

# Real-Time Seedless Post-Processing for Quantum Random Number Generators

Qian Li<sup>1</sup> and Hongyi Zhou<sup>2,\*</sup>

<sup>1</sup>*Shenzhen International Center For Industrial And Applied Mathematics,  
Shenzhen Research Institute of Big Data, Shenzhen, China.*

<sup>2</sup>*State Key Lab of Processors, Institute of Computing Technology,  
Chinese Academy of Sciences, 100190, Beijing, China.*

Quantum-proof randomness extraction is essential for handling quantum side information possessed by a quantum adversary, which is widely applied in various quantum cryptography tasks. In this study, we introduce a real-time two-source quantum randomness extractor against quantum side information. Our extractor is tailored for forward block sources, a novel category of min-entropy sources introduced in this work. These sources retain the flexibility to accommodate a broad range of quantum random number generators. Our online algorithms demonstrate the extraction of a constant fraction of min-entropy from two infinitely long independent forward block sources. Moreover, our extractor is inherently block-wise parallelizable, presenting a practical and efficient solution for the timely extraction of high-quality randomness. Applying our extractors to the raw data of one of the most commonly used quantum random number generators, we achieve a simulated extraction speed as high as 64 Gbps.

**Keywords:** Quantum random number generators; quantum adversary; online two-source extractors; constant extraction rate; infinitely long sources.

## I. INTRODUCTION

Randomness extraction is a fundamental concept in information theory and cryptography [1], playing a key role in various applications, from secure communication to cryptographic key generation. Traditionally, randomness extractors have been designed to efficiently distill uniform randomness from weak-randomness sources against classical side information. However, in the era of quantum technology, the threat posed by quantum adversaries has attracted lots of attention. A quantum adversary has a quantum memory that shares quantum correlations with the randomness source. The information can be extracted by performing various measurements on the quantum memory, which is called quantum side information. An extractor dealing with classical side information may fail in the case of quantum side information [2]. Therefore, the need for quantum-proof randomness extractors that are able to provide information-theoretic security in the presence of a quantum adversary [3] has emerged as a critical concern.

It has been shown that no deterministic extractor can extract even one bit of randomness from one general weak-randomness source with large min entropy, even not from a Santha-Vazirani source [4] that is more structural. Therefore, two main approaches to extracting randomness from weak-randomness sources have been explored: seeded extractors and two-source extractors. Seeded extractors rely on a relatively short sequence of random numbers called random seed, which may not be uniform [5] but should be truly random. During the extraction procedure, the seed must be periodically updated to maintain randomness and security, even for strong extractors where the seed can be reused [6]. Managing and updating the seed securely can be a logistical challenge, especially in long-term deployments or scenarios where the system may be physically tampered with. On the other hand, two-source extractors leverage two independent weak-randomness sources to extract high-quality random bits, which avoids the issues above.

Substantial researches have been dedicated to two-source extractors. The first explicit construction of a two-source extractor appeared in [7], where the authors proposed that randomness can be extracted from two independent sources with an entropy rate of at least  $1/2$  based on the Lindsey's lemma. The following researches mainly aim at relaxing the entropy rate requirement [8–13]. In 2016, Chattopadhyay and Zuckerman [13] provided an explicit two-source extractor that works for two independent sources of  $n$  bits with an entropy rate of  $\log^C n/n$  for a large enough constant  $C$ , which is nearly optimal. The two-source extractor against quantum adversaries was first considered in [14], which shows the two-source extractor proposed in [15] is still secure if the quantum side information about the two sources is in the product form. This result was further improved in [16, 17] for more general quantum adversaries.

---

\* [zhouhongyi@ict.ac.cn](mailto:zhouhongyi@ict.ac.cn)

The existing two-source extractors against quantum adversary either output just one bit or suffer from limitations on the real-time applicability, i.e., the extraction procedure cannot begin until all raw data are generated. The demand for such real-time extractors arises from the fact that the generation rate of random bits is constrained by the post-processing speed [18, 19], especially for trusted-device quantum random number generators [20], which can hinder the timely extraction of high-quality randomness for applications like quantum key distribution. In response to this challenge, there is a pressing demand for the development of online algorithms capable of extracting randomness in real time. Besides, even worse, we are not aware of any existing quantum-proof two-source extractor that can extract a constant fraction of min-entropy from the two independent sources.

In this research, we propose two *real-time* two-source quantum-proof randomness extractors. In contrast with conventional quantum-proof randomness extractors characterizing the input raw data as min-entropy sources, we find that a large class of quantum random number generators, including one of the most commonly used trusted-device quantum random number generators [21], can be characterized by the so-called *forward block sources* [22]. This insight allows us to develop more efficient online extraction algorithms. Both of our extractors process input raw data on-the-fly, partitioning it into blocks sequentially based on the time order of arrival: they employ the inner product function in specific finite fields on each block. In our first extractor, all blocks share a uniform length that is logarithmic in the overall input raw data length. This property makes the first extractor hardware-friendly and achieve a high extraction speed. While our second extractor utilizes incremental block lengths, rendering it capable of handling infinitely long input raw data. Both extractors can extract a constant fraction of min-entropy from the raw data. Additionally, the independence of applying the inner product function to each block facilitates natural parallel implementation of our extractors.

## II. PRELIMINARIES

Throughout the paper, we use capital and lowercase letters (e.g.,  $X$  and  $x$ ) to represent random variables and their assignments, respectively. We use  $U_m$  to represent the perfectly uniform random variable on  $m$ -bit strings and  $\rho_{U_m}$  to represent the  $m$ -dimensional maximally mixed state. The terms “quantum adversary” and “quantum side information” will be used interchangeably.

