}'_{\Theta_i}$ has the nonzero entry (+2 or −2) at one position, or equivalently $||\mathbf{s}_{\Theta_i}-\mathbf{s}'_{\Theta_i}|| = 2$ for a given $\mathbf{s}_{\Theta_i}$. This assumption, similar to the one in [75], is favorable for an adversary. From (27), the error probability under the assumption is given by

$$\begin{aligned}P_e^{(i)} &= \sum_{\mathbf{s}_{\Theta_i}}\Pr\left[\mathbf{s}_{\Theta_i}\right]\cdot\sum_{\mathbf{s}'_{\Theta_i}}\Pr\left[\mathbf{s}_{\Theta_i}\to\mathbf{s}'_{\Theta_i}\,|\,\mathbf{s}_{\Theta_i},\mathbf{s}'_{\Theta_i}\right]\cdot\Pr\left[\mathbf{s}'_{\Theta_i}\mid\mathbf{s}_{\Theta_i}\right]\\ &= \sum_{\mathbf{s}_{\Theta_i}}\Pr\left[\mathbf{s}_{\Theta_i}\right]\cdot\Pr\left[\mathbf{s}_{\Theta_i}\to\mathbf{s}'_{\Theta_i}\mid\mathbf{s}_{\Theta_i},\mathbf{s}'_{\Theta_i},||\mathbf{s}_{\Theta_i}-\mathbf{s}'_{\Theta_i}||=2\right]\\ &= \Pr\left[\mathbf{s}_{\Theta_i}\to\mathbf{s}'_{\Theta_i}\mid\mathbf{s}_{\Theta_i},\mathbf{s}'_{\Theta_i},||\mathbf{s}_{\Theta_i}-\mathbf{s}'_{\Theta_i}||=2\right]\\ &\geq Q\left(\frac{\sqrt{\sum_{k=0}^{M-1}4\left|\mathbf{U}_1\left(\omega_k,\theta_{i,\tau}\right)\right|^2}}{2\sqrt{\lambda_{\min}(\mathbf{K}_i)+\sigma^2}}\right)\\ &= Q\left(\sqrt{\frac{M\mu^2(\mathbf{U}_1)}{\lambda_{\min}(\mathbf{K}_i)+\sigma^2}}\right),\end{aligned} \tag{28}$$

where $\omega_k \in \Omega$ and $\theta_{i,\tau} \in \Theta_i$. In (28), we assumed that $\mathbf{s}_{\Theta_i}$ and $\mathbf{s}'_{\Theta_i}$ differ only at a position corresponding to the column index $\theta_{i,\tau}$ of $\mathbf{U}_1$. Note that $P_e^{(i)}$ is under the assumption that all the estimates from previous $i-1$ detections have been subtracted with no errors to yield $\mathbf{r}_i$ of (12). Then, the success probability of the ith detection is

$$\begin{aligned}P_s^{(i)} &= \Pr\left[\widehat{\mathbf{s}}_{\Theta_i}=\mathbf{s}_{\Theta_i}\mid\widehat{\mathbf{s}}_{\Theta_1}=\mathbf{s}_{\Theta_1},\cdots,\widehat{\mathbf{s}}_{\Theta_{i-1}}=\mathbf{s}_{\Theta_{i-1}}\right]\\ &= 1-P_e^{(i)} \leq 1-Q\left(\sqrt{\frac{M\mu^2(\mathbf{U}_1)}{\lambda_{\min}(\mathbf{K}_i)+\sigma^2}}\right),\end{aligned} \tag{29}$$

where $1 \leq i \leq n_s$. If a correct N-bit keystream is to be restored, all the component detections should be successful. Thus, the success probability of the SAMD is

$$\begin{aligned}P_{\mathrm{succ}} &= \Pr\left[\widehat{\mathbf{s}}_{\Theta_1}=\mathbf{s}_{\Theta_1},\cdots,\widehat{\mathbf{s}}_{\Theta_{n_s}}=\mathbf{s}_{\Theta_{n_s}}\right]\\ &= \prod_{i=1}^{n_s}\Pr\left[\widehat{\mathbf{s}}_{\Theta_i}=\mathbf{s}_{\Theta_i}\mid\widehat{\mathbf{s}}_{\Theta_1}=\mathbf{s}_{\Theta_1},\cdots,\widehat{\mathbf{s}}_{\Theta_{i-1}}=\mathbf{s}_{\Theta_{i-1}}\right]\\ &= \prod_{i=1}^{n_s}P_s^{(i)}.\end{aligned} \tag{30}$$

Finally, we obtain the upper bound of (17) by combining (29) and (30), which completes the proof.

