ISIT'2013 1569719785

Low-Density Random Matrices for Secret Key Extraction

Hongchao Zhou
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA
Email: hongchao@mit.edu

Venkat Chandar
Lincoln Laboratory
Massachusetts Institute of Technology
Lexington, MA
Email: vchandar@mit.edu

Gregory Wornell
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA
Email: gww@mit.edu

***Abstract*—Secret key extraction, the task of extracting a secret key from shared information that is partially known by an eavesdropper, has important applications in cryptography. Motivated by the requirements of high-speed quantum key distribution, we study secret-key extraction methods with simple and efficient hardware implementations, in particular, linear transformations based on low-density random matrices. We show that this method can achieve the information-theoretic upper bound (conditional Shannon entropy) on efficiency for a wide range of key-distribution systems. In addition, we introduce a numerical method that allows us to tightly estimate the quality of the generated secret key in the regime of finite block length, and use this method to demonstrate that low-density random matrices achieve very high performance for secret key extraction.**

I. Introduction

Secret key extraction, also known as privacy amplification, introduced by Bennett, Brassard and Bobert [3], is the task of extracting a secret key from a shared random sequence that is partially known by an eavesdropper. It plays an essential role in secret key distribution and a broad range of cryptographic applications. Specifically, let $X \in \{0,1\}^n$ be the common random sequence shared by Alice and Bob in a secret-key distribution protocol, and let Z be the information known by an eavesdropper Eve. The goal of secret key extraction is to extract a secret key $Y \in \{0,1\}^m$ that is virtually uniformly distributed given all the information known by Eve. Formally,

$$H(Y|G,Z) \geq m - \epsilon \tag{1}$$

for specified small ϵ [4], where $G : \{0,1\}^n \to \{0,1\}^m$ is the function selected to map X to Y. This information-theoretical security requirement, known as strong secrecy [9], is widely used in the studies of information theoretically (or unconditionally) secure systems, and usually stronger than the statistical-distance requirement used in randomness extractors [11]. This strong secrecy ensures that information the eavesdropper obtains about the secret key is negligibly small in an absolute sense. For instance, if we apply the key Y in the well-known one-time-pad scheme [12] for cryptographic communication, then (1) guarantees that the mutual information between the transmitted content and all possible information known by an eavesdropper is at most ϵ, i.e., the total amount of possible leaked information is upper bounded by ϵ bits.

This work was supported in part by AFOSR under Grant. No. FA9550-11-1-0183, and by the DARPA InPho program under Contract No. HR0011-10-C-0159.

An important technique for secret key extraction is universal hashing [4], [7], which was introduced by Carter and Wegman [6]. Randomness extractors [11] are another type of construction that possibly could be used. However, many existing constructions are not feasible for high-speed quantum key distribution, which requires a simple and efficient hardware implementation of secret key extraction [14]. Our interest falls into linear transformations based on sparse matrices, i.e., the extracted secret key Y is MX for a sparse $m \times n$ matrix M over $GF(2)$. In contrast to the other approaches, this approach should have an efficient hardware implementation based on FPGAs [15]. We demonstrate that the sparse matrix M can be randomly constructed, and it can achieve the information-theoretical upper bound on efficiency for a wide range of key-distribution systems.

We consider a widely used scenario for key distribution described in [1], [8]. It is assumed that Alice and Bob observe two memoryless sequences $A \in \mathcal{A}^N$ and $B \in \mathcal{B}^N$ from a secret channel, such as a quantum channel. Meanwhile, an eavesdropper Eve may observe a sequence $E \in \mathcal{E}^N$, which is correlated with A and B. For instance, Eve may passively listen to the secret channel and she gets a memoryless sequence E; or Eve actively selects a fraction of symbols in A to listen, hence she knows at most t symbols in A. As a generalization of both cases, we let E be an independent sequence, in which the distribution of each symbol is controlled by Eve. In a protocol of secret key distribution, Alice and Bob communicate over a public channel, and we assume that every message communicated between Alice and Bob can be eavesdropped by Eve. A simple scheme for key distribution is based on Slepian-Wolf coding [13], namely, Alice generates a message $R \in \{0,1\}^r$ based on A and sends it to Bob. By jointly decoding B and R, Bob can obtain the sequence A with a probability almost 1. Hence the key-distribution problem is converted into the problem of secret key extraction, where A is the shared memoryless sequence between Alice and Bob, and (E,R) is the information known by Eve. It is assumed that $R = h(X)$ for an arbitrary function $h : \{0,1\}^n \to \{0,1\}^r$.