**Definition II.1** (Conditional min-entropy). *Let  $Y$  be a classical random variable that takes value  $y$  with probability  $P_y$  and  $\mathcal{E}$  be a quantum system. The state of the composite system can be written as  $\rho_{Y\mathcal{E}} = \sum_y P_y |y\rangle\langle y| \otimes \rho_{\mathcal{E}}^y$ , where  $\{|y\rangle\}_y$  forms an orthonormal basis. The conditional min-entropy of  $Y$  given  $\mathcal{E}$  is  $H_{\min}(Y|\mathcal{E})_{\rho_{Y\mathcal{E}}} = -\log_2 p_{\text{guess}}(Y|\mathcal{E})$ , where  $p_{\text{guess}}(Y|\mathcal{E})$  is the maximum average probability of guessing  $Y$  given the quantum system  $\mathcal{E}$ . That is,*

$$p_{\text{guess}}(Y|\mathcal{E}) = \max_{\{E_{\mathcal{E}}^y\}_y} \left[ \sum_y P_y \text{Tr} \left( E_{\mathcal{E}}^y \rho_{\mathcal{E}}^y \right) \right], \quad (1)$$

where the maximization is taken over all positive operator-valued measures (POVMs)  $\{E_{\mathcal{E}}^y\}_y$  on  $\mathcal{E}$ . When  $\rho_{Y\mathcal{E}}$  is clear from the context, we will denote the conditional min-entropy as  $H_{\min}(Y|\mathcal{E})$  for brevity.

In this paper, we call the raw data generated by a QRNG a random source. A general random source is the min-entropy source, where the conditional min-entropy is lower bounded. We consider how to extract randomness from two sequences of raw data generated by two separated QRNGs. Here, we assume that the two separated QRNGs are in product, then the quantum side information is also in the product form. This scenario corresponds to two independent quantum adversaries, each tampering with one of the sources, and then jointly trying to guess the extractor’s output. Formally, we have the following definition.

**Definition II.2** (Product quantum side information [14]). *Let  $X$  and  $Y$  be two independent sources and  $\mathcal{E}$  be the quantum side information. We say that the quantum side information is product if the joint state  $\rho_{XY\mathcal{E}} = \rho_{X\mathcal{E}_1} \otimes \rho_{Y\mathcal{E}_2}$ . Note that then  $H_{\min}(X|\mathcal{E}) = H_{\min}(X|\mathcal{E}_1)$  and  $H_{\min}(Y|\mathcal{E}) = H_{\min}(Y|\mathcal{E}_2)$ .*

**Definition II.3** (Min-entropy quantum-proof two-source extractor). *A function  $\text{Ext} : \{0,1\}^t \times \{0,1\}^t \rightarrow \{0,1\}^m$  is a  $(t, k, \epsilon)$  min-entropy quantum-proof two-source extractor, if for every sources  $(X, Y)$  and product quantum side information  $\mathcal{E}$  where  $H_{\min}(X|\mathcal{E}) \geq k$  and  $H_{\min}(Y|\mathcal{E}) \geq k$ , we have*

$$\frac{1}{2} \|\rho_{\text{Ext}(X,Y)\mathcal{E}} - \rho_{U_m} \otimes \rho_{\mathcal{E}}\| \leq \epsilon, \quad (2)$$

where  $\|\cdot\|$  denotes the trace norm defined by  $\|A\| = \text{Tr}\sqrt{A^\dagger A}$ .

### III. MAIN RESULT

We use  $X = X_1 X_2 \cdots X_N \in (\{0, 1\}^b)^N$  and  $Y = Y_1 Y_2 \cdots Y_N \in (\{0, 1\}^b)^N$  to denote the raw data generated by two QRNGs, both consisting of  $N$  samples each of  $b$  bits. For a set  $I \subset \mathbb{N}^+$ , we write  $X_I$  for the restriction of  $X$  to the samples determined by  $I$ . For example, if  $I = \{1, 3, 7\}$ , then  $X_I = X_1 X_3 X_7$ . We use  $\mathcal{E}$  to denote the quantum system possessed by the quantum adversary.

#### A. Model the raw data as forward block sources

In contrast with conventional quantum-proof randomness extractors characterizing the input raw data as min-entropy sources, we model the input raw data as *forward block sources*, defined as follows.

**Definition III.1** (Forward block source). *A string of random variables  $X = X_1 \cdots X_N \in (\{0, 1\}^b)^N$  is a  $(b, N, \delta)$ -forward block source given a quantum system  $\mathcal{E}$  if for every  $1 \leq k \leq N$  and every  $x_1, x_2, \dots, x_{k-1}, \mathcal{E}$ ,*

$$H_{\min}(X_k, X_{k+1}, \dots, X_i | X_1 = x_1, \dots, X_{k-1} = x_{k-1}, \mathcal{E}) \geq (i - k + 1)\delta \cdot b. \quad (3)$$

Intuitively, Eq. (3) implies each sample  $X_i$  brings at least  $\delta b$  new randomness. The notion of forward block sources is more special than the min-entropy sources, so it is possible to design extractors with better performance. Fortunately, the notion of forward block sources is still quite general so that it can characterize the input raw data generated by a large class of QRNGs. When the independent and identically distributed (i.i.d.) assumption holds, the raw data straightforwardly satisfies the forward block source. For example, the trusted-devices QRNGs based on single photon detection [23–25], vacuum fluctuations [26–28], and photon arrival time [29–31] can generate independent raw data satisfying the forward block source. Besides, the raw data of the (semi-)device-independent QRNG under the independent and identically distributed (i.i.d.) assumption also satisfies the forward block source if non-zero min-entropy lower bound is calculated in test rounds. Moreover, for QRNGs with correlated raw data, one can construct appropriate physical models to check whether Eq. (3) is satisfied. For example, the ones based on homodyne detection [21] with finite bandwidth can both be shown to be a forward block source [22]. Finally, we remark that there does not exist any nontrivial deterministic extractor for a single block source.