Proof of Theorem 4

In (15), let $\mathbf{A}_i = \mathbf{R}_{\Delta_i}\mathbf{H}^T\mathbf{R}_\Omega^T$ with $\mathbf{U}_1 = \mathbf{H}$. Then, the singular values of $\mathbf{A}_i$ are equal to the square roots of the eigenvalues of $\mathbf{K}_i = \mathbf{A}_i^T\mathbf{A}_i$, where $\lambda_{\min}(\mathbf{K}_i) \geq 0$ for all i's. In other words, if $\sigma_{\min}(\mathbf{A}_i)$ denotes the minimum singular value of $\mathbf{A}_i$, then $\lambda_{\min}(\mathbf{K}_i) = \sigma_{\min}^2(\mathbf{A}_i)$.

To examine $\sigma_{\min}(\mathbf{A}_i)$ for $1 \le i \le n_s - 1$, we first define $\mathbf{B}_i = \mathbf{H}^T\mathbf{R}_\Omega^T$. Then, $\mathbf{B}_i$ is an $N \times M$ matrix satisfying $\mathbf{B}_i^T\mathbf{B}_i = \mathbf{R}_\Omega\mathbf{H} \cdot \mathbf{H}^T\mathbf{R}_\Omega^T = N\mathbf{I}$, which means that each column of $\mathbf{B}_i$ is mutually orthogonal. Also, it is clear that the l_2-norm of each row of $\mathbf{B}_i$ is $\sqrt{M}$, since each entry of $\mathbf{B}_i$ is ± 1. If Θ_i is a random selection, so is Δ_i, where $\mathbf{A}_i = \mathbf{R}_{\Delta_i}\mathbf{B}_i$ is an $(N - iJ) \times M$ matrix obtained by randomly subsampling $(N - iJ)$ rows from $\mathbf{B}_i$, where the selected row indices are specified by Δ_i. For such a matrix $\mathbf{A}_i$, Corollary 5.55 of [4] shows that for every $t \ge 0$,

$$\sigma_{\min}(\mathbf{A}_i) \ge \sqrt{N - iJ} - t\sqrt{M} \tag{31}$$

with probability at least $1 - 2Me^{-ct^2}$ for a constant $c > 0$. The corollary assumed that $t \ge \sqrt{c_1 \log M}$ and $N - iJ > c_2 M \log M$ for the bound to be nontrivial and nonnegative, where $0 < c_1 < c_2$. Thus, the bound of (31) is valid only for $i < \lceil \frac{N - c_2 M \log M}{J} \rceil = I_T$, and we set $\sigma_{\min}(\mathbf{A}_i) \ge 0$ if $i \ge I_T$, which gives the bound of (18) from $\lambda_{\min}(\mathbf{K}_i) = \sigma_{\min}^2(\mathbf{A}_i)$.

Acknowledgements
This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (no. NRF-2017R1A2B4004405).

Competing interests
The author declares that he has no competing interests.

Publisher's Note

Received: 17 May 2017 Accepted: 12 October 2017
Published online: 23 October 2017

