Complete Classification
of Generalized Santha-Vazirani Sources

Salman Beigi[*]   Andrej Bogdanov[†]   Omid Etesami[‡]   Siyao Guo[§]

Abstract

Let $\mathcal{F}$ be a finite alphabet and $\mathcal{D}$ be a finite set of distributions over $\mathcal{F}$. A Generalized Santha-Vazirani (GSV) source of type $(\mathcal{F}, \mathcal{D})$, introduced by Beigi, Etesami and Gohari (ICALP 2015, SICOMP 2017), is a random sequence $(F_1, \ldots, F_n)$ in $\mathcal{F}^n$, where F_i is a sample from some distribution $d \in \mathcal{D}$ whose choice may depend on $F_1, \ldots, F_{i-1}$.

We show that all GSV source types $(\mathcal{F}, \mathcal{D})$ fall into one of three categories: (1) non-extractable; (2) extractable with error $n^{-\Theta(1)}$; (3) extractable with error $2^{-\Omega(n)}$. This rules out other error rates like $1/\log n$ or $2^{-\sqrt{n}}$.

We provide essentially randomness-optimal extraction algorithms for extractable sources. Our algorithm for category (2) sources extracts with error ε from $n = \mathrm{poly}(1/\varepsilon)$ samples in time linear in n. Our algorithm for category (3) sources extracts m bits with error ε from $n = O(m + \log 1/\varepsilon)$ samples in time $\min\{O(nm2^m), n^{O(|\mathcal{F}|)}\}$.

We also give algorithms for classifying a GSV source type $(\mathcal{F}, \mathcal{D})$: Membership in category (1) can be decided in NP, while membership in category (3) is polynomial-time decidable.

1 Introduction

Randomness extractors turn a weak source of randomness into almost uniform independent random bits. One of the first classes of distributions that were considered in the context of randomness extraction are Santha-Vazirani (SV) sources [SV86], also called unpredictable-bit sources. An SV source is a sequence of random bits such that every bit in the sequence has entropy bounded away from zero, even when conditioned on any possible sequence of previous bits. As already pointed out in [SV86], deterministic (seedless) extraction of even a single almost unbiased bit from SV sources is impossible, although these sources have entropy that grows linearly with their length.[1]

---

[*]Institute for Research in Fundamental Sciences, Tehran.

[†]Department of Computer Science and Engineering and Institute for Theoretical Science and Communications, Chinese University of Hong Kong. Supported by HK RGC GRF grants CUHK14208215 and CUHK 14238716.

[‡]Institute for Research in Fundamental Sciences, Tehran.

[§]Northeastern University. Part of the work are done while Siyao Guo was a research fellow at the Simons Institute for the Theory of Computing, UC Berkeley.

[1]With respect to seeded extraction, a constant seed length is sufficient for all SV sources [Vad12].

In this work we consider deterministic extraction for a natural generalization of Santha-Vazirani sources which was introduced by Beigi, Etesami, and Gohari [BEG15, BEG17]. A *generalized Santha-Vazirani (GSV) source* is specified by a pair $(\mathcal{F}, \mathcal{D})$, where $\mathcal{F}$ is a finite set of *faces* and $\mathcal{D}$ is a finite set of *dice*, each of which is a probability distribution on $\mathcal{F}$. (We will assume that each face is assigned positive probability by at least one die.) A distribution $(F_1, \ldots, F_n)$, where the F_is are $\mathcal{F}$-valued correlated random variables, is admissible by the source if it is generated by the following type of *strategy*: For each $1 \le i \le n$, a die $d \in \mathcal{D}$ is chosen as a function of $F_1, \ldots, F_{i-1}$ and F_i is sampled according to the distribution d.

The case $|\mathcal{D}| = |\mathcal{F}| = 2$ recovers the definition of SV sources. In this instance, the dice are two-sided coins, one biased towards heads and the other one towards tails.

We call a GSV source $(\mathcal{F}, \mathcal{D})$ *extractable* with error ε from n samples if there exists a function $\mathrm{Ext} \colon \mathcal{F}^n \to \{-1, 1\}$ such that for every distribution $(F_1, \ldots, F_n)$ in the source, $|\mathbb{E}[\mathrm{Ext}(F_1, \ldots, F_n)]| \le \varepsilon$. We call a source *extractable* if for every error $\varepsilon > 0$ there exists a sample size n for which the source is extractable with these parameters.

Beigi, Etesami and Gohari [BEG17] showed that randomness extraction from a GSV source is possible assuming the following condition:

Definition 1. *A GSV source $(\mathcal{F}, \mathcal{D})$ satisfies the* Nonzero Kernel Positive Variance (NK$^+$) *condition if there exists a function $\psi \colon \mathcal{F} \to [-1, 1]$ such that $\mathbb{E}_d[\psi(F)] = 0$ and $\mathrm{Var}_d[\psi(F)] > 0$ for every die $d \in \mathcal{D}$.*

Here, $\mathbb{E}_d$ and Var_d denote expectation and variance with respect to the distribution of die d. On the other hand, they showed that extractability from such sources necessitates the following *Nonzero Kernel* (NK) condition:

> There exists a nonzero $\psi \colon \mathcal{F} \to [-1, 1]$ such that $\mathbb{E}_d[\psi(F)] = 0$ for every die $d \in \mathcal{D}$.

In particular, when all faces of all dice have positive probability (an assumption called "nondegeneracy" in [BEG17]), the (NK$^+$) and (NK) conditions coincide, providing a characterization of extractability for this class of sources. Their extractor requires $\Theta(1/\varepsilon^3)$ samples to achieve error ε.

There are, however, simple examples of GSV sources ((E1) and (E2) below) that satisfy (NK) but not (NK$^+$). The work [BEG17] does not address the extractability of such sources.

In the setting of GSV sources, the existence of extractors does not appear to easily follow from counting arguments, as is the case of other types of sources for which extraction is known to be possible in principle and the focus is on efficient constructions, such as affine sources [Bou07, Gab11], polynomial sources [DGW09, Dvi12] and independent blocks [Bou05, CZ16].

Our Contributions

Our first contribution is a complete characterization of extractability from GSV sources. To motivate our result, we first observe that the (NK) condition is, in general, insufficient

for extractability. Consider, for instance the two-diced, three-faced GSV source described by the distributions $d_1 = (0, 0, 1)$ and $d_2 = (\frac{1}{2}, \frac{1}{2}, 0)$. This source satisfies (NK) with the witness $\psi = (-1, 1, 0)$, but is clearly not extractable as the distribution in which d_1 is repeatedly tossed contains no entropy.

A slightly more interesting example is provided by the four-diced, three-faced GSV source

$$d_1 = (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0) \quad d_2 = (0, 0, \tfrac{1}{3}, \tfrac{2}{3}) \quad d_3 = (0, 0, \tfrac{2}{3}, \tfrac{1}{3}). \tag{E1}$$

This source also satisfies the (NK) condition (with $\psi = (-1, 1, 0, 0)$). However, it is not extractable because it contains a "hidden" SV source (over two faces): If die d_1 is tossed away and the first two faces are removed, dice d_2 and d_3 now fail the (NK) condition.

These two examples suggest the following method for coming up with non-extractable GSV sources: Start with any source that fails (NK), extend the dice with more faces of zero probability, and add any number of dice that assign positive probability to the new faces. To describe such sources, we introduce the following natural strengthening of (NK):

Definition 2. *A GSV source $(\mathcal{F}, \mathcal{D})$ satisfies the* Hereditary Nonzero Kernel (HNK) *condition if for all subsets $\mathcal{D}' \subseteq \mathcal{D}$ there exists a nonzero witness $\psi : \mathcal{F}' \to [-1, 1]$ such that $\mathbb{E}_d[\psi(F)] = 0$ for all $d \in \mathcal{D}'$, where $\mathcal{F}'$ is the set of faces to which at least one die in $\mathcal{D}'$ assigns nonzero probability.*

Clearly (HNK) is a necessary condition for extractability, because if $(\mathcal{F}, \mathcal{D})$ fails (HNK) then $(\mathcal{F}', \mathcal{D}')$ fails (NK). Our first theorem shows that (HNK) is also sufficient. Moreover, it gives a universal upper bound on the number of samples:

Theorem 1. *The following conditions are equivalent for a GSV source $(\mathcal{F}, \mathcal{D})$:*

1. *$(\mathcal{F}, \mathcal{D})$ satisfies HNK.*

2. *$(\mathcal{F}, \mathcal{D})$ is extractable.*

3. *For every ε, $(\mathcal{F}, \mathcal{D})$ is extractable with error ε from $n = \mathrm{poly}(1/\varepsilon)$ samples in time linear in n.*

In the course of proving Theorem 1 we introduce the analytic *Mean Variance Ratio (MVR)* condition that turns out to be equivalent to HNK (Proposition 3). We show that a quantitative variant of the MVR condition determines the best-possible quality of extraction, up to a quadratic gap, even for GSV sources that are not extractable to within arbitrary small error (Propositions 1 and 2).

It is natural to ask if $\mathrm{poly}(1/\varepsilon)$ samples are in general necessary for the extractor in part 3 of Theorem 1. Our second result shows not only that this is the case, but completely characterizes GSV sources that are extractable in a randomness-efficient manner.