In this paper, we study the application of low-density ran-

dom matrices to the secret-key-extraction problem described above. Specifically, let M be an $m \times n$ low-density random matrix over $GF(2)$, and compute the secret key as $Y = MX$. We show that as n becomes large enough, the number of secret bits that can be extracted by this method approaches $\min_e H(A|E = e) - r$, where r is the length of the message R. In addition, fixing n and m, a numerical upper bound of $m - H(Y|M, Z)$ is provided that can be used to guarantee the security required by practical key-distribution applications. Note that the method requires the independence of M and E, i.e., M needs to be constructed after E has been observed. Otherwise, the eavesdropper may attack the system based on the matrix M. However, it is not efficient for Alice and Bob to reconstruct M for every block, and it is desired to reuse the same construction. In practical high-speed key-distribution systems, one can store data for a certain number of blocks, and then construct a new matrix M for extracting secret bits from these cached blocks (e.g. based on FPGAs).

The approach based on low-density random matrices can be treated as an extension of the approach based on uniform random matrices, a standard construction of universal hash functions. In [4], it is shown that the number of secret bits that can be extracted by universal hashing is approximately $\min_{z\in\mathcal{Z}} R(X|Z = z)$, where $R(X|Z = z)$ is the Rényi entropy (of order two) of X conditioned on $Z = z$. It is also demonstrated that there is a gap between Rényi entropy and Shannon entropy (i.e., the information-theoretical limit), even when the Rényi entropy is replaced by the tighter smooth Rényi entropy introduced by Renner and Wolf [10]. With some constraints on the distribution P_{XZ}, certain techniques can be applied to close the gap. In particular, two types of constraints on P_{XZ} have been studied [4]. If X is uniformly distributed on $\{0,1\}^n$ and $Z = h(X)$ for an arbitrary eavesdropping function $h : \{0,1\}^n \to \{0,1\}^r$, then the number of secret bits that can be extracted by universal hashing is approximately $n - r$; if X is uniformly distributed on $\{0,1\}^n$ and Z is a sequence received by Eve by transmitting X over a binary symmetric channel with bit-error probability ϵ, then the number of secret bits that can be extracted by universal hashing is asymptotically $h(\epsilon)n$, where $h(\epsilon)$ is the binary entropy function. In this paper, these results are generalized: the constraint on P_{XZ} we consider is relaxed such that it can describe many practical key-distribution systems, and the number of extractable secret bits still approaches the information-theoretic limit specified by Shannon entropy.

II. Asymptotic Analysis

In this section, we study the asymptotic performance of low-density random matrices for secret key extraction. We first consider the case that $r = 0$. Then given E, A is an independent sequence.

Definition 1. *We say a source $X \in \mathcal{X}^n$ has the asymptotic semi-equipartition property if and only if for all $\epsilon > 0$, there exists large enough n and a set $\mathbb{S}$ such that*

$$P(X \in \mathbb{S}) \geq 1 - \frac{\epsilon}{n} \tag{2}$$

and

$$\log_2 \frac{1}{P(X = x)} \geq H(X)(1 - \epsilon), \forall x \in \mathbb{S}. \tag{3}$$

We claim that if a source is independent, then it has the asymptotic semi-equipartition property. When the distribution of each symbol of the source is identical, it is easy to prove this claim. However, although the source we consider is independent, the symbols may not be distributed identically. To show that an arbitrary independent source has this property, we first consider a simplified binary source.