References

1. DL Donoho, Compressed sensing. IEEE Trans. Inf. Theory. **52**(4), 1289–1306 (2006)
2. EJ Candes, J Romberg, T Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory. **52**(2), 489–509 (2006)
3. EJ Candes, T Tao, Near-optimal signal recovery from random projections: universal encoding strategies. IEEE Trans. Inf. Theory. **52**(12), 5406–5425 (2006)
4. YC Eldar, G Kutyniok, *Compressed sensing—theory and applications*. (Cambridge University Press, Cambridge, 2012)
5. J Tropp, JN Laska, M Duarte, J Romberg, RG Baraniuk, Beyond Nyquist: efficient sampling of sparse bandlimited signals. IEEE Trans. Inf. Theory. **56**(1), 520–544 (2010)
6. M Mishali, YC Eldar, From theory to practice: sub-Nyquist sampling of sparse wideband analog signals. IEEE J. Select Top. Sig. Process. **4**(2), 375–391 (2010)
7. J Haupt, W Bajwa, G Raz, R Nowak, Toeplitz compressed sensing matrices with applications to sparse channel estimation. IEEE Trans. Inf. Theory. **56**(11), 5862–5875 (2010)
8. MF Duarte, S Sarvotham, D Baron, MB Wakin, RG Baraniuk, in *Asilomar Conf. on Signals, Systems and computers*. Distributed compressed sensing of jointly sparse signals, (Pacific Grove, 2005), pp. 1537-1541
9. J Haupt, W Bajwa, M Rabbat, R Nowak, Compressed sensing for networked data. IEEE Sig. Process. Mag. **25**(2), 92–101 (2008)
10. C Caione, D Brunelli, L Benini, Compressive sensing optimization for signal ensembles in WSNs. IEEE Trans. Ind. Inform. **10**(1), 382–392 (2014)
11. M Duarte, M Davenport, D Takhar, JN Laska, T Sun, KF Kelly, RG Baraniuk, Single-pixel imaging via compressive sampling. IEEE Sig. Process. Mag. **25**(2), 83–91 (2008)
12. R Marcia, Z Harmany, R Willet, in *Proc. IS&T/SPIE Symp. Elec. Imag.: Comp. Imag*. Compressive coded aperture imaging, (San Jose, 2009)
13. M Lustig, D Donoho, J Pauly, in *Proc. Ann. Meeting of ISMRM*. Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories, (Seattle, 2006)
14. S Goginneni, A Nehorai, Target estimation using sparse modeling for distributed MIMO radar. IEEE Trans. Signal Process. **59**(11), 5315–5325 (2011)
15. Y Rachlin, D Baron, in *Proc. 46th Annu. Allerton Conf. Commun.Control, Comput*. The secrecy of compressed sensing measurements, (2008), pp. 813-817
16. A Orsdemir, HO Altun, G Sharma, MF Bocko, in *Proc. IEEE Military Commun. Conf. (MILCOM)*. On the security and robustness of encryption via compressed sensing, (2008), pp. 1-7
17. Y Zhang, LY Zhang, J Zhou, L Liu, F Chen, X He, A review of compressive sensing in information security field. IEEE Access Spec. Sect. Green Commun. Netw. 5G Wirel. **4**, 2507–2519 (2016)
18. AL Gibbson, FE Su, On choosing and bounding probability metrics. Int Stat Rev. **70**(3), 419–435 (2002)
19. S Goldwasser, S Micali, Probabilistic encryption. J. Comput. Syst. Sci. **28**, 270–299 (1984)
20. TM Cover, JA Thomas, *Elements of information theory*. (Wiley & Sons, Inc., Hoboken, 2006)
