On the Indistinguishability of Compressed Encryption With Partial Unitary Sensing Matrices

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

***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 bipolar keystream. For security analysis, the total variation distance, bounded by the Hellinger distance, is examined as a security measure for the indistinguishability of the CS-based cryptosystem. By developing an upper bound on the Hellinger distance, we show that the CS-based cryptosystem can be computationally secure in terms of the indistinguishability, as long as it has a sufficiently long keystream for each encryption with low compression and sparsity ratios.**

I. Introduction

Compressed sensing (CS) [1]−[3] is to recover a sparse signal from the measurements that are believed to be incomplete. With efficient measurement and stable reconstruction, the CS technique has been of interest in various research areas. A *CS-based cryptosystem* encrypts a plaintext through a CS measurement process by keeping the sensing matrix secret. With the knowledge of the matrix, the ciphertext can then be decrypted by a CS reconstruction process. In [4], Rachlin and Baron proved that the CS-based cryptosystem cannot be perfectly secure, but might be computationally secure. Orsdemir *et al.* [5] showed that it is computationally secure against a key search technique via an algebraic approach. By renewing a random Gaussian sensing matrix at each encryption, Bianchi *et al.* [6] analyzed the security of noiseless CS-based cryptosystems. A similar analysis has been made for a noiseless CS-based cryptosystem having a circulant sensing matrix for efficient CS processes [7][8]. In [9] and [10], wireless channel characteristics could be exploited for the security of CS-based cryptosystems. The CS technique can also be applied in database systems [11], where random noise has been intentionally added to CS measurements for differential privacy. In practice, a variety of CS-based cryptosystems concerning the security of multimedia, imaging, and smart grid data have been suggested in [12]−[18].

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 assumed to be obtained by a keystream generator of stream ciphers. Then, the initial seed of the 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, the *total variation (TV)* distance [19] between probability distributions of ciphertexts conditioned on a pair of plaintexts is examined as a security measure for the *indistinguishability* [20]. We investigate the TV distance by developing an upper bound on 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 and sparsity ratios.

Notations: A matrix (or a vector) is represented by a bold-face upper (or lower) case letter. $\mathbf{U}^T$ and $|\mathbf{U}|$ denote the transpose and the determinant of a matrix $\mathbf{U}$, respectively. $\mathbf{U}(k,t)$ 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., $\mu(\mathbf{U}) = \max\limits_{k,t} |\mathbf{U}(k,t)|$. $\operatorname{diag}(\mathbf{s})$ 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 $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ 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.

II. Mathematical Model

A. 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 by employing a sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$, which produces a ciphertext $\mathbf{r} = \mathbf{\Phi}\mathbf{x} + \mathbf{n} \in \mathbb{R}^M$, 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.

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

978-1-5090-5990-4/17/$31.00 ©2017 IEEE

Definition 1: The sensing matrix of our CS-based cryptosystem is defined by

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

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 Ω with $|\Omega| = M$. 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, $\mathbf{U} = \frac{1}{\sqrt{N}}\mathbf{U}_1\text{diag}(\mathbf{s})\mathbf{U}_2$ 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 Definition 1, one may consider $\widehat{\mathbf{x}} = \mathbf{U}_2\mathbf{x}$ as a plaintext, which is sparse with respect to the sparsifying basis $\mathbf{U}_2$.

In this paper, $\mathbf{U}_1 = \mathbf{H}$, or an $N \times N$ Hadamard matrix employing a binary m-sequence [22] of period $N-1 = 2^n-1$, 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

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

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 the out-of-phase autocorrelation of $\mathbf{d}$ is -1 [22], 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.

We assume that the keystream $\mathbf{s}$ is a segment of length N from an original keystream of extremely long period, which enables to update the keystream $\mathbf{s}$ at each CS encryption. Regarding the keystream of our CS-based cryptosystem, we make the following assumption.

Assumption 1: In stream ciphers [23][24], an original keystream is designed to have nice pseudorandomness properties [22] such as balance, large period, low autocorrelation, large linear complexity, etc. With the properties, we assume that each entry of the secret keystream $\mathbf{s}$ takes $+1$ or -1 independently and uniformly at random, which facilitates the security analysis of our CS-based cryptosystem.

If a keystream generator produces 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.