Lemma 1. *Let $X \in \{0,1\}^n$ be an independent source such that $P(X_i) \in [p - \frac{\delta}{2}, p + \frac{\delta}{2}]$ with $p \leq \frac{1}{2}$, and given any $\epsilon > 0$, we let $\mathbb{S} = \{x \in \{0,1\}^n : |x| \geq n(p - \frac{\delta}{2})(1 - \epsilon)\}$. Then for n large enough,*

$$P(X \in \mathbb{S}) \geq 1 - \frac{1}{2} e^{-2n\epsilon^2(p-\frac{\delta}{2})^2}.$$

The lemma can be proved based on the Hoeffding's inequality, and we can get a similar result when $p \geq \frac{1}{2}$.

Lemma 2. *Let $X \in \{0,1\}^n$ be a binary independent source, i.e., $P(X) = \prod_{i=1}^n P(X_i)$. If $H(X) = \Theta(n)$, then X has the asymptotic semi-equipartition property.*

Proof. By permutating all the bits in X, we let $P(X_i = 1) \leq P(X_j = 1)$ for all $i < j$. This operation does not affect the asymptotic semi-equipartition property of X.

For any $\delta > 0$, we partition X into strings $X^{(1)}, X^{(2)}, ..., X^{(\frac{1}{\delta})}$ such that $P(X_j^{(i)} = 1) \in [p_i - \frac{\delta}{2}, p_i + \frac{\delta}{2}]$ for all $1 \leq j \leq |X^{(i)}|$, where $X_j^{(i)}$ is the jth bit in $X^{(i)}$.

As $n \to \infty$, if $|X^{(i)}| = \Theta(n)$, then $|X^{(i)}| \to \infty$. In this case, we get a set $\mathbb{S}^{(i)}$ such that

$$\mathbb{S}^{(i)} = \begin{cases} \{x \in \{0,1\}^n : |x| \geq n(p_i - \frac{\delta}{2})(1 - \epsilon)\} & \text{if } p_i \leq \frac{1}{2}, \\ \{x \in \{0,1\}^n : |x| \leq n(p_i + \frac{\delta}{2})(1 + \epsilon)\} & \text{if } p_i > \frac{1}{2}. \end{cases}$$

According to Lemma 1, $P(X \in \mathbb{S}^{(i)}) \geq 1 - \frac{\epsilon}{|X_i|^3} \geq 1 - \frac{\epsilon}{n^2}$ for $n \to \infty$.

If $|X^{(i)}| = o(n)$, we let $\mathbb{S}^{(i)} = \{0,1\}^{|X^{(i)}|}$.

Then we define a set for $X \in \{0,1\}^n$, which is

$$\mathbb{S} = \mathbb{S}^{(1)} \times \mathbb{S}^{(2)} \times ... \times \mathbb{S}^{\frac{1}{\delta}}.$$

In this case,

$$P(X \in \mathbb{S}) \geq (1 - \frac{\epsilon}{n^2})^{\frac{1}{\delta}} \geq 1 - \frac{\epsilon}{\delta n^2} \geq 1 - \frac{\epsilon}{n}. \tag{4}$$

Now, let's consider a sequence $X \in \mathbb{S}$. For $X^{(i)} \in \mathbb{S}^{(i)}$ with $|X^{(i)}| \to \infty$, if $p_i \leq \frac{1}{2}$, then

$$\log_2 \frac{1}{P(X^{(i)})} \geq |X^{(i)}|((p - \frac{\delta}{2})(1 - \epsilon) \log_2 \frac{1}{p + \frac{\delta}{2}}$$

$$+((1 - (p - \frac{\delta}{2})(1 - \epsilon)) \log_2 \frac{1}{1 - p + \frac{\delta}{2}}).$$

Let $\epsilon, \delta \to 0$, then $\frac{\log_2 \frac{1}{P(X^{(i)})}}{H(X^{(i)})} \to 1$.