Theorem 2. *The following conditions are equivalent for a GSV source $(\mathcal{F}, \mathcal{D})$:*

1. *$(\mathcal{F}, \mathcal{D})$ satisfies NK^+.*

2. *For every ε, $(\mathcal{F}, \mathcal{D})$ is extractable with error ε from $o(1/\varepsilon^2)$ samples.*

3

3. *For every ε and m, $(\mathcal{F}, \mathcal{D})$ is extractable with error[2] ε and output length m from $n = O(\log(1/\varepsilon) + m)$ samples in time $\min\{O(nm2^m), n^{O(|\mathcal{F}|)}\}$.*

The sample complexity of the extractor in part 3 of Theorem 2 is optimal up to the leading constant: $\Omega(m)$ samples are necessary by entropy considerations, and $\Omega(1/\varepsilon)$ samples are necessary for non-trivial sources[3] by granularity considerations.

Condition NK^+ is strictly stronger than condition HNK. For example, the source

$$d_1 = (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0), \quad d_2 = (\tfrac{1}{4}, \tfrac{1}{12}, \tfrac{1}{3}, \tfrac{1}{3}), \quad d_3 = (\tfrac{1}{12}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{1}{3}). \tag{E2}$$

satisfies HNK but not NK^+.

Taken together, Theorems 1 and 2 completely classify non-trivial GSV sources into three categories: (1) non-extractable, (2) extractable with error $n^{-\Theta(1)}$, and (3) extractable with error $2^{-\Omega(n)}$, where n is the number of samples. This rules out the existence of GSV sources of other error rates like $1/\log n$ or $2^{-\sqrt{n}}$.

Moreover, sources can be classified algorithmically: Condition HNK can be decided by a coNP algorithm, while NK^+ is polynomial-time decidable (see Proposition 6).

Figure 1 indicates the relations between the different conditions for extractability of GSV sources uncovered in this work.
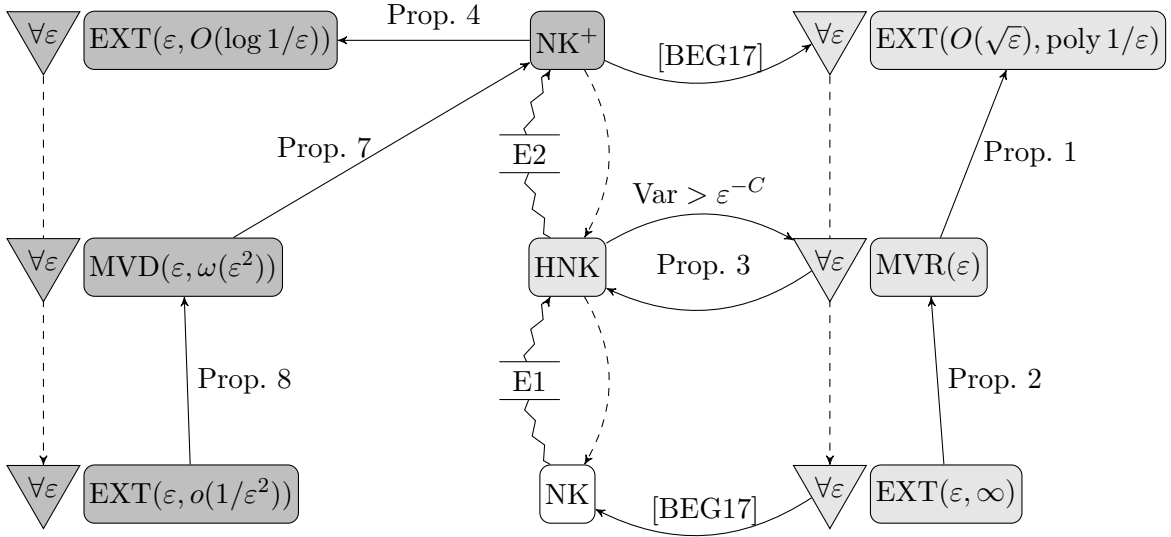

Figure 1: A map of our results. Straight arrows are implications (the dashed ones are immediate) and wiggly arrows are separations. $\mathrm{EXT}(\varepsilon, n)$ postulates extractability with error ε from n samples. Lightly and darkly shaded boxes represent equivalent conditions for extractability and randomness-efficient extractability, respectively.

---

[2]The error of an extractor that outputs multiple bits is the statistical (total variation) distance between its output distribution and the uniform distribution.

[3]The exception consists of one-die GSV sources that admit an event of probability exactly half, for which errorless extraction is possible.

Proof Techniques

Our proofs rely on a combination of probabilistic, algorithmic, and analytic methods.

Feasibility of extraction   The extractor of [BEG17] outputs the sign of $Z_T = \psi(F_1) + \cdots + \psi(F_T)$ at the earliest time T when $|Z_T|$ exceeds some pre-specified threshold M. Here, ψ is the witness for condition (NK$^+$), which ensures that $\mathbb{E}[\psi(F)]$ is always zero and $\text{Var}[\psi(F)]$ is always positive. Therefore (Z_t) is a martingale with growing variance, and the analysis of [BEG17] shows that the process terminates by time $n = O(1/\varepsilon^3)$ except with probability $\varepsilon/2$ when M is chosen as $\Theta(1/\varepsilon)$. Moreover, Z_T must take value in the range $(-(M+1), -M] \cup [M, M+1)$, so by the optional stopping time theorem, the bias of Z_T is $\varepsilon/2$ when $M = \Theta(1/\varepsilon)$.

In case only the weaker (HNK) condition holds, $\text{Var}[\psi(F)]$ could be zero for some dice and the value of Z_t may remain constant throughout the process. On the other hand, (HNK) provides not one but many witnesses ψ, one for every subset of the dice. Proposition 3 shows how all these witnesses can be combined into a single $\phi \colon \mathcal{F} \to [-1, 1]$ that has positive variance with respect to all the dice, but may have nonzero expectation. By a careful implementation of this strategy, it is ensured that the ratio $|\mathbb{E}_d[\phi(F)]|/\text{Var}_d[\phi(F)]$ can be made smaller than any pre-specified $\varepsilon > 0$. This is our Mean Variance Ratio (MVR) condition. Moreover, $\text{Var}_d[\phi(F)]$ can be lower bounded by ε^C for some constant C that depends only on the GSV source.

To prove Theorem 1 we apply the extractor of [BEG17] to the function ϕ. As ϕ may be biased with respect to some dice, (Z_t) may no longer be a martingale, rendering the optional stopping time theorem inapplicable. In Proposition 1 we demonstrate that the conclusion of the [BEG17] analysis still applies in our context. Intuitively, the (MVR) condition should imply that the variance of Z_t grows, and does so at a faster rate than the magnitude of its expectation. Therefore the stopping time should still be finite, and the component of extraction error incurred by $|\mathbb{E}[Z_T]|$ should be small. Owing to dependencies between the various steps, a rigorous implementation of these ideas requires substantial care.

Quality and quantity of extracted bits   For GSV sources that satisfy (NK$^+$) the extractor of [BEG17] inherently requires $\Omega(1/\varepsilon)$ samples: On the one hand, to ensure termination with high probability the boundary threshold M can be at most n, but on the other hand Z_T may fall anywhere in the range $(-(M+1), -M] \cup [M, M+1)$, thereby incurring an error of $\varepsilon = \Omega(1/M)$.[4] To improve the sample complexity, our bit extractor in Theorem 2 applies the update rule

$$Z_{t+1} = Z_t + \frac{\psi(F_t)}{2} \cdot (1 - |Z_t|)$$

and outputs the sign of Z_n for $n = O(\log 1/\varepsilon)$. Under (NK$^+$) the sequence (Z_t) is still a martingale, but now the range of Z_t is restricted to the open interval $(-1, 1)$. On average,

---

[4]A tempting alternative is for the "extractor" to simply output the sign of Z_n after looking at some predetermined number of samples. However, this "extractor" may incur error $\Omega(1)$ for almost any GSV source.

the deviation of the step size $Z_{t+1} - Z_t$ conditioned on Z_t is smaller the closer Z_t is to one of the boundary points $\{-1, 1\}$. We show that the logarithm of $1/(1 - |Z_t|)$ grows by a constant on average in every step and apply Azuma's inequality to conclude that Z_n is within $2^{-\Omega(n)}$ of 1 or -1 with probability $1 - 2^{-\Omega(n)}$. This ensures the bias of the output is inverse exponential in the number of samples.

To extract multiple bits, the state $\mathbf{Z}_t$ of the above process is extended to encode a probability distribution over $\{0, 1\}^m$. Initially $\mathbf{Z}_0$ is the uniform distribution. The distance measure $1 - |Z_t|$ is replaced by a carefully chosen quantity $\mathbf{D}_t \in \mathbb{R}^{2^m}$ which ensures that $\mathbf{Z}_t$ is a probability distribution that rapidly concentrates on a single entry in $\{0, 1\}^m$, which is the output of the extractor. Since $(\mathbf{Z}_t)$ is a multi-dimensional martingale, the output must be statistically close to uniform.

Lower bounds   Beigi, Etesami, and Gohari [BEG17] proved that if a source fails the (NK) condition, namely if for all ψ there exists a die d for which $|\mathbb{E}_d[\psi(F)]|/\mathrm{Var}_d[\psi(F)] = \Omega(1)$, then it is not extractable. In Proposition 2 we prove a quantitatively precise refinement of this statement: If $|\mathbb{E}_d[\psi(F)]|/\mathrm{Var}_d[\psi(F)] \geq \varepsilon$, then the extraction error must be at least $\Omega(\varepsilon)$. We conclude that extractability implies the (MVR) condition, which together with a compactness argument (see Proposition 3) gives (HNK), proving the "only if" direction of Theorem 1.