B. CS Decryption

For CS decryption, a noisy ciphertext $\mathbf{r} = \mathbf{\Phi}\mathbf{x} + \mathbf{n} \in \mathbb{R}^M$ is available for an adversary as well as a legitimate recipient, where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ is the measurement noise. A legitimate recipient of the ciphertext, 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 or the keystream, with no knowledge of $\mathbf{\Phi}$. Proposition 1 presents the reliability and the stability of our CS-based cryptosystem for a legitimate recipient.

Proposition 1: [25][26] In our CS-based cryptosystem, a stable decryption of a plaintext with at most K nonzero entries is theoretically guaranteed with bounded errors for a legitimate recipient, 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 $\mu(\mathbf{U}) = \mathcal{O}(\sqrt{\log N})$ 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. Therefore, if $M = \mathcal{O}(K \log^5 N)$, Proposition 1 guarantees a stable and robust decryption for this case.

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

III. Security Analysis

The notion of indistinguishability has been proposed as a concept of the computational security [20]. Assume that a cryptosystem produces a ciphertext by encrypting one of two possible plaintexts. It 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 [27]. 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.

Let us consider the *indistinguishability experiment* [27] with a constraint of K-sparse plaintexts. First, an adversary creates plaintexts $\mathbf{x}_1$ and $\mathbf{x}_2$ with at most K nonzero entries. Our CS-based cryptosystem then produces a ciphertext $\mathbf{r} = \mathbf{\Phi}\mathbf{x}_h + \mathbf{n}$ by randomly selecting h, where $h = 1$ or 2. Given $\mathbf{r}$, the adversary tries to figure out which plaintext, $\mathbf{x}_1$ or $\mathbf{x}_2$, was encrypted for the ciphertext, by carrying out a polynomial time test $\mathcal{D} : \mathbf{r} \to h$. Let $d_{\text{TV}}(p_1, p_2)$ be the total variation (TV) distance [19] 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 that the probability that an adversary can successfully distinguish the plaintexts by any kind of binary hypothesis test $\mathcal{D}$ is bounded by [28]

$$p_d \le \frac{1}{2} + \frac{d_{\text{TV}}(p_1, p_2)}{2}. \qquad (2)$$

Therefore, if $d_{\text{TV}}(p_1, p_2)$ 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 $d_{\text{TV}}(p_1, p_2)$ closer to zero would be more secure in terms of the indistinguishability. Since computing $d_{\text{TV}}(p_1, p_2)$ directly is difficult [29], we compute the *Hellinger* distance [21] to bound the TV distance.

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, $\mathbb{E}[\mathbf{\Phi}] = \frac{1}{\sqrt{MN}}\mathbf{R}_\Omega\mathbf{U}_1 \cdot \mathbb{E}[\text{diag}(\mathbf{s})] \cdot \mathbf{U}_2 = \mathbf{0}$ for a given $\mathbf{R}_\Omega$, as each entry of $\mathbf{s}$ takes ± 1 with probability $1/2$ under Assumption 1. Thus, $\mathbb{E}[\mathbf{r}|\mathbf{x}_h] = \mathbb{E}[\mathbf{\Phi}] \cdot \mathbf{x}_h + \mathbb{E}[\mathbf{n}] = \mathbf{0}$.

TABLE I
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 generator creates a keystream $\mathbf{s} \in \{-1,+1\}^N$. The keystream $\mathbf{s}$ is updated at each encryption by the extremely long entire keystream.
CS encryption:	With the keystream $\mathbf{s}$ and a plaintext $\mathbf{x} \in \mathbb{R}^N$, a ciphertext is generated by $\mathbf{r} = \mathbf{\Phi}\mathbf{x} + \mathbf{n} \in \mathbb{R}^M$, 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 $\mathbf{s}$.

Then, the Hellinger distance for the multivariate Gaussian distributions p_1 and p_2 is given by [30][31]

$$d_{\mathrm{H}}(p_1,p_2) = \sqrt{1 - \frac{|\mathbf{C}_1|^{\frac{1}{4}}|\mathbf{C}_2|^{\frac{1}{4}}}{|\mathbf{C}_3|^{\frac{1}{2}}}} \tag{3}$$

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, and $\mathbf{C}_3 = \frac{\mathbf{C}_1+\mathbf{C}_2}{2}$. The Hellinger distance is particularly useful by giving both upper and lower bounds on the TV distance [32], i.e.,

$$d_{\mathrm{H}}^2(p_1,p_2) \le d_{\mathrm{TV}}(p_1,p_2) \le d_{\mathrm{H}}(p_1,p_2)\sqrt{2 - d_{\mathrm{H}}^2(p_1,p_2)}. \tag{4}$$

In what follows, we present an upper bound on the Hellinger distance of (3), which leads to 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,

$$d_{\mathrm{H}}(p_1,p_2) \le \sqrt{1 - \left(\frac{2\sqrt{K\mu^2(\mathbf{U}_2)\cdot \mathrm{PNR}+1}}{K\mu^2(\mathbf{U}_2)\cdot \mathrm{PNR}+2}\right)^{\frac{M}{4}}} \tag{5}$$

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

To prove Theorem 1, we begin with the following lemma.