Similarly, if $p_i \geq \frac{1}{2}$, we can get the same conclusion.

Let ϵ, δ be small enough, as $n \to \infty$, based on the assumption that $H(X) = \Theta(n)$, we can get

$$\log_2 \frac{1}{P(X = x | X \in \mathbb{S})} \to H \geq H(X)(1 - \epsilon). \quad (5)$$

It shows that X has the asymptotic semi-equipartition property, following (4) and (5). ∎

The proof above can extend to an arbitrary large-alphabet independent source. The idea is that, instead of partitioning a probability space $[0, 1]$ into small segments with each of length δ, we partition the distribution space on $\mathcal{X}$ into small cubic, each with column $\delta^{|\mathcal{X}|-1}$.

Theorem 3. *Let $X \in \mathcal{X}^n$ be an independent source with an alphabet $\mathcal{X}$ of finite size, i.e., $P(X) = \prod_{i=1}^n P(X_i)$. If $H(X) = \Theta(n)$, then the binary representation of X has the asymptotic semi-equipartition property.*

Theorem 4. *Let $X \in \{0, 1\}^n$ be the binary representation of an independent source with finite alphabet, and let M be an $m \times n$ random matrix on $GF(2)$ where each entry equals one with probability p and $m = \Theta(n)$. Let $Y = MX$. If $\frac{1}{2} \geq p > K\frac{\log_2 n}{n}$ for any constant K and $\frac{m}{H(X)} < 1$, then*

$$\lim_{n\to\infty} m - H(Y|M) = 0.$$

Proof. According to Lemma 3, an independent source has the asymptotic semi-equipartition property, so does its binary representation X. As a result, we can find a set $\mathbb{S}$ that satisfies (2) and (3), where ϵ can be arbitrarily small.

Let us use $X_{\mathbb{S}}$ denote the random sequence X given the condition that $X \in \mathbb{S}$, i.e., for all $x \in \{0, 1\}^n$, $P(X_{\mathbb{S}} = x) = P(X = x | X \in \mathbb{S})$. The min-entropy of $X_{\mathbb{S}}$, defined by $\min_{x\in\{0,1\}^n} \log_2 \frac{1}{P(X_{\mathbb{S}}=x)}$, is at least

$$k = H(X)(1 - \epsilon) + \log_2(1 - \frac{\epsilon}{n}).$$

If we only consider the sequences in $\mathbb{S}$, then

$$\begin{aligned} & H(Y|M, X \in \mathbb{S}) \\ = & \sum_{M,y} P(M)P(y|M, X \in \mathbb{S}) \log_2 \frac{1}{P(y|M, X \in \mathbb{S})} \\ \geq & \log_2 \frac{1}{\sum_{M,y} P(M)P(Y = y|M, X \in \mathbb{S})^2} \\ = & \log_2 \frac{1}{P(MX_{\mathbb{S}} = MX'_{\mathbb{S}})}, \end{aligned}$$

where $X_{\mathbb{S}}$ and $X'_{\mathbb{S}}$ are identical and independent samples from $\mathbb{S}$.

According to the proof of Theorem 1 in [5], if

$$p = \min\{\frac{1}{m} \log_2 \frac{m}{\delta'} \ln \frac{K'n}{m}, \frac{1}{2}\}$$

with a sufficiently large constant K' and $0 < \delta' < 1$, then

$$P(MX_{\mathbb{S}} = MX'_{\mathbb{S}}) \leq \frac{1 + \delta' + K'2^{-k+m}}{2^m}, \quad (6)$$

where k is the min-entropy of $X_{\mathbb{S}}$ obtained above. Hence,

$$H(Y|M, X \in \mathbb{S}) \geq m - \log_2(1 + \delta' + K'2^{-k+m}).$$

Furthermore, we can get that for any $\epsilon > 0$, there exists large enough n such that

$$\begin{aligned} H(Y|M) & \geq H(Y|M, I_{X\in\mathbb{S}}) \\ & \geq P(X \in \mathbb{S})H(Y|M, X \in \mathbb{S}) \\ & \geq (1 - \frac{\epsilon}{n})(m - \log_2(1 + \delta' + K'2^{-k+m})). \end{aligned}$$