While this consequence was already established in [BEG17] by other, combinatorial methods, we obtain a further refinement that is used to prove the "only if" direction of Theorem 2. In Section 3.3 we introduce the mean-variance divergence (MVD) condition, which postulates that $|\mathbb{E}_d[\psi(F)]| < \varepsilon(\mathrm{Var}_d[\psi(F)] - \delta)$ for all dice. In Proposition 8 we show that if MVD fails then extraction with error ε requires $\Omega(1/\delta)$ samples. In Proposition 7 we use linear-algebraic duality to show that if (NK$^+$) fails then so does (MVD) with $\delta = O(\varepsilon^2)$, thereby completing the proof of Theorem 2.

2   A characterization of extractable GSV sources

In this Section we prove Theorem 1. The following analytic condition plays a central role in the proof:

Definition 3. *A GSV source* $(\mathcal{F}, \mathcal{D})$ *satisfies the* Mean-Variance Ratio condition *with parameter* $\varepsilon > 0$ *(MVR(ε)) if there exists a function* $\psi : \mathcal{F} \to [-1, 1]$ *such that for every die* $d \in \mathcal{D}$ *of a GSV source* $(\mathcal{F}, \mathcal{D})$,

$$\left| \mathbb{E}_d[\psi(F)] \right| < \epsilon \mathrm{Var}_d[\psi(F)]. \tag{MVR}$$

Proposition 1 in Section 2.1 shows that if a GSV source satisfies MVR(ε) then it is extractable with error $O(\sqrt{\varepsilon})$ from poly($1/\varepsilon$) samples. On the other hand, Proposition 2 in Section 2.2 shows that any GSV source that is extractable with error less than $\varepsilon/10$ (from any number of samples) satisfies MVR(ε). Thus the smallest ε for which MVR(ε) holds measures the best-possible quality of extraction of a GSV source to within a square.

In the case when MVR(ε) holds for all $\varepsilon > 0$, the source is extractable. Surprisingly, proposition 3 shows that $\forall \varepsilon \mathrm{MVR}(\varepsilon)$ implies HNK. HNK, in turn, implies a slightly

stronger form of $\forall \varepsilon \mathrm{MVR}(\varepsilon)$, which together with Proposition 1 establishes the extractability of HNK sources from ε^{-C} samples, where C is a constant that depends only on the source.

2.1 Feasibility of extraction

Proposition 1. *If GSV source $(\mathcal{F}, \mathcal{D})$ satisfies $MVR(\varepsilon)$, then it is extractable from n samples with error at most $3\sqrt{\varepsilon} + 4/\varepsilon vn + O(\varepsilon)$, where v is the minimum of $\mathrm{Var}_d[\psi(F)]$ over all $d \in \mathcal{D}$.*

Proof of Proposition 1. Define random variables $X_1, \ldots, X_n$ and $Z_0, \ldots, Z_n$ by $Z_0 = 0$ and $Z_i = Z_{i-1} + X_i$ for $i > 0$, where $X_i = 0$ if $|Z_i| \geq M$, $X_i = \psi(F_i)$ if $|Z_i| < M$, F_i $(1 \leq i \leq n)$ is the i-th output of the GSV source sequence, and $M = 1/\sqrt{\varepsilon}$. Under this definition, Z_n is uniformly bounded by $M + 1$. The extractor outputs the sign of Z_n.

To prove that the sign of Z_n has small bias, we begin by lower bounding $\mathrm{Var}[Z_n]$. We will use this lower bound to argue both that the expectation of Z_n in absolute value and that the probability that Z_n remains in the range $(-M, M)$ are small. These two facts will allow us to conclude that the sign of Z_n is almost unbiased.

Claim 1. $\mathrm{Var}[Z_n] \geq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}[\mathrm{Var}[X_i | Z_{i-1}]]$.

Proof. By the law of total variance we have

$$\mathrm{Var}[Z_i] = \mathrm{Var}\big[\mathbb{E}[Z_i | Z_{i-1}]\big] + \mathbb{E}\big[\mathrm{Var}[Z_i | Z_{i-1}]\big].$$

Furthermore,

$$\begin{aligned}
\mathrm{Var}\big[\mathbb{E}[Z_i | Z_{i-1}]\big] &= \mathrm{Var}\big[Z_{i-1} + \mathbb{E}[X_i | Z_{i-1}]\big] \\
&= \mathrm{Var}[Z_{i-1}] + \mathrm{Var}\big[\mathbb{E}[X_i | Z_{i-1}]\big] + 2\mathrm{Cov}\big(Z_{i-1}, \mathbb{E}[X_i | Z_{i-1}]\big).
\end{aligned}$$

Now we compute

$$\begin{aligned}
\mathrm{Cov}\big(Z_{i-1}, \mathbb{E}[X_i | Z_{i-1}]\big) &= \mathbb{E}\big[(Z_{i-1} - \mathbb{E}[Z_{i-1}]) \cdot \mathbb{E}[X_i | Z_{i-1}]\big] \\
&\geq -\mathbb{E}\big[|Z_{i-1} - \mathbb{E}[Z_{i-1}]| \cdot |\mathbb{E}[X_i | Z_{i-1}]|\big] \\
&\geq -\mathbb{E}\big[(2M + 2) \cdot \epsilon \mathrm{Var}[X_i | Z_{i-1}]\big] \\
&\geq -\mathbb{E}\big[\mathrm{Var}[X_i | Z_{i-1}]/4\big],
\end{aligned}$$

since $|\mathbb{E}[X_i | Z_{i-1}]| \leq \epsilon \mathrm{Var}[X_i | Z_{i-1}]$ in both cases $|Z_{i-1}| \geq M$ and $|Z_{i-1}| < M$. Combining the above three equations and noting that $\mathrm{Var}[\mathbb{E}[X_i | Z_{i-1}]] \geq 0$, we get

$$\mathrm{Var}[Z_i] \geq \mathrm{Var}[Z_{i-1}] - \frac{1}{2}\mathbb{E}\big[\mathrm{Var}[X_i | Z_{i-1}]\big] + \mathbb{E}[\mathrm{Var}[Z_i | Z_{i-1}]].$$

We also have $\mathrm{Var}[Z_i | Z_{i-1}] = \mathrm{Var}[Z_{i-1} + X_i | Z_{i-1}] = \mathrm{Var}[X_i | Z_{i-1}]$. Hence

$$\mathrm{Var}[Z_i] \geq \mathrm{Var}[Z_{i-1}] + \frac{1}{2}\mathbb{E}\big[\mathrm{Var}[X_i | Z_{i-1}]\big].$$

The claim now follows by induction on n. □

To upper bound $|\mathbb{E}[Z_n]|$, we can write

$$|\mathbb{E}[Z_n]| \leq \sum_{i=1}^{n} \mathbb{E}\big[|\mathbb{E}[X_i|Z_{i-1}]|\big] \leq \sum_{i=1}^{n} \mathbb{E}\big[\varepsilon \mathrm{Var}[X_i|Z_{i-1}]\big] \leq 2\varepsilon \mathrm{Var}[Z_n] \leq 2\varepsilon(M+1)^2. \quad (1)$$

The first inequality is the triangle inequality. The second inequality follows from assumption (MVR) when $|Z_i| < M$, and the fact that $\mathbb{E}[X_i|Z_{i-1}]$ is zero otherwise. The third inequality follows from Claim 1.

Let p be the probability that $|Z_i| < M$ for all $1 \leq i \leq n$, i.e., $p = \Pr\big[|Z_i| < M, 1 \leq i \leq n\big]$. Then

$$\mathrm{Var}[Z_n] \geq \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\big[\mathrm{Var}[X_i|Z_{i-1}]\big]$$

$$= \frac{1}{2}\sum_{i=1}^{n} \Pr\big[|Z_{i-1}| < M\big] \cdot \mathbb{E}\big[\mathrm{Var}[X_i|Z_{i-1}]\,\big|\,|Z_{i-1}| < M\big]$$

$$\geq \frac{1}{2}pnv,$$

where the first inequality follows from Claim 1, the second equality follows from the law of conditional expectations, and the third inequality follows because the event $|Z_{i-1}| < M$ contains, in particular the event $|Z_i| < M$ for all i of probability p, and conditioned on $|Z_i|$ the conditional variance of X_i is the variance of $\psi(F_i)$. Therefore,

$$p \leq \frac{2\mathrm{Var}[Z_n]}{nv} \leq \frac{2(M+1)^2}{nv}. \quad (2)$$

The bias of the extracted bit is at most

$$|\Pr[Z_n \geq 0] - \Pr[Z_n < 0]| \leq |p_+ - p_-| + p, \quad (3)$$

where $p_+ = \Pr[Z_n \geq M]$ and $p_- = \Pr[Z_n \leq M]$. To upper bound $|p_+ - p_-|$, we apply the law of conditional expectations to $\mathbb{E}[Z_n]$ to obtain that

$$\big|\mathbb{E}[Z_n] - (p_+ - p_-)M\big| = \big|p_+\mathbb{E}[Z_n - M|Z_n \geq M] + p_-\mathbb{E}[Z_n + M|Z_n \leq -M] + p\mathbb{E}[Z_n||Z_n| < M]\big|$$

$$\leq p_+ + p_- + pM$$

$$\leq pM + 1.$$

By the triangle inequality and (1),

$$|p_+ - p_-| \leq \frac{1}{M} \cdot \big(\mathbb{E}[Z_n] + pM + 1\big) \leq 2\varepsilon\frac{(M+1)^2}{M} + p + \frac{1}{M}.$$

By (3), the bias of the extractor is at most $2\varepsilon(M+3) + 2p + 1/M$. Assuming, without loss of generality, that $\varepsilon < 1$ and using (2) we obtain the desired bound for $M = 1/\sqrt{\varepsilon}$. $\quad\square$

2.2 Impossiblilty of extraction

Proposition 2. *Let ε be a sufficiently small constant. Assume* $\mathrm{MVR}(\varepsilon)$ *fails for a source* $(\mathcal{F}, \mathcal{D})$. *Then* $(\mathcal{F}, \mathcal{D})$ *is not extractable with error better than* $\varepsilon/10$ *from any number of samples.*

Proof of Proposition 2. Assuming $\mathrm{MVR}(\varepsilon)$ fails we will prove the following claim:

Claim 2. *For every n, every extractor* $\mathrm{Ext}\colon \mathcal{F}^n \to \{0,1\}$, *and every $0 \le \alpha \le 1$, if* $\mathbb{E}_{A_-}[\mathrm{Ext}] \ge \alpha$ *for every strategy A_-, then there exists a strategy A_+ for which* $\mathbb{E}_{A_+}[\mathrm{Ext}] \ge \alpha + (\varepsilon/(1+\varepsilon)) \cdot \alpha(1-\alpha)$.