Lemma 1: Let $\lambda_1(\mathbf{C}_h) \ge \cdots \ge \lambda_M(\mathbf{C}_h)$ be the eigenvalues of $\mathbf{C}_h$, where $h = 1$ or 2. Then,

$$\begin{aligned} \lambda_{\min} &= \min_h \min_{\mathbf{x}_h} \lambda_M(\mathbf{C}_h) = \sigma^2, \\ \lambda_{\max} &= \max_h \max_{\mathbf{x}_h} \lambda_1(\mathbf{C}_h) = \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M} + \sigma^2 \end{aligned} \tag{6}$$

Proof: Like Lemma 1 of [33], the covariance matrix of $\mathbf{r}$ is

$$\mathbf{C}_h = \mathbb{E}[\mathbf{r}\mathbf{r}^T|\mathbf{x}_h] = \mathbf{R}_\Omega\widetilde{\mathbf{C}}_h\mathbf{R}_\Omega^T + \sigma^2\mathbf{I}, \quad h = 1,2 \tag{7}$$

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

In (7), $\widehat{\mathbf{C}}_h = \mathbf{R}_\Omega\widetilde{\mathbf{C}}_h\mathbf{R}_\Omega^T$ is an $M \times M$ principal submatrix of $\widetilde{\mathbf{C}}_h$, where a successive application of the interlacing inequality [34] yields $\lambda_{t+N-M}(\widetilde{\mathbf{C}}_h) \le \lambda_t(\widehat{\mathbf{C}}_h) \le \lambda_t(\widetilde{\mathbf{C}}_h)$ for $1 \le t \le M$. Thus, $\min_h \min_{\mathbf{x}_h} \lambda_M(\widehat{\mathbf{C}}_h) = \min_h \min_{\mathbf{x}_h} \lambda_N(\widetilde{\mathbf{C}}_h) = 0$ from $v_{h,N-1} \ge 0$. On the other hand, $\max_h \max_{\mathbf{x}_h} \lambda_1(\widehat{\mathbf{C}}_h) = \max_h \max_{\mathbf{x}_h} \lambda_1(\widetilde{\mathbf{C}}_h) = \max_h \max_{\mathbf{x}_h} \frac{v_{h,0}}{M}$. By the Cauchy-Schwarz inequality, we obtain $\frac{v_{h,0}}{M} = \frac{|\widehat{x}_{h,\pi(0)}|^2}{M} = \frac{1}{M}\left|\sum_{k\in\mathcal{S}} x_{h,k}\mathbf{U}_2(\pi(0),k)\right|^2 \le \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M}$, where $\mathcal{S}$ is the set of nonzero entries of $\mathbf{x}_h$ with $|\mathcal{S}| \le K$. Finally, we have (6), as $\lambda_t(\mathbf{C}_h) = \lambda_t(\widehat{\mathbf{C}}_h) + \sigma^2$ from $\mathbf{C}_h = \widehat{\mathbf{C}}_h + \sigma^2\mathbf{I}$. □

Proof of Theorem 1: Let $\lambda_1(\mathbf{C}_3) \ge \cdots \ge \lambda_M(\mathbf{C}_3)$ 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 (6) and the Weyl inequality [34]. In (3), 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,

$$\Gamma_d = \prod_{t=1}^{M}\lambda_t(\mathbf{C}_3) \le \left(\frac{\mathrm{tr}(\mathbf{C}_3)}{M}\right)^M = \left(\frac{\mathrm{tr}(\mathbf{C}_1)+\mathrm{tr}(\mathbf{C}_2)}{2M}\right)^M \tag{8}$$