If $\frac{1}{2} \geq p > K\frac{\log_2 n}{n}$ for any constant K, then we can let δ' be arbitrarily small and K' be a large constant. In this case, if $\frac{m}{H(X)} < (1 - 2\epsilon)$, it is easy to prove that

$$m - H(Y|M) \to 0.$$

This completes the proof. ∎

Theorem 5. *Let $X \in \{0, 1\}^n$ be the binary representation of an independent source with finite alphabet, and let $R = h(X)$ for an arbitrary function $h : \{0, 1\}^n \to \{0, 1\}^r$. Assume that $Y = MX$ with an $m \times n$ random matrix M on $GF(2)$ where each entry equals one with probability p and $m = \Theta(n)$. If $\frac{1}{2} \geq p > K\frac{\log_2 n}{n}$ for any constant K and $\frac{m}{H(X)-r} < 1$, then*

$$\lim_{n\to\infty} m - H(Y|M, R) = 0.$$

Proof. Similar to the proof of Theorem 4, we first consider the set $\mathbb{S}$ that satisfies (2) and (3). Based on the value of R, we divide $\mathbb{S}$ into 2^r sets, such that for all $h \in \{0, 1\}^r$,

$$\mathbb{S}_h = \{x \in \mathbb{S}, h(x) = h\}.$$

The min-entropy of the set $\mathbb{S}_h$ is

$$k_h = \min_{x\in\mathbb{S}_h} \log_2 \frac{P(\mathbb{S}_h)}{P(x)} \geq H(X)(1 - \epsilon) + \log_2 P(\mathbb{S}_h).$$

Hence,

$$P(\mathbb{S}_h) \leq 2^{k_h - H(X)(1-\epsilon)}.$$

As a result, for $k > 0$,

$$\sum_{h:k_h<k} P(\mathbb{S}_h) \leq 2^r 2^{k-H(X)(1-\epsilon)}.$$

If $p = \min\{\frac{1}{m} \log_2 \frac{m}{\delta'} \ln \frac{K'n}{m}, \frac{1}{2}\}$, following the same proof as that for Theorem 4, we can get

$$\begin{aligned} & H(Y|M, R) \\ \geq & P(X \in \mathbb{S})H(Y|M, X \in \mathbb{S}, R) \\ \geq & \sum_h P(X \in \mathbb{S}_h)H(Y|M, X \in \mathbb{S}_h) \\ \geq & \sum_{h:k_h\geq k} P(X \in \mathbb{S}_h)H(Y|M, X \in \mathbb{S}_h) \\ \geq & (1 - \frac{\epsilon}{n} - 2^r 2^{k-H(X)(1-\epsilon)}) \\ & \times(m - \log_2(1 + \delta' + K'2^{-k+m})). \end{aligned}$$

When $\frac{1}{2} \geq p > K\frac{\log_2 n}{n}$ for any constant K as $n \to \infty$, we can let δ' be arbitrarily small and K' be a large constant.

We set $k = H(X)(1 - \epsilon)^2 - r$ and $\frac{m}{k} < 1$. As $n \to \infty$, we can let ϵ be arbitrarily small. Hence, based on the above inequality, we can get $m - H(Y|M, R) \to 0$. ∎