To derive the theorem from the claim, assume that $\mathbb{E}[\mathrm{Ext}] \ge \alpha = 1/2 - \varepsilon/10$ with respect to every strategy. By Claim 2 there must then exist a strategy for which

$$\mathbb{E}[\mathrm{Ext}] \ge \frac{1}{2} - \frac{\varepsilon}{10} + \frac{\varepsilon}{1+\varepsilon} \cdot \frac{1 - \varepsilon^2/100}{4}$$

which is at least $1/2 + \varepsilon/10$. $\qquad\square$

Proof of Claim 2. We prove the claim by induction on n. When $n = 0$ the claim holds by checking the cases $\mathrm{Ext} = 0$ and $\mathrm{Ext} = 1$. We now assume it holds for $n - 1$ and prove it for n. Let d_- be the choice of the first die that minimizes $\mathbb{E}_{A_-}[\mathrm{Ext}]$. Then

$$\alpha \le \mathbb{E}_{d_-}[\alpha(F)],$$

where $\alpha(f)$ is the advantage of Ext conditioned on the first outcome being f.

We now describe the strategy A_+. By $\overline{\mathrm{MVR}(\varepsilon)}$ applied to the function $\psi(f) = \alpha(f) - \alpha$, there exists a die d_+ such that

$$\mathbb{E}_{d_+}[\alpha(F) - \alpha] \ge \varepsilon \mathrm{Var}_{d_+}[\alpha(F)]. \tag{4}$$

The adversary A_+ tosses this die first. She then plays the strategy that maximizes $\mathbb{E}_{A_+}[\mathrm{Ext}]$ conditioned on the outcome of the first die. By our inductive assumption, the conditional advantage of A_+ when the first outcome is f must be at least $\alpha(f) + (\varepsilon/(1+\varepsilon)) \cdot \alpha(f)(1-\alpha(f))$ so that

$$\mathbb{E}_{A_+}[\mathrm{Ext}] \ge \mathbb{E}_{d_+}\left[\alpha(F) + \frac{\varepsilon}{1+\varepsilon} \cdot \alpha(F)(1 - \alpha(F))\right].$$

We can write

$$\mathbb{E}_{d_+}\left[\alpha(F) + \frac{\varepsilon}{1+\varepsilon} \cdot \alpha(F)(1 - \alpha(F))\right] - \left(\alpha + \frac{\varepsilon}{1+\varepsilon} \cdot \alpha(1-\alpha)\right)$$
$$= \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)\mathbb{E}_{d_+}[\alpha(F) - \alpha] - \frac{\varepsilon}{1+\varepsilon}\mathrm{Var}_{d_+}[\alpha(F)] - \frac{\varepsilon}{1+\varepsilon}\left(\mathbb{E}_{d_+}[\alpha(F)^2] - \alpha^2\right). \tag{5}$$

We can upper bound the last term by

$$\left(\mathbb{E}_{d_+}[\alpha(F)^2] - \alpha^2\right) = \mathbb{E}_{d_+}[\alpha(F) + \alpha] \cdot \mathbb{E}_{d_+}[\alpha(F) - \alpha] \le 2\mathbb{E}_{d_+}[\alpha(F) - \alpha]$$

9

since all the αs are between zero and one, and the second term is non-negative because by the minimality of d_-, $\mathbb{E}_{d_+}[\alpha(F)] \geq \mathbb{E}_{d_-}[\alpha(F)] \geq \alpha$. We can therefore lower bound the left hand size of (5) by

$$\left(1 - \frac{\varepsilon}{1+\varepsilon}\right)\mathbb{E}_{d_+}[\alpha(F) - \alpha] - \frac{\varepsilon}{1+\varepsilon}\mathrm{Var}_{d_+}[\alpha(F)]$$

which, by (4), must be non-negative. It follows that the advantage of A_+ is at least $\alpha + (\varepsilon/(1+\varepsilon))\alpha(1-\alpha)$, concluding the inductive step. $\qquad\square$

2.3 Proof of Theorem 1

Proposition 3. *The following conditions are equivalent for a GSV source $(\mathcal{F}, \mathcal{D})$:*

1. *For all $\varepsilon > 0$, $(\mathcal{F}, \mathcal{D})$ satisfies $\mathrm{MVR}(\varepsilon)$: There exists a $\psi : \mathcal{F} \to [-1, 1]$ such that for all dice d, $|\mathbb{E}_d[\psi(F)]| < \varepsilon\mathrm{Var}_d[\psi(F)]$.*

2. *There exists a constant C such that for sufficiently small $\varepsilon > 0$, there exists a $\psi : \mathcal{F} \to [-1, 1]$ such that for all dice d, $|\mathbb{E}_d[\psi(F)]| < \varepsilon\mathrm{Var}_d[\psi(F)]$ and $\mathrm{Var}_d[\psi(F)] \geq \varepsilon^C$.*

3. *$(\mathcal{F}, \mathcal{D})$ satisfies HNK.*

Proof. We will show that 1 implies 3 and 3 implies 2. This will establish equivalence as 2 is a stronger condition than 1.

1 implies 3: Assume that $(\mathcal{F}, \mathcal{D})$ satisfies $\mathrm{MVR}(\varepsilon)$. This condition is hereditary, namely if it holds for $(\mathcal{F}, \mathcal{D})$ then it holds for all $(\mathcal{F}', \mathcal{D}')$ in the assumption of HNK. So in proving 3, we may and will assume, without loss of generality, that $(\mathcal{F}', \mathcal{D}') = (\mathcal{F}, \mathcal{D})$. We will moreover assume (by scaling and flipping sign if necessary) that ψ attains the value 1.

Now consider an infinite decreasing sequence (ε_k) that converges to zero. By assumption, for every k there exists a ψ_k such that $|\mathbb{E}_d[\psi_k(F)]| < \varepsilon_k\mathrm{Var}_d[\psi_k(F)]$. By the pigeonhole principle there must exist a face f for which the set of indices $K = \{k : \psi_k(f) = 1\}$ is infinite. By compactness of $[-1, 1]^{\mathcal{F}}$ there must exist an infinite subset $K' \subseteq K$ for which the subsequence ψ_k over $k \in K'$ converges to a limit ψ. Then ψ is nonzero as $\psi(f)$ must equal one. On the other hand, for every $\varepsilon > 0$ there exists a sufficiently large $k \in K'$ such that for every die d,

$$|\mathbb{E}[\psi_d(F)]| \leq |\mathbb{E}_d[\psi_k]| + \varepsilon \leq \varepsilon\mathrm{Var}_d[\psi_k(F)] + \varepsilon,$$

so $\mathbb{E}_d[\psi_d(F)]$ must equal zero for every d.

3 implies 2: The proof is by strong induction on the number of dice $|\mathcal{D}|$ with $C = 3 \cdot 2^{|\mathcal{D}|} - 3$. In the base case $|\mathcal{D}| = 1$, all faces must be assigned nonzero probability by the unique die d. Take any witness ψ for HNK. Then $\mathbb{E}_d[\psi(F)] = 0$, but ψ must take nonzero value on at least one of the faces, so $\mathrm{Var}_d[\psi(F)] > 0$. Condition 2 is then satisfied for sufficiently small $\varepsilon > 0$.

For the inductive step, take any ψ that is a witness for HNK with respect to the whole source $(\mathcal{F}, \mathcal{D})$. Let $\mathcal{D}'$ be the subset of dice d such that $\mathrm{Var}_d[\psi(F)] = 0$ and v be the minimum of $\mathrm{Var}_d[\psi(F)]$ over $d \notin \mathcal{D}'$. Then $\mathcal{D}'$ is a proper subset of $\mathcal{D}$ (otherwise, there is

a face that is assigned no probability by any die). If $\mathcal{D}'$ is empty, condition 2 follows by the same argument as in the base case. If not, then by the inductive hypothesis we can choose $\psi' \colon \mathcal{F}' \to [-1, 1]$ such that

$$|\mathbb{E}_d[\psi'(F)]| < (v\varepsilon^2/8) \cdot \mathrm{Var}_d[\psi(F)] \quad \text{and} \quad \mathrm{Var}_d[\psi(F)] \geq (v\varepsilon^2/8)^{3 \cdot 2^{|\mathcal{D}'|} - 3}. \qquad (6)$$

We will show that the function $\phi = \psi + (v\varepsilon/8) \cdot \psi'$ satisfies the conclusion of condition 2. Here, ψ' is naturally extended as a function on $\mathcal{F}$ by assigning zero on all inputs in $\mathcal{F} \setminus \mathcal{F}'$. The proof is by cases.