#### B. Gadget: a construction of min-entropy quantum-proof two-source extractor

In this subsection, we provide a construction of min-entropy quantum-proof two-source extractor (see Definition III.2), which will be used as a gadget. Let  $\mathbb{F}_{2^q}$  denote the finite field on  $2^q$  elements. We identify  $\mathbb{F}_{2^q}$  with  $\{0, 1\}^q$  and  $\mathbb{F}_{2^q}^n$  with  $(\{0, 1\}^q)^n$  where the addition operation “ $+_q$ ” in the field  $\mathbb{F}_{2^q}$  is the bit-wise parity operation in  $\{0, 1\}^q$ .

**Definition III.2** (Inner-product extractor). *The  $(n, q)$  inner-product extractor  $\text{Ext}_{IP}^{n,q} : (\{0, 1\}^q)^n \times (\{0, 1\}^q)^n \rightarrow \{0, 1\}^q$  is defined as  $\text{Ext}_{IP}^{n,q}(x, y) = \sum_{i=1}^n x_i y_i$ . Here,  $x = x_1 \cdots x_n \in (\{0, 1\}^q)^n$  and  $y = y_1 \cdots y_n \in (\{0, 1\}^q)^n$ , and the addition operations and multiplication operations are in the field  $\mathbb{F}_{2^q}$ .*

**Theorem III.3.**  $\text{Ext}_{IP}^{n,q}$  is a  $(qn, \delta qn, \sqrt{3} \cdot 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8})qn + 2q})$  min-entropy quantum-proof two-source extractor.

Theorem III.3 immediately follows from the following two lemmas, namely Lemma III.4 and Lemma III.5.

**Lemma III.4** (Classical-Quantum XOR-Lemma [14]). *Let  $Z$  be a classical random variable of  $q$  bits, and  $\mathcal{E}$  be a quantum system. Then*

$$\|\rho_{Z\mathcal{E}} - \rho_{U_q} \otimes \rho_{\mathcal{E}}\| \leq 2^q \sum_{S \in \{0, 1\}^q \setminus \{0^q\}} \|\rho_{(S \cdot Z) \circ \mathcal{E}} - \rho_{U_1} \otimes \rho_{\mathcal{E}}\|.$$

Here  $S \cdot Z := \sum_{i=1}^n S_i \cdot Z_i \pmod{2}$  is the inner product over  $\mathbb{F}_2$ .

The following lemma will also be used, whose proof can be found in Appendix A.

**Lemma III.5.** *Let  $X, Y$  be random variables on  $(\{0, 1\}^q)^n$  and  $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$  be the product quantum side information where  $H_{\min}(X|\mathcal{E}_1) \geq \delta qn$  and  $H_{\min}(Y|\mathcal{E}_2) \geq \delta qn$ . Let  $Z = \text{Ext}_{IP}^{n,q}(X, Y)$ . Then for any  $S \in \{0, 1\}^q \setminus \{0^q\}$ , we have*

$$\frac{1}{2} \|\rho_{(S \cdot Z) \circ \mathcal{E}} - \rho_{U_1} \otimes \rho_{\mathcal{E}}\| \leq \sqrt{3} \cdot 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8})qn}.$$

### C. Online Seedless Extractors for forward block sources

In this subsection, we design two online seedless quantum-proof extractors that can both extract a constant fraction of the min-entropy from forward block sources. The two extractors both proceed in the following fashion: partition the input raw data  $X, Y$  into blocks, and apply the inner-product extractor  $\text{Ext}_{IP}$  to each block separately. Thus the two extractors can be paralleled naturally. The two extractors are described in Algorithms 1 and 2 respectively. Given the input raw data  $X, Y$  each of length  $N$ , the first extractor, named  $\text{Ext}_{bs}^{\text{eq}}$ , evenly partitions the input raw data into blocks each of size  $O(\log N)$ , and the second extractor, named  $\text{Ext}_{bs}^{\text{neq}}$ , requires incremental block lengths. Compared to the first extractor, the second one is less hardware-friendly and sacrifices the extraction speed in general. On the other hand, it enjoys the property that it can deal with infinite raw data, without the need to determine the raw data length prior to extraction.

---

#### Algorithm 1: $\text{Ext}_{bs}^{\text{eq}}$

---

```

1 Input: Two independent  $(b, N, \delta)$ -forward block sources  $X = X_1 \cdots X_N$  and  $Y = Y_1 \cdots Y_N$ . And  $0 < \epsilon < 1$ ;
2 Let  $n := \lceil \frac{24}{2\delta-1} \rceil$ ,  $q := b \lceil \log_2 \left( \frac{N}{\epsilon n} \right) / b \rceil$ , and  $i := 1$ ;
3 for  $\ell = 1$  to  $Nb/qn$  do
4   Let  $I_\ell := [i, i - 1 + qn/b]$ ;
5   Compute  $Z^{(\ell)} := \text{Ext}_{IP}^{n,q}(X_{I_\ell}, Y_{I_\ell})$ ;
6   Let  $i := i + n$ ;
7   Output  $Z^{(\ell)}$ .

```

---



---

#### Algorithm 2: $\text{Ext}_{bs}^{\text{neq}}$

---

```

1 Input: Two  $(b, \infty, \delta)$ -forward block sources  $X = X_1 \cdots X_\infty$  and  $Y = Y_1 \cdots Y_\infty$  independent conditioned on quantum
adversary  $\mathcal{E}$ ;
2 Parameter:  $q_1, \Delta \in \mathbb{N}^+$ ;
3 Let  $i := 1$  and  $n := \lceil \frac{24}{2\delta-1} \rceil$ ;
4 for  $\ell = 1$  to  $\infty$  do
5   Let  $I_\ell := [i, i - 1 + q_\ell n/b]$ ;
6   Compute  $Z^{(\ell)} := \text{Ext}_{IP}^{n,q_\ell}(X_{I_\ell}, Y_{I_\ell})$ ;
7   Let  $i := i + q_\ell n/b$  and  $q_{\ell+1} := q_\ell + \Delta \cdot b$ ;
8   Output  $Z^{(\ell)}$ .