where the inequality is from the arithmetic mean-geometric mean inequality. For $h = 1$ or 2, the tth diagonal entry of $\widetilde{\mathbf{C}}_h$ is given by $\frac{1}{MN}\sum_{k=0}^{N-1}|\widehat{x}_{h,k}|^2\mathbf{U}_1^2(k,t) = \frac{1}{M}||\mathbf{x}_h||^2 = \frac{\mathcal{E}_x}{M}$, where $\mathbf{U}_1^2(k,t) = 1$ for $0 \le t \le N-1$. Also, $\widehat{\mathbf{C}}_h$ 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 $\mathrm{tr}(\mathbf{C}_h) = \mathrm{tr}(\widehat{\mathbf{C}}_h) + M\sigma^2 = \mathcal{E}_x + M\sigma^2$, where (8) becomes

$$\Gamma_d \le \left(\frac{\mathcal{E}_x}{M} + \sigma^2\right)^M. \tag{9}$$

In Γ_n, the geometric mean-harmonic mean inequality yields

$$|\mathbf{C}_h|^{\frac{1}{2}} = \left(\prod_{t=1}^{M}\lambda_t(\mathbf{C}_h)\right)^{\frac{1}{2}} \ge \left(\frac{1}{\frac{1}{M}\sum_{t=1}^{M}\lambda_t^{-1}(\mathbf{C}_h)}\right)^{\frac{M}{2}} \tag{10}$$

where $h = 1$ or 2. By the Kantorovich inequality [35],

$$\frac{1}{M}\sum_{t=1}^{M}\lambda_t^{-1}(\mathbf{C}_h) \le \frac{M}{4\,\mathrm{tr}(\mathbf{C}_h)}\left(\tau + \frac{1}{\tau} + 2\right) \tag{11}$$

where $\tau = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{K\mu^2(\mathbf{U}_2)\cdot\mathcal{E}_x}{M\sigma^2} + 1 = K\mu^2(\mathbf{U}_2)\cdot\mathrm{PNR} + 1$. By (10) and (11),

$$\Gamma_n \ge \left(\frac{4\sqrt{\mathrm{tr}(\mathbf{C}_1)\cdot\mathrm{tr}(\mathbf{C}_2)}}{M(\tau+\frac{1}{\tau}+2)}\right)^M = \left(\frac{4\left(\frac{\mathcal{E}_x}{M}+\sigma^2\right)}{\tau+\frac{1}{\tau}+2}\right)^M. \tag{12}$$

By combining Γ_d and Γ_n, (9) and (12) yield

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

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

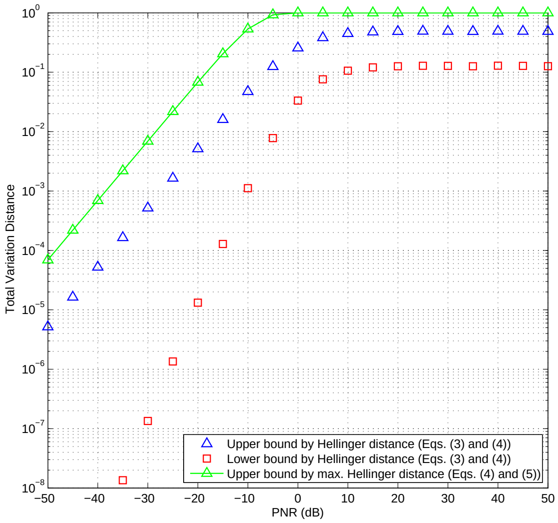

Fig. 1. The upper and lower bounds of total variation distance over PNR with $\mathbf{U}_2 = \mathbf{W}$, where $N = 1024$, $M = 48$, and $K = 4$. The number of tested plaintext pairs is 10000.

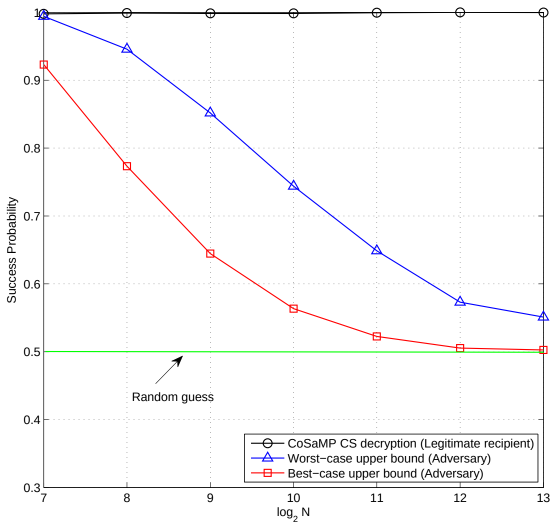

Fig. 2. The success probability over N at $\mathrm{PNR} = 25$ dB, where $\mathbf{U}_2 = \mathbf{W}$ and $M = 48$. Each plaintext has at most $K = \lfloor 8.5M/\log_2^2 N \rfloor$ nonzero entries. The number of tested plaintext pairs is 10000.