If $d \in \mathcal{D}'$, then $\mathbb{E}_d[\phi(F)] = (v\varepsilon/8)\mathbb{E}_d[\psi'(F)]$, while $\mathrm{Var}_d[\phi(F)] = (v\varepsilon/8)^2\mathrm{Var}_d[\psi'(F)]$. From these two equalities and (6) it follows that $\mathbb{E}_d[\phi(F)] < \varepsilon\mathrm{Var}_d[\phi(F)]$. On the other hand, $\mathrm{Var}_d[\phi(F)] \geq (v\varepsilon/8)^2 \cdot (v\varepsilon^2/8)^{3 \cdot 2^{|\mathcal{D}'|} - 3} \geq \varepsilon^{3 \cdot 2^{|\mathcal{D}|} - 3}$ for sufficiently small ε.

If $d \notin \mathcal{D}'$, then $|\mathbb{E}_d[\phi(F)]| \leq (v\varepsilon/8)|\mathbb{E}_d[\psi'(F)]| \leq v\varepsilon/8$, while

$$\begin{aligned}
\mathrm{Var}_d[\psi(F)] &\geq \mathrm{Var}_d[\psi'(F)] - 2|\mathrm{Cov}_d[\psi(F), (v\varepsilon/8) \cdot \psi'(F)]| \\
&= \mathrm{Var}_d[\psi'(F)] - \frac{v\varepsilon}{4} \cdot |\mathrm{Cov}_d[\psi(F), \psi'(F)]| \\
&\geq \mathrm{Var}_d[\psi'(F)] - \frac{v\varepsilon}{2} \\
&\geq \frac{v}{2},
\end{aligned}$$

where the last inequality follows from our definition of v. In particular, $\mathrm{Var}_d[\psi(F)] \geq \varepsilon^{3 \cdot 2^{|\mathcal{D}|} - 3}$ for sufficiently small ε. On the other hand, $|\mathbb{E}_d[\psi(F)]| \leq v\varepsilon/8 \leq (\varepsilon/4) \cdot \mathrm{Var}_d[\psi(F)]$, as desired. $\qquad \square$

Proof of Theorem 1. If $(\mathcal{F}, \mathcal{D})$ satisfies HNK, then it also satisfies condition 2 of Proposition 3. By Proposition 1, $(\mathcal{F}, \mathcal{D})$ is extractable with error $O(\sqrt{\varepsilon}) + n/\varepsilon^{C+1}$. The forward direction follows by setting $n = \varepsilon^{C+1.5}$.

For the reverse direction, if $(\mathcal{F}, \mathcal{D})$ fails to satisfy HNK, by Proposition 3, then it also fails to satisfy $\mathrm{MVR}(\varepsilon)$ for some $\varepsilon > 0$. So by Proposition 2 it is not extractable. $\qquad \square$

Alternatively, the reverse direction of Theorem 1 can be derived from Theorem 6 of [BEG17] because if $(\mathcal{F}, \mathcal{D})$ fails (NHK) then it contains some $(\mathcal{F}', \mathcal{D}')$ which fails (NK).

3   Randomness-efficient extraction

In this Section we prove Theorem 2. In Section 3.1, we begin with improving the quality of the extractor of [BEG17] for (NK$^+$) GSV sources to exponentially small error. Then in Section 3.2, we show how to improve the number of extracted bits and prove the implication $1 \to 3$ in Theorem 2.

In Section 3.3 we state and prove a necessary condition for the quality of extraction and use it to prove the remaining implication $2 \to 1$ in Theorem 2.

3.1 An optimal bit extractor

Proposition 4. *For every $\varepsilon > 0$, every GSV source that satisfies (NK^+) is extractable with error ε from $O(\log(1/\varepsilon)/v^2)$ samples where v is the minimum of $\mathrm{Var}_d[\psi(F)]$ over all $d \in \mathcal{D}$.*

Proof of Proposition 4. Define random variables $Z_0, \ldots, Z_n$ by $Z_0 = 0$ and $Z_{t+1} = Z_t + (\psi(F_t)/2) \cdot (1 - |Z_t|)$ where F_t $(1 \le t \le n)$ is the t-th output of the GSV source sequence. The extractor outputs the sign of Z_n.

Under (NK^+) the sequence (Z_t) is still a martingale so that the expectation of Z_n is 0. But now the range of Z_t is restricted to the open interval $(-1, 1)$. To prove that the sign of Z_n has small bias, we begin by showing on average, the logarithm of $1/D_t$ grows by a constant on average in every step where $D_t = 1 - |Z_t|$ is the distance between Z_t and its sign. Then we will use to argue the expectation of D_n is exponentially small. This fact together with $\mathbb{E}[Z_n] = 0$ allows us to conclude that the sign of Z_n is exponentially close to unbiased.

Claim 3. $\mathbb{E}[\ln(1/D_t) - \ln(1/D_{t-1}) \mid D_1, \ldots, D_{t-1}] \ge v/24$.

Proof. Observe that, $D_t = 1 - |Z_t| \le 1 - \mathrm{sign}(Z_{t-1}) \cdot Z_t$. By expanding Z_t to $Z_{t-1} + (\psi(F_t)/2) \cdot D_{t-1}$, and replacing $1 - \mathrm{sign}(Z_{t-1}) \cdot Z_{t-1}$ by D_{t-1}, it follows that

$$D_t \le (1 - \mathrm{sign}(Z_{t-1}) \cdot \frac{\psi(F_t)}{2}) \cdot D_{t-1}. \tag{7}$$

Because $|Z_t| \in (-1, 1)$ and $D_t > 0$, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\ln \frac{D_{t-1}}{D_t} \;\middle|\; D_1, \ldots, D_{t-1}\right] &\ge \mathbb{E}\left[-\ln\left(1 - \mathrm{sign}(Z_{t-1}) \cdot \frac{\psi(F_i)}{2}\right) \;\middle|\; D_1, \ldots, D_{t-1}\right] \\
&\ge \mathbb{E}\left[\mathrm{sign}(Z_{t-1}) \cdot \frac{\psi(F_i)}{2} + \frac{1}{6}\left(\mathrm{sign}(Z_{t-1}) \cdot \frac{\psi(F_i)}{2}\right)^2\right] \\
&= \mathbb{E}\left[\mathrm{sign}(Z_{t-1}) \cdot \frac{\psi(F_i)}{2}\right] + \frac{1}{6}\mathbb{E}\left[\left(\frac{\psi(F_i)}{2}\right)^2\right] \\
&\ge 0 + v/24.
\end{aligned}
$$

The second inequality follows from $-\ln(1-x) \ge x + x^2/6$ and that $\frac{\psi(C_i)}{2} \in (-1/2, 1/2)$, the third equality follows from the linearity of expectation, and the last inequality follows from $\mathbb{E}_d[\psi(F)] = 0$, for all $d \in \mathcal{D}$, and the definition of v. $\qquad\square$

Let us define variable $X_t = \ln(1/D_t) - (vt/24)$, for $t = 0, \ldots, n$. By Claim 3, $\mathbb{E}[X_t | D_0, \ldots, D_{t-1}] \ge X_{t-1}$ so that the sequence (X_t) forms a sub-martingale with respect to $D_0, \ldots, D_n$. Moreover, because by triangle inequality, $D_t = 1 - |z_t| \ge D_{t-1}(1 - |\frac{\psi(F_t)}{2}|) \ge D_{t-1}/2$ and by (7), $D_t \le D_{t-1}(1 + \frac{1}{2})$, $|X_t - X_{t-1}|$ is upper bounded by $c = \ln 2 + v/24$. By Azuma's inequality, for any $k \ge 0$

$$\Pr[X_n - X_0 \le -k] \le e^{-k^2/2nc^2}.$$

Then by plugging in $k = vn/12$, $X_0 = 0$ and $c = \ln 2 + v/24$, we obtain $\Pr[D_n \geq e^{-vn/12}] = 2^{-\Omega(v^2 n)}$. Therefore $\mathbb{E}[D_n] = 2^{-\Omega(v^2 n)}$. And we can conclude $|\mathbb{E}[\operatorname{sign}(Z_n)]| \leq \mathbb{E}[D_n] + |\mathbb{E}[Z_n]| = 2^{-\Omega(v^2 n)}$ by triangle inequality.

$\square$

3.2 Extracting more bits

We now explain how the extractor of the previous subsection can be modified to extract multiple bits. We will prove the following proposition, which is a more detailed restatement of the implication $3 \rightarrow 1$ in Theorem 2.

Proposition 5. *For every $\varepsilon > 0$ and m, every GSV source that satisfies (NK^+) is extractable with error ε and output length m from $O((\log(1/\varepsilon) + m)/v^2)$ samples where v is the minimum of $\operatorname{Var}_d[\psi(F)]$ over all $d \in \mathcal{D}$.*

Extraction procedure. Let $M = 2^m$. We describe how to output an almost uniform distribution over $[M]$ given (NK^+) GSV sources. Define random variables $\mathbf{Z}_0, \ldots, \mathbf{Z}_n$ over $\mathbb{R}^M$ by $\mathbf{Z}_0 = \frac{1}{M} \cdot \mathbf{1} \in \mathbb{R}^M$ be the vector all of whose coordinates are $1/M$ and

$$\mathbf{Z}_t = \mathbf{Z}_{t-1} + \frac{\psi(F_t)}{2} \cdot \mathbf{D}_t,$$

where F_t ($1 \leq t \leq n$) is the t-th output of the GSV source sequence and $\mathbf{D}_t \in \mathbb{R}^M$ is a vector defined below. The extractor outputs the index of the largest coordinate in $\mathbf{Z}_n$.