```

---

Note that if we impose  $\Delta = 0$  and let  $q_1 = b \lceil \log_2 \left( \frac{N}{\epsilon n} \right) / b \rceil$ , then  $\text{Ext}_{bs}^{\text{neq}}$  becomes  $\text{Ext}_{bs}^{\text{eq}}$ . We first analyze the second extractor  $\text{Ext}_{bs}^{\text{neq}}$ .

**Theorem III.6.** *The extractor  $\text{Ext}_{bs}^{\text{neq}}$  satisfies the following properties. For any  $k \in \mathbb{N}$ ,*

$$\frac{1}{2} \|\rho_{Z^{(1)} \circ Z^{(2)} \circ \dots \circ Z^{(k)}} \circ \mathcal{E} - \rho_{U_{m_k}} \otimes \rho_{\mathcal{E}}\| \leq \sqrt{3} \cdot \sum_{\ell=1}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8})q_\ell n + 2q_\ell} \leq \sqrt{3} \cdot \sum_{\ell=1}^k 2^{-\frac{1}{4} - q_\ell} < \sqrt{3} \cdot \frac{2^{-\frac{1}{4} - q_1}}{1 - 2^{-\Delta b}}, \quad (4)$$

where  $m_k = \sum_{\ell=1}^k q_\ell = kq_1 + \frac{(k-1)k\Delta b}{2}$ .

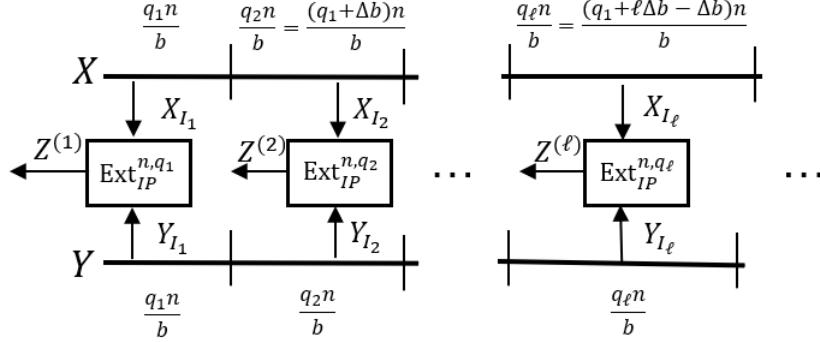


FIG. 1: Illustration of  $\text{Ext}_{\text{bs}}^{\text{neq}}$ .

*Proof.* Fix  $k \in \mathbb{N}$ . For convenience of presentation, for any  $0 \leq \ell \leq k$ , we use  $Z^{>\ell}$  to represent  $Z^{(\ell+1)} \circ Z^{(\ell+2)} \circ \dots \circ Z^{(k)}$ ,  $X^{\leq \ell}$  to represent  $X_{I_1} X_{I_2} \cdots X_{I_\ell}$ , and  $Y^{\leq \ell}$  to represent  $Y_{I_1} Y_{I_2} \cdots Y_{I_\ell}$ . Besides, let  $\tilde{q}_\ell := q_{\ell+1} + q_{\ell+2} + \dots + q_k$ .

In fact, we will prove that for any  $0 \leq \ell \leq k$ , it has

$$\frac{1}{2} \left\| \rho_{Z^{>\ell} X^{\leq \ell} Y^{\leq \ell} \mathcal{E}} - \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{X^{\leq \ell} Y^{\leq \ell} \mathcal{E}} \right\| \leq \sqrt{3} \cdot \sum_{j=\ell+1}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_j n + 2q_j}, \quad (5)$$

which implies the theorem immediately by letting  $\ell = 0$ .

The proof is by an induction on  $\ell$ . The base case when  $\ell = k$  is trivial. The induction proceeds as follows. Assume Eq. (5) is true. Then due to the contractivity of trace-preserving quantum operations, we have

$$\frac{1}{2} \left\| \rho_{Z^{>\ell} \text{Ext}_{IP}(X_{I_\ell}, Y_{I_\ell}) X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} - \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{\text{Ext}_{IP}(X_{I_\ell}, Y_{I_\ell}) X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} \right\| \leq \sqrt{3} \cdot \sum_{j=\ell+1}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_j n + 2q_j}. \quad (6)$$

On the other hand, by Definition III.1, for any assignment  $x^{\leq \ell-1}$  of  $X^{\leq \ell-1}$  and  $y^{\leq \ell-1}$  of  $Y^{\leq \ell-1}$ , we have

$$H_{\min}(X_{I_\ell} | X^{\leq \ell-1} = x^{\leq \ell-1}, \mathcal{E}) \geq \delta q_\ell n, \text{ and } H_{\min}(Y_{I_\ell} | Y^{\leq \ell-1} = y^{\leq \ell-1}, \mathcal{E}) \geq \delta q_\ell n. \quad (7)$$

Then, recalling that  $\text{Ext}_{IP}$  is a min-entropy quantum-proof two-source extractor (Theorem III.3), it follows that

$$\frac{1}{2} \left\| \rho_{\text{Ext}_{IP}(X_{I_\ell}, Y_{I_\ell}) X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} - \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} \right\| \leq \sqrt{3} \cdot 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_\ell n + 2q_\ell}.$$