IV. 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. For CS encryption, we use the $N \times N$ Hadamard matrix $\mathbf{U}_1 = \mathbf{H}$ that employs a binary m-sequence of period $N - 1 = 2^n - 1$. In CS decryption, the CoSaMP recovery algorithm [36] has been employed for a legitimate recipient to decrypt each ciphertext with the knowledge of $\mathbf{\Phi}$.

Figure 1 displays the upper and lower bounds of total variation (TV) distance over PNR with $\mathbf{U}_2 = \mathbf{W}$, where $N = 1024$, $M = 48$, and $K = 4$. To compute the Hellinger distance (3), the covariance matrix of (7) has been used. Averaged over 10000 pairs of randomly generated plaintexts with at most K nonzero entries per each, the Hellinger distance yields the upper and lower bounds of TV distance by (4). We also sketch the theoretical upper bound on the TV distance, obtained by the maximum Hellinger distance of (5). The figure shows that the TV distance approaches to zero as noise level grows, which implies that our CS-based cryptosystem will be indistinguishable at low PNR. As PNR increases, however, the upper and lower bounds increase and finally converge to certain levels, respectively. More extensive simulations agreed with the implication of Theorem 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 2 depicts the upper bounds on the success probability of an adversary in the indistinguishability experiment, where $\mathbf{U}_2 = \mathbf{W}$ and $\mathrm{PNR} = 25$ dB. In the figure, the best- and worst-case upper bounds of (2) are from the minimum and maximum achievable TV distances of (4), respectively, obtained by the Hellinger distance of (3). With $M = 48$, the maximum sparsity is set as $K = \lfloor cM/\log_2^2 N \rfloor$ at each $N = 2^n$ for reliable nonuniform CS recovery [37], where $c = 8.5$. For comparison, we sketch the empirical success probability of CS decryption by a legitimate recipient, where a decrypted plaintext $\widehat{\mathbf{x}}$ has been declared as a success if $||\mathbf{x} - \widehat{\mathbf{x}}||^2/||\mathbf{x}||^2 < 10^{-2}$. In the test, the secret keystream has been generated by the *self-shrinking generator* [38] of a 128-stage LFSR. Figure 2 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. Thus, we conclude 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.

V. Conclusions

This paper has proposed a CS-based cryptosystem that encrypts a plaintext with a partial unitary sensing matrix embedding a secret keystream. We showed that our CS-based cryptosystem can offer a theoretically reliable decryption performance for a legitimate recipient. To examine the indistinguishability, we have studied the total variation distance as a security measure, by developing an upper bound on the Hellinger distance. Finally, we demonstrated that our CS-based cryptosystem can be computationally secure in terms of the indistinguishability, if the keystream length for each encryption is sufficiently large with low compression and sparsity ratios. A further research will be required for specifying a region of the system parameters by tightening the distance bound.

References

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

[2] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[3] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.

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

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

[6] T. Bianchi, V. Bioglio, and E. Magli, "On the security of random linear measurements," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP)*, pp. 3992-3996, May 2014.

[7] T. Bianchi and E. Magli, "Analysis of the security of compressed sensing with circulant matrices," in *Proc. IEEE Workshop on Inf. Forens. Security (WIFS)*, pp. 1-6, Dec. 2014.

[8] T. Bianchi, V. Bioglio, and E. Magli, "Analysis of one-time random projections for privacy preserving compressed sensing," *IEEE Trans. Inf. Forens. Security*, vol. 11, no. 2, pp. 313-327, Feb. 2016.

[9] G. Reeves, N. Goela, N. Milosavljevic, and M. Gastpar, "A compressed sensing wire-tap channel," in *Proc. IEEE Inf. Theory Workshop (ITW)*, pp. 548-552, Oct. 2011.

[10] S. Agrawal and S. Vishwanath, "Secrecy using compressive sensing," in *Proc. IEEE Inf. Theory Workshop (ITW)*, pp. 563-567, Oct. 2011.