Let $\mathbf{s}_{i-1} : [M] \rightarrow [M]$ be the function which on input $j \in [M]$ outputs the index of the j-th smallest coordinate in $\mathbf{z}_{i-1}$ (so that $\mathbf{z}_{i-1}[\mathbf{s}_{i-1}(1)] \leq \mathbf{z}_{i-1}[\mathbf{s}_{i-1}(2)] \leq \cdots \leq \mathbf{z}_{i-1}[\mathbf{s}_{i-1}(M)]$); Then $\mathbf{d}_i$ is given by

$$\mathbf{d}_i[\mathbf{s}_{i-1}(j)] = (-1)^j \cdot \mathbf{z}_{i-1}[\mathbf{s}_{i-1}(j)], \qquad \text{if } 1 \leq j \leq M - 1,$$

and $\mathbf{d}_i[\mathbf{s}_{i-1}(M)] = -\sum_{j=1}^{M-1} \mathbf{d}_i[\mathbf{s}_{i-1}(j)]$.

As an example, consider $m = 1$. For each $\mathbf{z}_t$, if $\mathbf{z}_{t-1}[1] \leq \mathbf{z}_{t-1}[2]$, then $\mathbf{d}_t[1] = -\mathbf{z}_{t-1}[1]$ and $\mathbf{d}_t[2] = \mathbf{z}_{t-1}[1]$, otherwise $\mathbf{d}_t[2] = -\mathbf{z}_{t-1}[2]$ and $\mathbf{d}_t[1] = \mathbf{z}_{t-1}[2]$. The extractor outputs 2 if $\mathbf{z}_n[1] \leq \mathbf{z}_n[2]$ otherwise output 1. This procedure simulates the extractor in Proposition 4 by considering variable $Z_t = \mathbf{Z}_t[2] - \mathbf{Z}_t[1]$ instead and observing that $Z_t = Z_{t-1} + \frac{\psi(F)}{2} \cdot (1 - |Z_t|)$.

Analysis. Our proof relies on following two claims. The first claim says the values of $\mathbf{z}_t$ are never negative and add up to one, so at each time they represent a probability distribution over the M possibilities. The second claims says the total mass in this probability distribution gets concentrated in only one of the M possibilities after reading enough symbols from the GSV source.

Claim 4. *For any $0 \leq t \leq n$, $\mathbf{z}_t[1] + \cdots + \mathbf{z}_t[M] = 1$ and $\mathbf{z}_t[j] > 0$ for any $j \in [M]$*

Claim 5. *For a sufficiently large constant C, and $n \geq C \cdot (m + \log(1/\varepsilon))/v^2$, with probability at least $1 - \varepsilon/2$, $\mathbf{Z}_n[\mathbf{S}_n(M)] \geq 1 - \varepsilon/2$.*

We first assume those two claims and prove Proposition 5.

Proof of Proposition 5. For any $j \in [M]$, the sequence of random variables $\mathbf{Z}_0[j], \ldots, \mathbf{Z}_n[j]$ forms a martingale because $\mathbb{E}_d[\psi(F)] = 0$ for any $d \in \mathcal{D}$. So that $\mathbf{Z}_n[j] = \mathbf{Z}_0[j] = 1/M$. By Claim 4, $\mathbf{z}_n$ is a probability distribution. Consider a modification Ext' of our extractor which outputs j with probability $\mathbf{Z}_n[j]$ for any $j \in [M]$. Then the output of Ext' is uniformly distributed over $[M]$, because for any $j \in [M]$, Ext' outputs j with probability $\mathbb{E}[\mathbf{Z}_n[j]] = 1/M$.

Conditioned on $\mathbf{Z}_n$, the output of our extractor is different from Ext' with probability at most $1 - \mathbf{Z}_n[\mathbf{S}_n(M)]$. Thus the error of our extractor is at most $1 - \mathbb{E}[\mathbf{Z}_n[\mathbf{S}_n(M)]]$. By Claim 5 and $\mathbf{Z}_n \geq 0$, for $n = O((m + \log(1/\varepsilon^2)/v^2))$, $\mathbb{E}[\mathbf{Z}_n[\mathbf{S}_n(M)]] \geq (1 - \varepsilon/2)^2 \geq 1 - \varepsilon$. Proposition 5 follows. $\qquad\square$

Now we prove Claim 4 and 5.

Proof of Claim 4. We prove by induction on t. The base case holds by the definition of $\mathbf{z}_0$. Moreover, assuming that coordinates of $\mathbf{z}_{t-1}$ sum to one, we find that $\mathbf{z}_t[1] + \cdots + \mathbf{z}_t[M] = 1$ due to $\sum_{j=1}^{M} \mathbf{d}_t[j] = 0$. Also, $\mathbf{z}_t > 0$ follows from $\mathbf{z}_{t-1} > 0$, $|\psi| \leq 1$ and that $|\mathbf{d}_t[j]| \leq \mathbf{z}_{t-1}[j]$ for all $j \in [M]$. The latter inequality, when written in the form $|\mathbf{d}_t[\mathbf{s}_{t-1}(j)]| \leq \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(j)]$ is easy to verify: It is immediate for $1 \leq j \leq M - 1$ and for $j = M$ we have

$$\left|\mathbf{d}_t[\mathbf{s}_{i-1}(M)]\right| = \left|\sum_{j=1}^{M-1} (-1)^j \cdot \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(j)]\right| \leq \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(M-1)] \leq \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(M)],$$

where the middle inequality comes from the fact that $\mathbf{z}_{t-1}[\mathbf{s}_{t-1}(j)]$'s have been sorted and $\mathbf{z}_{t-1} > 0$ so that we can rewrite $\sum_{j=1}^{M-1} (-1)^{j+1} \cdot \mathbf{z}_{i-1}[\mathbf{s}_{i-1}(j)]$ and show $0 < \sum_{j=1}^{M-1} (-1)^{j+1} \cdot \mathbf{z}_{i-1}[\mathbf{s}_{i-1}(j)] \leq \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(M-1)]$ as follows.

$$\mathbf{z}_{t-1}[\mathbf{s}_{t-1}(1)] + \sum_{j=2}^{M/2} \left(\mathbf{z}_{t-1}[\mathbf{s}_{t-1}(2j-1)] - \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(2j-2)]\right) > 0,$$

$$\mathbf{z}_{t-1}[\mathbf{s}_{t-1}(M-1)] - \sum_{j=2}^{M/2} \left(\mathbf{z}_{t-1}[\mathbf{s}_{t-1}(2j-2)] - \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(2j-3)]\right) \leq \mathbf{z}_{t-1}[\mathbf{s}_{t-1}(M-1)].$$

$\qquad\square$

Proof of Claim 5. To prove $\Pr[\mathbf{Z}_n[\mathbf{S}_n(M)] \geq 1 - \varepsilon/2] \geq 1 - \varepsilon/2$, it is sufficient to show,

$$\Pr\left[\exists j_0 \in [M], \text{ s.t. } \forall j \neq j_0, \mathbf{Z}_n[j] \leq \frac{\varepsilon}{2M}\right] \geq 1 - \varepsilon/2. \tag{8}$$

To show this, for any $j \in [M]$, we define the sequence $\mathbf{X}_0[j], \ldots, \mathbf{X}_n[j]$ where $\mathbf{X}_t[j] = \ln(1/\mathbf{Z}_t[j])$, and the sequence $\mathbf{Y}_0[j], \ldots, \mathbf{Y}_n[j]$ where $\mathbf{Y}_0[j] = 0$ and for $t \geq 1$ $\mathbf{Y}_t[j] = \mathbf{Y}_{t-1}[j] + \mathbf{X}_t[j] - \mathbb{E}[\mathbf{X}_t[j] \mid \mathbf{X}_{t-1}[j], \ldots, \mathbf{X}_0[j]]$.

Observe that the sequence of $\mathbf{Y}_0[j], \ldots, \mathbf{Y}_n[j]$ forms a martingale. In addition, for any $j \in [M]$ and $1 \leq t \leq n$, $|\mathbf{Y}_t[j] - \mathbf{Y}_{t-1}[j]| \leq \ln 3$, because

$$\mathbf{Y}_t[j] - \mathbf{Y}_{t-1}[j] = \mathbf{X}_t[j] - \mathbf{X}_{t-1}[j] + \mathbb{E}[\mathbf{X}_{t-1}[j] - \mathbf{X}_t[j] \mid \mathbf{X}_{t-1}[j], \ldots, \mathbf{X}_0[j]]$$

14

and $\mathbf{X}_t[j] - \mathbf{X}_{t-1}[j] = \ln\left(1 + \frac{\mathbf{D}_t[j]}{\mathbf{Z}_{t-1}[j]} \cdot \frac{\psi(F_t)}{2}\right) \in [\ln\frac{1}{2}, \ln\frac{3}{2}]$ ($|\mathbf{D}_t[j]| \leq |\mathbf{Z}_{t-1}[j]|$ has been established).