21. L Le Cam, *Asymptotic methods in statistical decision theory*. (Springer-Verlag, New York, 1986)
22. T Bianchi, V Bioglio, E Magli, in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP)*. On the security of random linear measurements, (2014), pp. 3992–3996
23. T Bianchi, V Bioglio, E Magli, Analysis of one-time random projections for privacy preserving compressed sensing. IEEE Trans. Inf. Forens. Sec. **11**(2), 313–327 (2016)
24. T Bianchi, E Magli, in *IEEE Workshop on Information Forensics and Security (WIFS)*. Analysis of the security of compressed sensing with circulant matrices, (2014), pp. 1–6
25. V Cambareri, M Mangia, F Pareschi, R Rovatti, G Setti, Low complexity multiclass encryption by compressed sensing. IEEE Trans. Signal Process. **63**(9), 2183–2195 (2015)
26. V Cambareri, M Mangia, F Pareschi, R Rovatti, G Setti, On known-plaintext attacks to a compressed sensing-based encryption: a quantitative analysis. IEEE Trans. Inf. Forensics Secur. **10**(10), 2182–2195 (2015)
27. Y Zhang, J Zhou, F Chen, LY Zhang, K-W Wong, X He, Embedding cryptographic features in compressive sensing. Neurocomputing. **205**, 472–480 (2016)
28. L Zeng, X Zhang, L Chen, Z Fan, Y Wang, Scrambling-based speech encryption via compressed sensing. EURASIP J. Adv. Signal Process. **2012**, 257 (2012)
29. LY Zhang, K-W Wong, Y Zhang, Q Lin, *Joint quantization and diffusion for compressed sensing measurements of natural images*, (2015), pp. 2744–2747
30. LY Zhang, K-W Wong, Y Zhang, J Zhou, Bi-level protected compressive sampling. IEEE Trans. Multimed. **18**(9), 1720–1732 (2016)
31. M Bloch, J Barros, *Physical-layer security—from information theory to security engineering*. (Cambridge University Press, 2011)
32. Y Zou, J Zhu, X Wang, L Hanzo, A survey on wireless security: technical challenges, recent advances, and future trends. Proc. IEEE. **104**(9), 1727–1765 (2016)
33. S Agrawal, S Vishwanath, in *Proc. IEEE Inf. Theory Workshop (ITW)*. Secrecy using compressive sensing, (2011), pp. 563–567
34. G Reeves, N Goela, N Milosavljevic, M Gastpar, in *Proc. IEEE Inf. Theory Workshop (ITW)*. A compressed sensing wire-tap channel, (2011), pp. 548–552
35. R Dautov, GR Tsouri, in *Proc. Int. Conf. Comput. Netw. Commun*. Establishing secure measurement matrix for compressed sensing using wireless physical layer security, (2013), pp. 354–358
36. SN George, DP Pattathil, A secure LFSR based random measurement matrix for compressive sensing. Sens. Imag. **15**(1), 1–29 (2014)
37. YD Li, Z Zhang, M Winslett, Y Yang, in *Proc.10th Annu. ACM Workshop Privacy Electron. Soc. (WPES)*. Compressive mechanism: utilizing sparse representation in differentical privacy, (2011), pp. 177–182
38. H Li, R Mao, L Lai, R Qui, in *Proc. IEEE SmartGridComm*. Compressed meter reading for delay-sensitive and secure load report in smart grid, (2010), pp. 114–119