[11] Y. D. Li, Z. Zhang, M. Winslett, and Y. Yang, "Compressive mechanism: Utilizing sparse representation in differentical privacy," in *Proc. 10th Annu. ACM Workshop Privacy Electron. Soc. (WPES)*, pp. 177-182, 2011.

[12] R. Dautov and G. R. Tsouri, "Establishing secure measurement matrix for compressed sensing using wireless physical layer security," in *Proc. Int. Conf. Comput. Netw. Commun.*, pp. 354-358, Jan. 2013.

[13] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Low complexity multiclass encryption by compressed sensing," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2183-2195, May 2015.

[14] S. N. George and D. P. Pattathil, "A secure LFSR based random measurement matrix for compressive sensing," *Sens. Imag.*, vol. 15, no. 1, pp. 1-29, 2014.

[15] Y. Zhang, J. Zhou, F. Chen, L. Y. Zhang, K.-W. Wong, and X. He, "Embedding cryptographic features in compressive sensing," *Neurocomputing*, vol. 205, pp. 472-480, 2016.

[16] H. Li, R. Mao, L. Lai, and R. Qui, "Compressed meter reading for delay-sensitive and secure load report in smart grid," in *Proc. IEEE SmartGridComm*, Oct. 2010.

[17] J. Gao, X. Zhang, H. Liang, and X. Shen, "Joint encryption and compressed sensing in smart grid data transmission," in *Proc. IEEE GLOBECOM, Commun. Inf. Syst. Security Symp.*, pp. 662-667, Dec. 2014.

[18] Y. Zhang, L. Y. Zhang, J. Zhou, L. Liu, F. Chen, and X. He, "A review of compressive sensing in information security field," *IEEE Access, Special Section on Green Communications and Networking for 5G Wireless*, vol. 4, pp. 2507-2519, 2016.

[19] A. L. Gibbson and F. E. Su, "On choosing and bounding probability metrics," *International Statistical Review*, vol. 70, no. 3, pp. 419-435, 2002.

[20] S. Goldwasser and S. Micali, "Probabilistic encryption," *Journ. Comput. Syst. Sciences*, vol. 28, pp. 270-299, 1984.

[21] L. Le Cam, *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York, 1986.

[22] S. W. Golomb and G. Gong, *Signal Design for Good Correlation - for Wireless Communication, Cryptography and Radar*. Cambridge University Press, 2005.

[23] L. Chen and G. Gong, *Communication System Security*, Chapman & Hall/CRC, 2012.

[24] A. Klein, *Stream Ciphers*, Springer-Verlag, London, 2013.

[25] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Comm. Pure Appl. Math.*, vol. 61, no. 8, pp. 1025-1045, Aug. 2008.

[26] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053-4085, Sep. 2011.

[27] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*, 2nd Ed., Chapman & Hall/CRC, 2015.

[28] L. Le Cam, "Convergence of estimates under dimensionality restrictions," *The Annals of Statistics*, vol. 1, no. 1, pp. 38-53, 1973.

[29] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer Science+Business Media, LLC 2008.

[30] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol.COM-15, no. 1, pp. 52-60, Feb. 1967.

[31] K. T. Abou-Moustafa and F. P. Ferrie, "A note on metric properties for some divergence measures: The Gaussian case," *JMLR: Asian Conference on Machine Learning*, vol. 25, pp. 1-15, 2012.

[32] A. Guntuboyina, S. Saha, and G. Schiebinger, "Sharp inequalities for f-divergences," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 104121, Jan. 2014.

[33] N. Y. Yu, "Indistinguishability of compressed encryption with circulant matrices for wireless security," *IEEE Signal Process. Lett.*, vol. 24, no. 2, pp. 181-185, Feb. 2017.

[34] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd Ed., Cambridge University Press, Cambridge, 2013.

[35] G. Strang, "On the Kantorovich inequality," *Proc. Amer. Math. Soc.*, vol. 11, pp. 468, 1960.

[36] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. and Comput. Harmon. Anal.*, vol. 26, pp. 301-321, 2009.

[37] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer Science+Business Media, New York, 2013.

[38] W. Meier and O. Staffelbach, "The self-shrinking generator," *Advances in Cryptology-Eurocrypt'94*, Lecture Notes in Computer Science (LNCS), vol. 950, pp. 205-214, Springer-Verlag, 1995.