By Azuma's inequality, it holds that for any k, $\Pr[|\mathbf{Y}_n[j]| > k] \leq e^{-k^2/2n(\ln 3)^2}$. By union bound, it holds that

$$\Pr[\forall j, |\mathbf{Y}_n[j]| \leq k] \geq 1 - M \cdot e^{-k^2/2n(\ln 3)^2}. \tag{9}$$

We claim that there exists $j_0 \in [M]$ such that for any $j \neq j_0$,

$$\mathbf{X}_n[j] - \mathbf{Y}_n[j] \geq \frac{nv}{48}. \tag{10}$$

Then (9) and (10) imply

$$\Pr[\exists j_0 \in [M], \text{ s.t. } \forall j \neq j_0, \mathbf{X}_n[j] \geq \frac{nv}{48} - t] \geq 1 - M \cdot e^{-k^2/2n(\ln 3)^2}.$$

Plugging in $k = nv/96$, we obtain (8) for $n \geq C(m + \log(1/\varepsilon))$ where $C \geq \max(96, 192(\ln 3)^2/v^2) \cdot \ln 2$. It remains to prove (10). By expanding the recursive relation of $\mathbf{Y}_t$, for any $t \geq 1$, we have

$$\mathbf{Y}_t[j] - \mathbf{X}_t[j] = \sum_{k=0}^{t-1} \left(\mathbf{X}_k[j] - \mathbb{E}\left[\mathbf{X}_{k+1}[j] \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right]\right). \tag{11}$$

Intuitively, $\mathbf{Y}_t[j] - \mathbf{X}_t[j]$ accumulates the shifts of $\mathbf{X}_0, \ldots, \mathbf{X}_{t-1}$ from being a martingale. Now for any k we compute

$$
\begin{aligned}
\mathbf{X}_k[j] - \mathbb{E}\left[\mathbf{X}_{k+1}[j] \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right] &= \mathbb{E}\left[\ln\frac{\mathbf{Z}_{k+1}[j]}{\mathbf{Z}_k[j]} \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right] \\
&= \mathbb{E}\left[\ln\left(1 + \frac{\mathbf{D}_{k+1}[j]}{\mathbf{Z}_k[j]} \cdot \frac{\psi(F_k)}{2}\right) \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right] \\
&\leq -\frac{1}{6}\mathbb{E}\left[\left(\frac{\mathbf{D}_{k+1}[j]}{\mathbf{Z}_k[j]} \cdot \frac{\psi(F_k)}{2}\right)^2 \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right],
\end{aligned}
$$

where the second equation is by $\mathbf{Z}_{k+1}[j] = \mathbf{Z}_k[j] + \mathbf{D}_{k+1}[j] \cdot \frac{\psi(C_k)}{2}$ and the third equation is because, by Taylor expansion, $\ln(1+x) = \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1}}{\ell!}x^\ell$ and $\ln(1+x) \leq -\frac{x^2}{6}$ for $x = \frac{\mathbf{D}_{k+1}[j]}{\mathbf{Z}_k[j]} \cdot \frac{\psi(C_k)}{2} \in (-1/2, 1/2)$ (note that we have already established $|\mathbf{D}_{k+1}[j]| \leq |\mathbf{Z}_k[j]|$). Furthermore, for any k and $j \neq \mathbf{s}_k[M]$, because $|\mathbf{d}_{k+1}[j]| = |\mathbf{z}_k[j]|$, we find that

$$\mathbf{X}_k[j] - \mathbb{E}\left[\mathbf{X}_{k+1}[j] \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right] \leq -\frac{1}{6}\mathbb{E}\left[\left(\frac{\psi(F_k)}{2}\right)^2 \,\middle|\, \mathbf{X}_0[j], \ldots, \mathbf{X}_k[j]\right] \leq -\frac{v}{24}, \tag{12}$$

Because for every k, there exists at most a single j such that $|\mathbf{d}_{k+1}[j]| \neq |\mathbf{z}_k[j]|$. By averaging argument, there exists at most a single j_0, such that $|\mathbf{d}_{k+1}[j]| \neq |\mathbf{z}_k[j]|$ happens at least $n/2$ times for $k = 0, \ldots, n$. For other $j \neq j_0$, $|\mathbf{d}_{k+1}[j]| = |\mathbf{z}_k[j]|$ happens at least $n/2$ times. Thus by (12), (11), for any $j \neq j_0$,

$$\mathbf{Y}_n[j] - \mathbf{X}_n[j] \leq \sum_{k:|\mathbf{d}_{k+1}[j]|=|\mathbf{z}_k[j]|} \left(-\frac{v}{24}\right) + \sum_{k:|\mathbf{d}_{k+1}[j]|\neq|\mathbf{z}_k[j]|} 0 \leq -\frac{v}{24} \cdot \frac{n}{2} = -\frac{nv}{48}.$$

The desired conclusion follows. $\qquad\square$

A more efficient implementation. A straightforward implementation of the extractor in Proposition 5 is simultaneous manipulation of 2^m martingales, which requires time $nm2^m$. To extract $\omega(\log n)$ number of bits, the extractor will run in super-polynomial time which is inefficient.

We give a more efficient implementation of our extractor which orchestrates all the 2^m martingale updates simultaneously in time $\min\{nm2^m, n^{O(|\mathcal{F}|)}\}$. When the number of dice $|\mathcal{F}|$ is a constant, the extractor runs in polynomial time even for extracting linear number of bits.

For every $0 \leq t \leq n$, we use two lists (L_t, LL_t) to keep track of martingales in $\mathbf{Z}_t$. LL_t keeps track of martingales who have been one of the largest martingales in last t steps. And L_t keeps necessary information for other martingales. Every element in LL_t is a pair (v, t_v) which represents a martingale with value v whose first time being the largest martingale is in $\mathbf{Z}_{t_v}$. Every element in L_t is a pair of $(v, c(v))$ where v is a value and $c(v)$ is the number of martingales in $\mathbf{Z}_t$ (but not in LL_t) with value v. In particular, we define $LL_0 = \emptyset$ and $L_0 = \{(1/2^m, 2^m)\}$.

The main observation is that every martingale in L_t has been updated in a multiplicative way where the multiplicative factor comes from a fixed set $\{1 \pm \psi(f)/2 : f \in \mathcal{F}\}$. Thus the size L_t is at most $\binom{t+2|\mathcal{F}|}{2|\mathcal{F}|} = n^{O(|\mathcal{F}|)}$. Moreover, there are at most t martingales in LL_t and at most 2^m martaingales in total. Therefore we will operate on at most $\min(2^m, n^{O(|\mathcal{F}|)} + n)$ objects in every step. In particular, sorting values among LL_{t-1} and L_{t-1} runs in $\min(m2^m, n^{O(|\mathcal{F}|)})$ and after that, the order of jth martingale with value v in LL_{t-1} or a martingale in L_{t-1} can be obtained in time $\min(2^m, n^{O(|\mathcal{F}|)})$.

Given LL_{t-1}, L_{t-1}, it is sufficient to continue the update rule to obtain LL_t, L_t. We go over $(v, c(v)) \in L_{t-1}$ in order (increasing in v) and we derive at most two groups of martingales with values $v_1 = v \cdot (1 + \psi(f)/2)$ and $v_2 = v \cdot (1 - \psi(f)/2)$. Moreover, knowing the order in Z_{t-1} allows us to know the size of each group and to update $(v_1, c(v_1))$ and $(v_2, c(v_2))$ in L_t accordingly. Updating values in LL_{t-1} is straightforward. In the end, if the largest value v is in L_t (if several martingales have the same value, take the one with highest order in step $t-1$), then we move the martingale from LL_t to L_t by adding (v, t) into LL_t and updating $(v, c(v))$ to $(v, c(v) - 1)$ in L_t.

From LL_n, we obtain the largest martingale in $\mathbf{Z}_n$. In order to track back its identity, for $t = 0$, we define the order of all martingales by their indexes and for every $1 \leq t \leq n$, we define the order martingales with the same value in LL_t by their orders in $\mathbf{Z}_{t-1}$. We prove by strong induction that given the order of a martingale in $\mathbf{Z}_t$ for $0 \leq t \leq n$, we can track back its order in $\mathbf{Z}_0$ which is its identify.

The base case is true for $t = 0$. Suppose it holds for $t' \leq t - 1$. Given a martingale in LL_t with value v, if $t_v \leq t - 1$, we can apply induction hypothesis to track back the largest one in $\mathbf{Z}_{t_v}$. If $t_v = t$, then we run our updating procedures on LL_{t-1} to identify the last martingale in L_t whose value becomes v. Similarly, for any j, given the jth martingale with value v in L_t, we run the updating procedures on L_{t-1} to identify the jth martingale whose value becomes v in L_t. Given L_{t-1} and LL_{t-1}, knowing its order among martingales with the same value is sufficient to identify its order among all martingales in $\mathbf{Z}_{t-1}$. So we can apply induction hypothesis to track back its identity in Z_0.

3.3 A lower bound on the quality of extraction

The implication $2 \to 1$ in Theorem 2 follows readily from Propositions 7 and 8 below. These refer to an analytic condition that characterizes randomness-efficient extractability called the mean-variance divergence (MVD) condition, which can be viewed as the suitable analogue of the MVR condition in Section 2.

The *kernel* of GSV source $(\mathcal{F}, \mathcal{D})$, denoted by $\operatorname{Ker} \mathcal{D}$, is the set of all $\psi \colon \mathcal{F} \to \mathbb{R}$ such that $\mathbb{E}_d[\psi_d(F)] = 0$ for all dice $d \in \mathcal{D}$.

Proposition 6. *A GSV source $(\mathcal{F}, \mathcal{D})$ satisfies (NK$^+$) if and only if for every die $d \in \mathcal{D}$ there exists a function $\psi_d \in \operatorname{Ker} \mathcal{D}$ that is not constant on the support of d.*

Proof. The forward direction follows by setting all ψ_d to equal the witness ψ for the (NK$^+$) condition.