Thus,

$$\frac{1}{2} \left\| \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{\text{Ext}_{IP}(X_{I_\ell}, Y_{I_\ell}) X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} - \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{U_{q_\ell}} \otimes \rho_{X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} \right\| \leq \sqrt{3} \cdot 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_\ell n + 2q_\ell}. \quad (8)$$

Finally, combining inequalities (6) and (8) and applying the triangle inequality, we conclude that

$$\frac{1}{2} \left\| \rho_{Z^{>\ell} \text{Ext}_{IP}(X_{I_\ell}, Y_{I_\ell}) X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} - \rho_{U_{\tilde{q}_\ell}} \otimes \rho_{U_{q_\ell}} \otimes \rho_{X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} \right\| \leq \sqrt{3} \cdot \sum_{j=\ell}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_j n + 2q_j} \quad (9)$$

That is,

$$\frac{1}{2} \left\| \rho_{Z^{>\ell-1} X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} - \rho_{U_{\tilde{q}_{\ell-1}}} \otimes \rho_{X^{\leq \ell-1} Y^{\leq \ell-1} \mathcal{E}} \right\| \leq \sqrt{3} \cdot \sum_{j=\ell}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8}) q_j n + 2q_j},$$

So Eq. (5) also holds for  $\ell - 1$ . The proof is finished.  $\square$

As can be seen from the proof, to extract infinitely long randomness, the parameter  $\Delta$  of  $\text{Ext}_{\text{bs}}^{\text{neq}}$  must be strictly positive, since the upper bound  $\sqrt{3} \cdot \sum_{\ell=1}^k 2^{-\frac{1}{4} - (\frac{\delta}{4} - \frac{1}{8})q_\ell n + 2q_\ell}$  on the error converges if and only if  $\Delta > 0$ . Moreover, the infinitely long output string  $Z^{(1)} \circ Z^{(2)} \circ \dots$  can be arbitrarily close to the uniform distribution by choosing a sufficiently large constant  $q_1$ . In addition,  $\text{Ext}_{\text{bs}}^{\text{neq}}$  extracts  $\frac{q_\ell}{2\delta q_\ell n} \approx \frac{2\delta-1}{48\delta}$  fraction of the min-entropy of the raw data  $X_{I_\ell}, Y_{I_\ell}$ . Finally, we remark there is a tradeoff between the error and the computational complexity by setting  $\Delta$ . As  $\Delta$  increases, the error becomes smaller, but the size of block increases, which leads to heavy computations.

We intuitively explain Theorem III.6, i.e., how our algorithms work for the forward block sources. When the input raw data  $X, Y$  are forward block sources, the block  $X_{I_\ell}$  and  $Y_{I_\ell}$  are independent and each has large min-entropy conditioned on any assignment  $(x^{<\ell}, y^{<\ell})$  of previous blocks  $(X^{<\ell}, Y^{<\ell})$ , which implies that the output  $Z^{(\ell)}$  of this block is (approximately) uniform conditioned on any assignment  $(x^{<\ell}, y^{<\ell})$ . Note that  $z^{<\ell}$  can be fully determined by  $(x^{<\ell}, y^{<\ell})$ . So  $Z^{(\ell)}$  is (approximately) uniform conditioned on any assignment  $z^{<\ell}$  of  $Z^{<\ell}$ . By induction on  $\ell$ , we can conclude that the whole output  $Z^{(1)} \circ \dots \circ Z^{(k)} \circ \dots$  is (approximately) uniform.

As a corollary of Theorem III.6, we have the following result for  $\text{Ext}_{\text{bs}}^{\text{eq}}$ .

**Theorem III.7.** *The output  $Z^{(1)} \circ \dots \circ Z^{(Nb/qn)}$  of  $\text{Ext}_{\text{bs}}^{\text{eq}}$  satisfies that*

$$\frac{1}{2} \|\rho_{Z^{(1)} \circ Z^{(2)} \circ \dots \circ Z^{(Nb/qn)} \circ \mathcal{E}} - \rho_{U_m} \otimes \rho_{\mathcal{E}}\| \leq \epsilon, \quad (10)$$

where  $m := \frac{Nb}{n}$ .

#### IV. SIMULATIONS OF THE REAL-TIME RANDOMNESS GENERATION RATE

In this section, we provide an analysis of the computational efficiency of our extractors. For simplicity of discussion, we focus on the first extractor  $\text{Ext}_{\text{bs}}^{\text{eq}}$ .

##### A. Asymptotic analysis.

Suppose we apply  $\text{Ext}_{\text{bs}}^{\text{eq}}$  on two independent  $(b, N, \delta)$ -forward block sources. The extractor will set  $n = O(1)$  and  $q = O(\log N)$ , and then compute the inner product function  $\text{Ext}_{IP}^{n,q}$  for  $Nb/qn = O(Nb/\log N)$  times. The computation of  $\text{Ext}_{IP}^{n,q}$  involves  $(n-1)$  additions ' $+_q$ ' and  $n$  multiplications ' $\cdot_q$ ' over the finite field  $\mathbb{F}_{2^q}$ . The addition ' $+_q$ ' is the bit-wise parity in  $\{0, 1\}^q$  and thus can be implemented by  $q$  parity gates. The multiplication ' $\cdot_q$ ' can be implemented by a  $O(q \log q \log \log q)$ -size Boolean circuit (see the last inequality in Section 2 and the third paragraph on Page 3 in [32]). So  $\text{Ext}_{IP}^{n,q}$  can be computed by a  $O(nq \log q \log \log q)$ -size Boolean circuit. Therefore, it requires  $O(Nb/\log N) \cdot O(nq \log q \log \log q) = O(Nb \log \log N \log \log \log N)$  logic operations in total to execute  $\text{Ext}_{\text{bs}}^{\text{eq}}$ .