Corollary 6. *Let $X \in \{0,1\}^n$ be the binary representation of a memoryless sequence $A \in \mathcal{A}^N$ shared between Alice and Bob. Let (E, R) be the information known by Eve, where $E \in \mathcal{E}^N$ is an independent sequence and $R = h(X)$ for an arbitrary function $h : \{0,1\}^n \to \{0,1\}^r$. Assume that $Y = MX$ with an $m \times n$ random matrix M on $GF(2)$ where each entry equals one with probability p and $m = \Theta(n)$. If $\frac{1}{2} \geq p > K\frac{\log_2 n}{n}$ for any constant K and $\frac{m}{H(A|E=e)-r} < 1$ for any $e \in \mathcal{E}^n$, then*

$$\lim_{n\to\infty} m - H(Y|M, E, R) = 0.$$

Specifically, when E is a memoryless sequence, it can be proved that the approach based on low-density random matrices can extract a secret key of length $H(A|E) - r$ asymptotically. According to [13], if we can construct an optimal Slepian-Wolf code, then the shortest length of R is approximately $r = H(A|B)$. Combining an optimal Slepian-Wolf code with the secret-key-extraction technique described in this paper yields a secret key of length

$$m = H(A|E) - H(A|B)$$

asymptotically. It achieves the information-theoretical lower bound for secret key distribution derived in [8] when $I[A; E] \leq I[B; E]$.

III. Numerical Evaluation

The previous section studies the asymptotic performance of low-density random matrices in secret key extraction. However, for practical security-related applications, it is crucial to know how good the approach is in the regime of finite block length, i.e., we need to guarantee that little information about the generated secret key is known by any eavesdropper. For instance, given a source of length $n = 1000$, we would like to know how small the secret-key length m should be to guarantee that $m - H(Y|M, Z) \leq \delta$, namely, the maximal information leaked to Eve is at most δ bits. One idea of getting an upper bound of δ is based on the Rényi entropy of A conditioned on E as derived in [4], however, it yields an upper bound of δ too loose to be used in practical applications especially when there is a big gap between Rényi entropy and Shannon entropy.

In this section, we introduce a numerical method to evaluate the performance of the approach based on low-density random matrices in the regime of finite block length when the distribution P_{AE} is given.

Lemma 7. *Let $X \in \{0,1\}^n$ be a source with min-entropy k, i.e., $P(X = x) \leq 2^{-k}$ for all $x \in \{0,1\}^n$, and let $R = h(X)$ for an arbitrary function $h : \{0,1\}^n \to \{0,1\}^r$. We assume that $Y = MX$, where M is an $m \times n$ random matrix on $GF(2)$ where each entry equals one with probability p, then*

$$H(Y|M, R) = \begin{cases} \lambda(k) & \text{if } r = 0, \\ \max_v (1 - 2^{r+v-k})\lambda(v) & \text{otherwise.} \end{cases} \tag{7}$$

Here

$$\lambda(v) = v + \log_2 \frac{2^m}{\sum_{j=0}^m \binom{m}{j} \sum_{x\in S(v,n)} (1-2p)^{j|x|}},$$

and $S(v, n)$ is the subset of $\{0,1\}^n$ such that $|S(v, n)| = 2^v$ and for all $x' \in \{0,1\}^n / S(v, n), x \in S(v, n)$, $|x'| \geq |x|$.

Proof. (1) First, we prove that $H(Y|M) \geq \lambda(k)$.

It follows two inequalities:

$$H(Y|M) \geq \log_2 \frac{1}{P(MX_{\mathbb{S}} = MX'_{\mathbb{S}})}$$

proved in Theorem 4, and

$$P(MX_{\mathbb{S}} = MX'_{\mathbb{S}}) \leq \frac{1}{2^m} \sum_{j=0}^m \binom{m}{j} \frac{\sum_{x\in S(k,n)} (1-2p)^{j|x|}}{|S(k, n)|}, \tag{8}$$

which follows the proof of Theorem 1 in [5].

(2) Using a similar proof as that of Theorem 5, we can get

$$H(Y|M, R) \geq \max_v (1 - 2^{r+v-k})\lambda(v).$$

This completes the proof. ∎

Based on the above lemma, if the min-entropy of a source X is k, then we can get a lower bound of $H(Y|M, R)$ as a function of k, m, r, p, where m is the output length, r is the message length, and p is the density of the random matrix. Hence, we denote this bound as $B(k, m, r, p)$.