For the reverse direction, let $\psi = \sum_{d \in \mathcal{D}} N_d \psi_d$ where N_d are independent random variables, each uniformly distributed over some finite set $\mathcal{N} \subseteq \mathbb{R}$ of size more than $|D|$. By linearity, ψ is in $\operatorname{Ker} \mathcal{D}$. Moreover, for each die d and each possible choice of the values $N_{d'}$ for $d' \neq d$, the sum $\sum N_d \psi_d$ can be constant on the support of d for at most one choice of N_d (for if two such choices existed then ψ_d itself must be constant on the support of d). Therefore, ψ is constant on d with probability at most $1/|\mathcal{N}|$. Since $|\mathcal{N}| > |\mathcal{D}|$, the existence of an (NK$^+$) witness ψ follows from the union bound. $\square$

Claim 6. *If (NK$^+$) fails for GSV source $(\mathcal{F}, \mathcal{D})$ then there exists a die $d \in \mathcal{D}$ such that for every pair of faces f^*, f_* in the support of $\mathcal{D}$ there exists a function $\beta \colon \mathcal{D} \to \mathbb{R}$ such that for all functions $\psi \colon \mathcal{F} \to \mathbb{R}$,*

$$\psi(f^*) - \psi(f_*) = \sum_{d' \in \mathcal{D}} \beta(d') \cdot \mathbb{E}_{d'}[\psi(F)]. \tag{13}$$

Proof. If $f^* = f_*$ the conclusion holds with $\beta = 0$. Otherwise, let $\mathcal{C}_d$ denote the linear space of functions that are constant on the support of die d. By Proposition 6, if (NK$^+$) fails then there exists a die d for which all functions $\psi \in \operatorname{Ker} \mathcal{D}$ also belong to $\mathcal{C}_d$, i.e., $\operatorname{Ker} \mathcal{D} \subseteq \mathcal{C}_d$. Then $\mathcal{C}_d^\perp \subseteq (\operatorname{Ker} \mathcal{D})^\perp$, where $\perp$ indicates the dual subspace. The space $(\operatorname{Ker} \mathcal{D})^\perp$ is the span of the probability mass functions pmf_d of all the dice. Therefore every $\phi \in \mathcal{C}_d^\perp$ can be written as a linear combination

$$\phi = \sum_{d \in \mathcal{D}} \beta(d) \cdot \operatorname{pmf}_d.$$

Then for every $\psi \colon \mathcal{F} \to \mathbb{R}$,

$$\sum_{f \in \mathcal{F}} \phi(f) \cdot \psi(f) = \sum_{d \in \mathcal{D}, f \in \mathcal{F}} \beta(d) \cdot \operatorname{pmf}_d(f) \cdot \psi(f) = \sum_{d \in \mathcal{D}} \beta(d) \cdot \mathbb{E}_d[\psi(F)].$$

The claim follows by specializing ϕ to the function that takes value 1 on f^*, -1 on f_*, and 0 elsewhere. This function is dual to $\mathcal{C}_d$. $\square$

The MVD(ε, δ) (mean-variance divergence) condition postulates that there exists a function $\psi \colon \mathcal{F} \to [-1, 1]$ such that for every die $d \in \mathcal{D}$,

$$\left| \mathbb{E}_d[\psi(F)] \right| < \epsilon(\operatorname{Var}_d[\psi(F)] - \delta). \tag{MVD}$$

Proposition 7. *If GSV source $(\mathcal{F}, \mathcal{D})$ fails (NK^+) then there exists a constant C such that for every $\varepsilon > 0$, $(\mathcal{F}, \mathcal{D})$ fails $MVD(\varepsilon, C\varepsilon^2)$.*

Proof. Assume $(\mathcal{F}, \mathcal{D})$ fails (NK^+). Let d be the die stipulated by Claim 6 and C be the maximum of $(\sum_{d' \in \mathcal{D}} |\beta(d')|)^2$ over all pairs of faces f^*, f_* in the support of d.

Towards a contradiction suppose that $(\mathcal{F}, \mathcal{D})$ satisfies $MVD(\varepsilon, \delta)$. Then the witness $\psi \colon \mathcal{F} \to [-1, 1]$ for $MVD(\varepsilon, \delta)$ must satisfy the conditions $\mathrm{Var}_d[\psi(F)] > \delta$ and $|\mathbb{E}_{d'}[\psi(F)]| < \varepsilon \mathrm{Var}_{d'}[\psi(F)]$ for all dice $d' \in \mathcal{D}$. Let f^* and f_* be faces in the support of d that maximize and minimize the value of ψ, respectively. By Claim 6, relation (13) holds for some β that may depend on f^* and f_* but not on ψ. Then

$$\sqrt{\delta} < \sqrt{\mathrm{Var}_d[\psi(F)]} \le \psi(f^*) - \psi(f_*) = \sum_{d' \in \mathcal{D}} \beta(d') \cdot \mathbb{E}_{d'}[\psi(F)]$$

$$\le \sum_{d' \in \mathcal{D}} |\beta(d')| \cdot |\mathbb{E}_{d'}[\psi(F)]| < \sum_{d' \in \mathcal{D}} |\beta(d')| \cdot \varepsilon \mathrm{Var}_{d'}[\psi(F)] \le \sqrt{C}\varepsilon,$$

where the last inequality follows from the definition of C and the boundedness of ψ. Therefore $MVD(\varepsilon, \delta)$ fails for $\delta = C\varepsilon^2$. $\qquad\square$

Proposition 8. *Assume that $(\mathcal{F}, \mathcal{D})$ fails $MVD(\varepsilon, \delta)$. Then every extractor with error $\varepsilon/20$ for $(\mathcal{F}, \mathcal{D})$ requires $1/8\delta$ samples, assuming $\varepsilon > 0$ is sufficiently small.*

Proof. The proof is a direct extension of the proof of Proposition 2. The main technical tool is the following claim:

Claim 7. *For every extractor $Ext \colon \mathcal{F}^n \to \{0, 1\}$, and every $0 \le \alpha \le 1$, if $\mathbb{E}_{A_-}[Ext] \ge \alpha$ for every strategy A_-, then there exists a strategy A_+ for which*

$$\mathbb{E}_{A_+}[Ext] \ge \alpha + \frac{\varepsilon}{1 + \varepsilon} \cdot \big(\alpha(1 - \alpha) - \delta n\big).$$

The proof of Claim 7 is a notationally intensive direct extension of the proof of Claim 2. We omit the details.

By Claim 7 it follows that for every $\varepsilon > 0$, if no strategy A_- has error less than $\alpha = 1/2 - \varepsilon/20$ against Ext then there exists a strategy A_+ with advantage at least

$$\mathbb{E}[Ext] \ge \frac{1}{2} - \frac{\varepsilon}{20} + \frac{\varepsilon}{1 + \varepsilon} \cdot \Big(\frac{1 - \varepsilon^2/400}{4} - \frac{1}{8}\Big),$$

which is at least $1/2 + \varepsilon/20$ for sufficiently small ε. $\qquad\square$

4 Open Questions

In this work, we completely classify GSV sources in terms of their extractability. We point out the following questions for further investigation:

- Is the sample complexity of $o(1/\varepsilon^2)$ in part 2 of Theorem 2 tight? Example E2 gives an upper bound of $O(1/\varepsilon^7)$. This non-NK$^+$ extractable source satisfies MVR(ε) with minimum variance ε^2 for every ε (with witness $\psi = (\varepsilon, -\varepsilon, 1, -1)$), so $O(1/\varepsilon^7)$ are sufficient for extraction error ε by Proposition 1.

- The number of required samples in Theorem 1 is of the form $\varepsilon^{-O(2^{|\mathcal{D}|})}$, where $|\mathcal{D}|$ is the number of dice (see the proof of Proposition 3). Is this exponential dependence in $|\mathcal{D}|$ necessary?

- The multi-bit extractor in Theorem 2 runs in time $\min(nm2^m, n^{O(|\mathcal{F}|)})$. Can the dependence on the number of faces be improved, possibly by applying known seeded extraction algorithms?

- Proposition 1 states that sources satisfying condition $\mathrm{MVR}(\varepsilon)$ admit extraction with error $O(\sqrt{\varepsilon})$, while by Proposition 2 extraction error $\Omega(\varepsilon)$ is necessary. Can this quadratic gap be narrowed?

Acknowledgments

References

[BEG15]   Salman Beigi, Omid Etesami, and Amin Gohari. Deterministic randomness extraction from generalized and distributed santha-vazirani sources. In *International Colloquium on Automata, Languages, and Programming*, pages 143–154. Springer, 2015.

[BEG17]   Salman Beigi, Omid Etesami, and Amin Gohari. Deterministic randomness extraction from generalized and distributed santha-vazirani sources. *SIAM Journal on Computing*, 46(1):1–36, 2017.

[Bou05]   Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[Bou07]   Jean Bourgain. On the construction of affine extractors. *GAFA Geometric And Functional Analysis*, 17(1):33–57, 2007.

[CZ16]   Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 670–683, 2016.

[DGW09]   Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *Computational Complexity*, 18(1):1–58, 2009.

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology - EUROCRYPT 2006, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28 - June 1, 2006, Proceedings*, pages 486–503, 2006.

[Dvi12]   Zeev Dvir. Extractors for varieties. *Computational complexity*, 21(4):515–572, 2012.

[Gab11]   Ariel Gabizon. Deterministic extractors for affine sources over large fields. In *Deterministic Extraction from Weak Random Sources*, pages 33–53. Springer, 2011.

[SV86]   Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *J. Comput. Syst. Sci.*, 33(1):75–87, 1986.

[Vad12]   Salil P Vadhan. *Pseudorandomness*, volume 56. Now, 2012.