We compare the computational complexity of  $\text{Ext}_{\text{bs}}^{\text{eq}}$  with the seeded extractor proposed in [22], which evenly partitions the input raw data into blocks each of size  $O(\log N)$  and applies the Toeplitz-hashing extractor to each block separately with the same seed of  $O(\log N)$  bits. Note that it needs  $O(\log^2 N)$  logic operations to multiply a  $O(\log N) \times O(\log N)$  dimensional Toeplitz matrix with a  $O(\log N)$  dimensional vector. Thus, it requires  $O(Nb/\log N) \cdot O(\log^2 N) = O(Nb \log N)$  logic operations in total to execute the seeded extractor. So  $\text{Ext}_{\text{bs}}^{\text{eq}}$  is asymptotically faster than the seeded extractor.

##### B. Simulations.

In the following, we make a simulation estimating the extraction speed of the first extractor  $\text{Ext}_{\text{bs}}^{\text{eq}}$  implemented in the Xilinx Kintex-7 XC7K480T Field Programmable Gate Array (FPGA), a common application in industry. We use the raw data from Ref. [21] as the random source, which is a forward block source with parameters  $b = 16$  and  $\delta = 10.74/16 \approx 0.67$ . We consider a raw data length of  $N = 2^{51}$  bits and a total security parameter  $\epsilon = 2^{-30}$ , which means the final output data is  $2^{-30}$ -close to a uniform distribution. Correspondingly, the algorithmic parameters  $n, q$  of  $\text{Ext}_{\text{bs}}^{\text{eq}}$  are specified to  $n = \lceil 24/(2\delta - 1) \rceil = 71$  and  $q := b \lceil \log_2 (N/\epsilon n) / b \rceil = 80$ . By Theorem III.7,  $\text{Ext}_{\text{bs}}^{\text{eq}}$  uses no seed and will output about 0.45 PB random bits.

Precisely,  $\text{Ext}_{\text{bs}}^{\text{eq}}$  simply divides the input raw data  $X, Y$  into blocks each containing  $71 \times 80$  bits, and then apply  $\text{Ext}_{IP}^{71,80}$  on each block  $(X_{I_t}, Y_{I_t})$ . The inner product function  $\text{Ext}_{IP}^{71,80}$  involves  $71 - 1 = 70$  addition operations and 71 multiplication operations in the field  $\mathbb{F}_{2^{80}}$ . The addition operation in  $\mathbb{F}_{2^{80}}$  is the bit-wise parity operation in  $\{0, 1\}^{80}$ , and so can be computed in 80 ‘ $\oplus$ ’ operations. Each multiplication operation in  $\mathbb{F}_{2^{80}}$  can be computed by at most 4885 ‘ $\oplus$ ’ or ‘ $\wedge$ ’ operations [32]. Thus, each  $\text{Ext}_{IP}^{71,80}$  can be implemented by at most  $80 \times (71 - 1) + 4885 \times 71 = 352435$  ‘ $\oplus$ ’ or ‘ $\wedge$ ’ operations.

The parameters of the Xilinx Kintex-7 XC7K480T FPGA are as follows. The clock rate is set to be 200 MHz; the number of Look-Up-Tables (LUTs) is  $3 \times 10^5$ ; each LUT can perform 5 basic logical operations (e.g., ‘ $\oplus$ ’ or ‘ $\wedge$ ’) simultaneously. To make full use of the FPGA, we can perform the matrix multiplications of  $[3 \times 10^5 \times 5 / 352435] = 4$  blocks in parallel. Therefore, the extraction speed of  $\text{Ext}_{\text{bs}}^{\text{eq}}$  is  $200 \times 10^6 \times 4 \times 80 = 64$  Gbps, which is significantly improved by one order of magnitude compared to the state-of-the-art result with a speed of 18.8 Gbps [19]. As a result, our online extraction is adequate for the post-processing of the QRNG in [21].

## V. CONCLUSION

In conclusion, we introduce a real-time two-source quantum-proof randomness extractor, a crucial advancement for applications such as quantum key distribution. The key contribution lies in the design of online algorithms based on the concept of a forward block source. We demonstrate that a constant fraction of min-entropy can be extracted from two arbitrarily long independent forward block sources. Notably, this extractor accommodates a broad class of quantum random number generators. The inherent parallelizability of our extractor enhances its practicality for real-world applications.

Looking forward, there are several open questions. Firstly, our extractor currently necessitates a min-entropy rate  $\delta > 1/2$ . It remains an intriguing challenge to explore whether this restriction can be dropped or imitated without compromising the security and efficiency of the extraction process. Secondly, the design of an online extractor for more general sources, or even the most general min-entropy source, remains an open avenue for future research. These questions underscore the ongoing pursuit of robust quantum-proof randomness extraction techniques with broader applicability and enhanced flexibility. Our work lays a foundation for these inquiries, marking a significant step towards securing communication and cryptographic protocols in the quantum era.