Now, we study the lower bound of $H(Y|M, R)$ when X is the binary representation of an independent sequence. For such a source X, we define a function $\Phi : [0, \infty) \to [0, 1]$ as

$$\Phi(k) = \sum_{x\in\{0,1\}^n} P(X = x) I_{(P(X=x)\leq 2^{-k})}.$$

This function can be calculated analytically or estimated based on Monte-Carlo simulation. Our goal is to get a lower bound on $H(Y|M, R)$ based on this function Φ (assume it is given).

Our idea is to partition all the sequences in $\{0,1\}^n$ into groups, denoted by $S_0, S_1, ..., S_t$. We do this by selecting a group of values $g_1, g_2, ..., g_t$ with $0 < g_1 < g_2 < ... < g_t < \infty$ to divide the probability space into intervals, then for all $0 \leq i \leq t$,

$$S_i = \{x \in \{0,1\}^n | 2^{-g_{i+1}} < P(X = x) \leq 2^{-g_i}\},$$

where $g_0 = 0$ and $g_{t+1} = \infty$. It is easy to get

$$P(S_i) = \Phi(g_i) - \Phi(g_{i+1})$$

and given $X \in S_i$, the min-entropy of X is at least $k_i = g_i + \log_2 P(S_i)$. Based on which, we can get

$$H(Y|M, R) \geq \sum_{i=0}^t P(S_i) B(k_i, m, r, p),$$

which is a lower bound of $H(Y|M, R)$. We can maximize this lower bound by selecting a good combination of $t, g_1, g_2, ..., g_t$ based on the function Φ.

In our model, X is the binary representation of a memoryless sequence A, and E is an independent sequence. Hence,

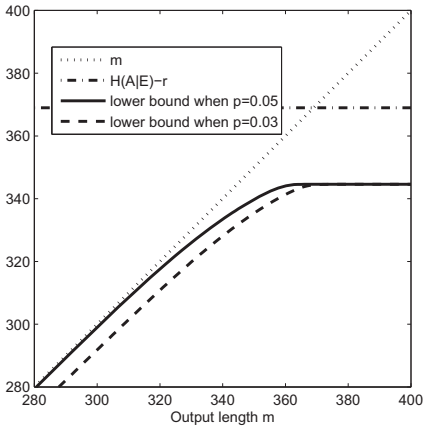

Fig. 1. The lower bound of $H(Y|M,E,R)$ for the source described in Example 1.

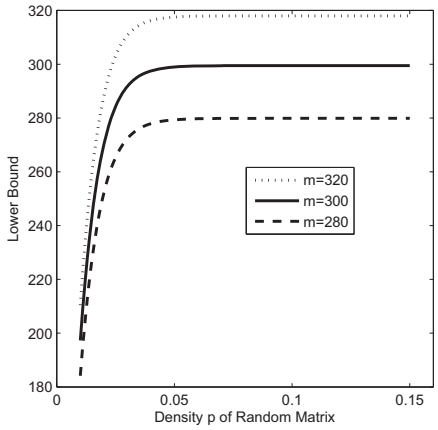

Fig. 2. The lower bound of $H(Y|M,E,R)$ for the source described in Example 1.

given $E=e$ with $e\in\mathcal{E}^N$, A can be treated as an independent sequence. Based on the approach above, we can calculate a lower bound for $H(Y|M,E=e,R)$. However, it is impossible to compute a lower bound for $H(H|M,E,R)$ based on the expression

$$H(Y|M,E,R)=\sum_{e\in\mathcal{E}^N}P(E=e)H(Y|M,E=e,R),$$

since it is too complex to enumerate all $e\in\mathcal{E}^N$ and the alphabet $\mathcal{E}$ may not be finite. A simple method is that one can generate a group of samples of E, denoted by $e_1,e_2,...,e_M$. If M is large, then

$$H(Y|M,E,R)\geq\min_{i=1}^{M}H(Y|M,E=e_i,R)$$

with a probability almost one. Hence, we can use the minimal one of the lower bounds for $H(Y|M,E=e_i,R)$ with $1\leq i\leq M$ to estimate a lower bound for $H(Y|M,E,R)$.