39. J Gao, X Zhang, H Liang, X Shen, in *Proc.IEEE GLOBECOM, Commun. Inf. Syst. Security Symp*. Joint encryption and compressed sensing in smart grid data transmission, (2014), pp. 662–667
40. SW Golomb, G Gong, *Signal design for good correlation—for wireless communication, cryptography and radar*. (Cambridge University Press, New York, 2005)
41. R Rueppel, O Staffelbach, Products of linear recurring sequences with maximum complexity. IEEE Trans. Inf. Theory. **33**(1), 124–131 (1987)
42. T Herlestam, in *Advances in Cryptology-Eurocrypt'85*. On functions of linear shift register sequences. Lecture Notes in Computer Science (LNCS), vol. 219 (Springer-Verlag, 1986), pp. 119–129
43. T Beth, F Piper, in *Advances in Cryptology-Eurocrypt'84*. The stop-and-go generator. Lecture Notes in Computer Science (LNCS), vol. 209 (Springer-Verlag, 1985), pp. 88–92
44. D Gollmann, WG Chambers, Clock-controlled shift registers: a review. IEEE J. Sel. Areas Commun. **7**(4), 525–533 (1989)
45. D Coppersmith, H Krawczys, Y Mansour, in *Advances in Cryptology - Eurocrypt'93*. The shrinking generator. Lecture Notes in Computer Science (LNCS), vol. 773 (Springer-Verlag, 1993), pp. 22–39
46. W Meier, O Staffelbach, in *Advances in Cryptology-Eurocrypt'94*. The self-shrinking generator. Lecture Notes in Computer Science (LNCS), vol. 950 (Springer-Verlag, 1995), pp. 205–214
47. L Chen, G Gong, *Communication system security*. (Chapman & Hall/CRC, Boca Raton, 2012)
48. A Klein, *Stream ciphers*. (Springer-Verlag, London, 2013)
49. M Rudelson, R Vershynin, On sparse reconstruction from Fourier and Gaussian measurements. Comm. Pure Appl. Math. **61**(8), 1025–1045 (2008)
50. MF Duarte, YC Eldar, Structured compressed sensing: from theory to applications. IEEE Trans. Signal Process. **59**(9), 4053–4085 (2011)
51. J Katz, Y Lindell, *Introduction to modern cryptography, 2nd Ed.*, (Boca Raton, 2015)
52. L Le Cam, Convergence of estimates under dimensionality restrictions. Ann. Stat. **1**(1), 38–53 (1973)
53. A DasGupta, *Asymptotic theory of statistics and probability*. (Springer Science+Business Media, LLC, New York, 2008)
54. MS Pinsker, *Information and information stability of random variables and processes (in Russian)*. (U.S.S.R., Izv.Akad. Nauk, Moscow, 1960)
55. AA Fedotov, P Harremoës, F Topsoe, Refinements of Pinsker's inequality. IEEE Trans. Inf. Theory. **49**(6), 1491–1498 (2003)
56. Y Singer, MK Warmuth, in *Proc. Advances in Neural Information Processing Systems 11 (NIPS'98)*. Batch and on-line parameter estimation of Gaussian mixtures based on the joint entropy, (1998), pp. 578–584
57. S Foucart, H Rauhut, *A mathematical introduction to compressive sensing*. (Springer Science+Business Media, New York, 2013)
58. T Kailath, The divergence and Bhattacharyya distance measures in signal selection. IEEE Trans. Commun. Technol. COM-15(1), 52–60 (1967)
59. KT Abou-Moustafa, FP Ferrie, in *JMLR: Asian Conference on Machine Learning*. A note on metric properties for some divergence measures: The Gaussian case, vol. 25, (2012), pp. 1–15
60. A Guntuboyina, S Saha, G Schiebinger, Sharp inequalities for f-divergences. IEEE Trans. Inf. Theory. **60**(1), 104–121 (2014)
61. S Arora, L Babai, J Stern, Z Sweedyk, The hardness of approximate optima in lattices, codes, and systems of linear equations. J. Comput. Syst. Sci. **54**(2), 317–331 (1997)
62. M Ajtai, in *Proc. of 30th Ann. ACM Symp. Theory Comput*. The shortest vector problem in L_2 is NP-hard for randomized reductions, (1998), pp. 10–19
63. M Damen, K Abed-Meraim, J-C Belfiore, Generalized sphere decoder for asymmetrical space-time communication architecture. IET Electron. Lett. **36**(2), 166–167 (2000)
64. P Dayal, MK Varanasi, in *Proc.of 41st Annula Allerton Conf. on Comm. Control, and Comput*. A fast generalized sphere decoder for optimum decoding for under-determined MIMO systems, (2003)
65. T Cui, C Tellambura, An efficient generalized sphere decoder for rank-deficient MIMO systems. IEEE Commun. Lett. **9**(5), 423–425 (2005)
66. M Ajtai, in *Proc. 28th Annu. ACM Symp. Theory Comput*. Generating hard instances of lattice problems, (1996), pp. 99–108
67. O Goldreich, S Goldwasser, S Halevi, in *Proc.17th Annu. Int. Cryptography Conf*. Publick-key cryptosystems from lattice reduction problems, (1997), pp. 112–131
68. R Fischlin, J Seifert, in *Proc. 7th IMA Int. Conf*. Tensor-based trapdoors for CVP and their application to public key cryptography, (1999), pp. 244–257
69. RG Gallager, *Stochastic processes: theory for applications*. (Cambridge University Press, 2013)
70. HL van Trees, KL Bell, Z Tian, *Detection, estimation, and modulation theory: part I—detection, estimation, and filtering theory, Second Ed.* (Wiley & Sons, Inc., Hoboken, 2013)
71. D Needell, JA Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples. Appl. Comput. Harmon. Anal. **26**, 301–321 (2009)
72. NY Yu, Indistinguishability of compressed encryption with circulant matrices for wireless security. IEEE Signal Process. Lett. **24**(2), 181–185 (2017)
73. RA Horn, CR Johnson, *Matrix Analysis*, 2nd Ed. (Cambridge University Press, Cambridge, 2013)
74. G Strang, On the Kantorovich inequality. Proc. Amer. Math. Soc. **11**, 468 (1960)
75. J Choi, Secure transmission via compressive sensing in multicarrier systems. IEEE Signal Process. Lett. **23**(10), 1315–1319 (2016)