- 
- [1] R. Impagliazzo and D. Zuckerman, in *FOCS*, Vol. 30 (1989) pp. 248–253.
  - [2] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. De Wolf, in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing* (2007) pp. 516–525.
  - [3] R. Konig and R. Renner, *IEEE Transactions on Information Theory* **57**, 4760 (2011).
  - [4] M. Santha and U. V. Vazirani, *J. Comput. Syst. Sci.* **33**, 75–87 (1986).
  - [5] A. De, C. Portmann, T. Vidick, and R. Renner, *SIAM Journal on Computing* **41**, 915 (2012).
  - [6] N. Nisan and D. Zuckerman, *Journal of Computer and System Sciences* **52**, 43 (1996).
  - [7] B. Chor and O. Goldreich, *SIAM Journal on Computing* **17**, 230 (1988).
  - [8] R. Raz, in *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing* (2005) pp. 11–20.
  - [9] J. Bourgain, *International Journal of Number Theory* **01**, 1 (2005).
  - [10] B. Barak, R. Impagliazzo, and A. Wigderson, *SIAM Journal on Computing* **36**, 1095 (2006).
  - [11] X. Li, in *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA* (IEEE Computer Society, 2013) pp. 100–109.
  - [12] X. Li, in *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, edited by V. Guruswami (IEEE Computer Society, 2015) pp. 863–882.
  - [13] E. Chattopadhyay and D. Zuckerman, in *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, edited by D. Wichs and Y. Mansour (ACM, 2016) pp. 670–683.
  - [14] R. Kasher and J. Kempe, *Theory Comput.* **8**, 461 (2010).
  - [15] Y. Dodis, A. Elbaz, R. Oliveira, and R. Raz, in *Approximation, Randomization, and Combinatorial Optimization, Algorithms and Techniques, 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2004, and 8th International Workshop on Randomization and Computation, RANDOM 2004, Cambridge, MA, USA, August 22-24,*

- 2004, *Proceedings*, Lecture Notes in Computer Science, Vol. 3122, edited by K. Jansen, S. Khanna, J. D. P. Rolim, and D. Ron (Springer, 2004) pp. 334–344.
- [16] K.-M. Chung, X. Li, and X. Wu, arXiv preprint arXiv:1411.2315 (2014).
  - [17] R. A. Friedman, C. Portmann, and V. B. Scholz, in *11th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2016, September 27-29, 2016, Berlin, Germany*, LIPIcs, Vol. 61, edited by A. Broadbent (Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016) pp. 2:1–2:34.
  - [18] X.-G. Zhang, Y.-Q. Nie, H. Zhou, H. Liang, X. Ma, J. Zhang, and J.-W. Pan, *Review of Scientific Instruments* **87**, 076102 (2016), <https://aip.scitation.org/doi/pdf/10.1063/1.4958663>.
  - [19] B. Bai, J. Huang, G.-R. Qiao, Y.-Q. Nie, W. Tang, T. Chu, J. Zhang, and J.-W. Pan, *Applied Physics Letters* **118**, 264001 (2021).
  - [20] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, *npj Quantum Information* **2**, 1 (2016).
  - [21] T. Gehring, C. Lupo, A. Kordts, D. Solar Nikolic, N. Jain, T. Rydberg, T. B. Pedersen, S. Pirandola, and U. L. Andersen, *Nature Communications* **12**, 605 (2021).
  - [22] Q. Li, X. Sun, X. Zhang, and H. Zhou, Improved real-time post-processing for quantum random number generators (2023), [arXiv:2301.08621 \[quant-ph\]](https://arxiv.org/abs/2301.08621).
  - [23] J. G. Rarity, P. Owens, and P.R.Tapster, *Journal of Modern Optics* **41**, 2435 (1994), <http://dx.doi.org/10.1080/09500349414552281>.
  - [24] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, *Journal of Modern Optics* **47**, 595 (2000), <http://dx.doi.org/10.1080/09500340008233380>.
  - [25] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, *Review of Scientific Instruments* **71**, 1675 (2000).
  - [26] C. Gabriel, C. Wittmann, C. Marquardt, and G. Leuchs, *Nature Photonics* **4**, 711 (2010).
  - [27] T. Symul, S. M. Assad, and P. K. Lam, *Applied Physics Letters* **98**, 231103 (2011).
  - [28] M. Jofre, M. Curty, F. Steinlechner, G. Anzolin, J. P. Torres, M. W. Mitchell, and V. Pruneri, *Optics Express* **19**, 20665 (2011).
  - [29] M. Wahl, M. Leifgen, M. Berlin, T. Rohlicke, H.-J. Rahn, and O. Benson, *Applied Physics Letters* **98**, 171105 (2011).
  - [30] S. Li, L. Wang, L.-A. Wu, H.-Q. Ma, and G.-J. Zhai, *Journal of the Optical Society of America A* **30**, 124 (2013).
  - [31] Y.-Q. Nie, H.-F. Zhang, Z. Zhang, J. Wang, X. Ma, J. Zhang, and J.-W. Pan, *Applied Physics Letters* **104**, 051110 (2014).
  - [32] J. Rifà and J. Borrell, in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 9th International Symposium, AAEC-9, New Orleans, LA, USA, October 7-11, 1991, Proceedings*, Lecture Notes in Computer Science, Vol. 539, edited by H. F. Mattson, T. Mora, and T. R. N. Rao (Springer, 1991) pp. 352–359.

### Appendix A: Proof of Lemma III.5

In this section, we prove Lemma III.5. Let  $X, Y$  be random variables on  $(\{0, 1\}^q)^n$  and  $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$  be the product quantum side information where  $H_{\min}(X|\mathcal{E}_1) \geq \delta qn$  and  $H_{\min}(Y|\mathcal{E}_2) \geq \delta qn$ . Let  $Z = \text{Ext}_{IP}^{n,q}(X, Y)$ , and fix  $S \in \{0, 1\}^q \setminus \{0^q\}$  arbitrarily. We will use “ $\times$ ” to denote the multiplication operation in the field  $\mathbb{F}_{2^q}$ .

Note that for any  $S \in \{0, 1\}^q$ , there exists a  $a \in \mathbb{F}_{2^q}$  such that  $S \cdot Z$  corresponds exactly to the first bit of  $a \times Z$  viewed as a vector in  $\{0, 1\}^q$ . Moreover this correspondence between  $S$  and  $a$  is a bijection in the sense that different  $S \in \{0, 1\}^q$  correspond to different  $a \in \mathbb{F}_{2^q}$ . The special case when  $S = 0^q$  corresponds to the case  $a = 0$ . Thus, for any  $S \in \{0, 1\}^q \setminus \{0^q\}$ , there exists some non-zero  $a \in \mathbb{F}_{2^q}$  such that  $S \cdot Z = [a \times Z]_1$ . Here,  $[a \times Z]_1$  is used to denote the first bit of  $a \times Z$  viewed as a vector in  $\{0, 1\}^q$ .