Example 1. *Let $A\in\{0,1\}^n, E\in\{0,1\}^n$ with $n=1000$ be memoryless sequences such that*

$$P(A_i=1)=\frac{1}{2}, P(E_i\neq A_i|A_i)=\delta=0.1.$$

Given $E=e$, the function Φ is fixed for all $e\in\{0,1\}^n$, and it can be calculated. We partition all the sequences in $\{0,1\}^n$ into 101 groups such that for all $0\leq i\leq 99$, $S_i=\{x\in\{0,1\}^n|\|x\|=i\}$, and for $i=100$, $S_i=\{x\in\{0,1\}^n|\|x\|\geq i\}$. Based on this partition, if $p=0.1$, $r=100$ and $m=300$, then $H(Y|M,E,R)=H(Y|M,E=e,R)\geq 299.53$.

In the above example, we let the message length be $r=100$. By setting the density of the random matrix p as 0.05 and changing m for different values, we get Fig. 1, which shows that $H(Y|M,R,E)$ can quickly converge to m as m decreases. However, it requires p larger than a threshold. For instance, if we set p as 0.03, it can not guarantee that $H(Y|M,R,E)$ converges to m.

To see the phase change caused by p, we fix the value of m and change p dynamically. As a result, we get Fig. 2. It shows that the phase-change point on p, where the performance of the approach based on random matrices drops dramatically, does not strongly depends on m. If the density of the random matrix M is lower than this point, the quality of the generated secret key may not be acceptable. If the density of the random matrix M is larger than this point, the quality of the generated secret key is prone to be stable.

REFERENCES

[1] R. Ahlswede and I. Csiszár,"Common randomness in information theory and cryptography–Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1121–1132, Jul. 1993.

[2] I. Ali-Khan, C. J. Broadbent, and J. C. Howell, "Large-alphabet quantum key distribution using energy-time entangled bipartite states," *Phys. Rev. Lett.* vol. 98, 060503, 2007.

[3] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, no. 2, p. 210, 1988.

[4] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.

[5] A. Bogdanov and S. Guo, "Sparse extractor families for all the entropy," arXiv:1207.6260, 2012.

[6] J. L. Carter and M. N.Wegman, "Universal classes of hash functions," *J. Comp. Syst. Sci.*, vol. 18, no. 2, pp. 143–154, 1979.

[7] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," in *Proc. ACM Symposium on the Theory of Computing (STOC)*, pp. 12-24, 1989.

[8] U. M. Maurer, "Secret key agreement by public discussion from common information", *IEEE Trans. Info. Theory*, vol. 39, pp. 733–742, 1993.

[9] U. M. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in Proc. *Advances in Cryptology-EUROCRYPT*, Bruges, Brugge, Belgium, pp. 356–373, May 2000.

[10] R. Renner and S. Wolf, "Smooth Rényi entropy and applications," in *Proc. International Symposium on Information Theory*, pp. 232, 2004.

[11] R. Shaltiel, "Recent developments in explicit constructions of extractors," in *Current trends in theoretical computer science. The Challenge of the New Century*, vol 1: Algorithms and Complexity, 2004.

[12] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.* vol. 28, pp. 656–715, 1949.

[13] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources", *IEEE Trans. Info. Theory*, vol. 19, pp. 471-480, 1973.

[14] A. Tanaka, M. Fujiwara, et.al. "High-speed quantum key distribution system for 1-Mbps real-time key generation," *IEEE J. Quantum Electronics*, vol. 48, 2012.

[15] L. Zhuo and V. K. Prasanna, "Sparse matrix-vector multiplication on FPGAs," in *Proceedings of the ACM/SIGDA international symposium on Field-programmable gate arrays*, pp. 63–74, 2005.