First, we show that the function  $[a \times \text{Ext}_{IP}^{n,q}(\cdot, \cdot)]_1$  is a hadamard function.

**Definition A.1** (hadamard function). *A function  $f : \{0, 1\}^t \times \{0, 1\}^t \rightarrow \{0, 1\}$  is called a hadamard function if for any two distinct  $x, x' \in \{0, 1\}^t$ , it has  $\sum_{y \in \{0, 1\}^t} (-1)^{f(x,y)+f(x',y)} = 0$ .*

**Proposition A.2.** *For any non-zero  $a \in \mathbb{F}_{2^q}$ ,  $[a \times \text{Ext}_{IP}^{n,q}(\cdot, \cdot)]_1$  is a hadamard function.*

*Proof.* For simplicity of notations, we let  $f_a(x, y) := [a \times \text{Ext}_{IP}^{n,q}(\cdot, \cdot)]_1$ . Given any two distinct  $x = x_1 \cdots x_n, x' = x'_1 \cdots x'_n \in (\{0, 1\}^q)^n$ , w.l.o.g. we assume  $x_1 \neq x'_1$ . Recalling that the addition operation “ $+_q$ ” in the field  $\mathbb{F}_{2^q}$  corresponds to the bitwise-parity operation in  $\{0, 1\}^q$ , we have

$$\begin{aligned} \sum_{y \in (\{0, 1\}^q)^n} (-1)^{f_a(x,y)+f_a(x',y)} &= \sum_{y \in (\{0, 1\}^q)^n} (-1)^{f_a(x,y)-f_a(x',y)} = \sum_{y \in (\{0, 1\}^q)^n} (-1)^{f_a(x-x',y)} \\ &= \sum_{y \in (\{0, 1\}^q)^n} (-1)^{[a \times \text{Ext}_{IP}^{n,q}(x-x',y)]_1} = \sum_{y \in (\{0, 1\}^q)^n} (-1)^{[a \times (x_1-x'_1) \times y_1]_1 + [\sum_{i=2}^n a \times (x_i-x'_i) \times y_i]_1} \\ &= \left( \sum_{y_1 \in \{0, 1\}^q} (-1)^{[a \times (x_1-x'_1) \times y_1]_1} \right) \cdot \left( \sum_{y_2, \dots, y_n \in \{0, 1\}^q} (-1)^{[\sum_{i=2}^n a \times (x_i-x'_i) \times y_i]_1} \right). \end{aligned}$$

Note that  $a \times (x_1 - x'_1) \neq 0$ , so

$$|\{y_1 \in \{0,1\}^q | [a \times (x_1 - x'_1) \times y_1]_1 = 1\}| = |\{y_1 \in \{0,1\}^q | [a \times (x_1 - x'_1) \times y_1]_1 = 0\}| = 2^{q-1}.$$

Thus we get the conclusion immediately by Definition A.1.  $\square$

**Definition A.3** (One-bit-output two-source extractor). *A function  $f : \{0,1\}^t \times \{0,1\}^t \rightarrow \{0,1\}$  is a  $(t, k, \epsilon)$  one-bit-output two-source extractor, if for every pair of independent sources  $(X, Y)$  with min-entropy  $\geq k$  each, we have*

$$(1 - \epsilon)/2 \leq P[f(X, Y) = 1] \leq (1 + \epsilon)/2.$$

**Lemma A.4** ([7]). *Let  $f : \{0,1\}^t \times \{0,1\}^t \rightarrow \{0,1\}$  be a hadamard function. Then  $f$  is a  $(t, k, 2^{1-\frac{2k-t}{2}})$  one-bit-output two-source extractor for any  $k \leq t$ .*

*In other words, let  $X, Y$  be independent random variables on  $\{0,1\}^t$  with min-entropy  $\geq k$  each. Then  $(1 - \epsilon)/2 \leq P[f(X, Y) = 1] \leq (1 + \epsilon)/2$  where  $\epsilon = 2^{1-\frac{2k-t}{2}}$ .*

So by the Lemma A.4, we know that the function  $[a \times \text{Ext}_{IP}^{n,q}(\cdot, \cdot)]_1$  is a  $(qn, k, 2^{1-\frac{2k-qn}{2}})$  one-bit-output two-source extractor for any  $k \leq qn$ .

Finally, Lemma III.5 follows from the Lemma A.5 immediately by letting  $k = \left(\frac{1}{4} + \frac{\delta}{2}\right)qn + \frac{1}{2}$  and  $\epsilon = 2^{\frac{1}{2} - (\frac{\delta}{2} - \frac{1}{4})qn}$ .

**Lemma A.5** (Corollary 27 of [14] for  $m = 1$ ). *Let  $f$  be a  $(t, k, \epsilon)$  one-bit-output two-source extractor. Then it is a  $(t, k + \log \epsilon^{-1}, \sqrt{3\epsilon/2})$ -min-entropy quantum-proof two-source extractor. That is, for any random variables  $X, Y$  on  $\{0,1\}^t$  and any product quantum side information  $\mathcal{E}$  where  $H_{\min}(X|\mathcal{E}) \geq k + \log \epsilon^{-1}$  and  $H_{\min}(Y|\mathcal{E}) \geq k + \log \epsilon^{-1}$ , we have  $\frac{1}{2}\|\rho_{f(X,Y)\mathcal{E}} - \rho_{U_1} \otimes \rho_{\mathcal{E}}\| \leq \sqrt{3\epsilon/2}$ .*