On the Impossibility of Extracting Classical Randomness Using a Quantum Computer

Yevgeniy Dodis[1] and Renato Renner[2]

[1] New York University, USA. Email: `dodis@cs.nyu.edu`
[2] University of Cambridge, UK. Email: `r.renner@damtp.cam.ac.uk`

Abstract. In this work we initiate the question of whether quantum computers can provide us with an almost perfect source of classical randomness, and more generally, suffice for classical cryptographic tasks, such as encryption. Indeed, it was observed [SV86,MP91,DOPS04] that classical computers are insufficient for either one of these tasks when all they have access to is a realistic *imperfect* source of randomness, such as the Santha-Vazirani source.

We answer this question in the *negative*, even in the following very restrictive model. We generously assume that quantum computation is error-free, and *all the errors come in the measurements*. We further assume that all the measurement errors are not only small but also *detectable*: namely, all that can happen is that with a small probability $p_\perp \leq \delta$ the (perfectly performed) measurement will result in some distinguished symbol $\perp$ (indicating an "erasure"). Specifically, we assume that if an element x was supposed to be observed with probability p_x, in reality it might be observed with probability $p'_x \in [(1-\delta)p_x, p_x]$, for some small $\delta > 0$ (so that $p_\perp = 1 - \sum_x p'_x \leq \delta$).

Our negative "quantum" result also implies a new "classical" result of independent interest: namely, even a much more restrictive form of (classical) Santha-Vazirani sources is not sufficient for randomness extraction and cryptography.

1 Introduction

Randomness is important in many areas of computer science, such as algorithms, cryptography and distributed computing. A common abstraction typically used in these applications is that there exists some source of unbiased and independent random bits. However, in practice this assumption seems to be problematic: although there seem to be many ways to obtain somewhat random data, this data is almost never uniformly random, its exact distribution is unknown, and, correspondingly, various algorithms and protocols have to be based on *imperfect sources of randomness*.

Not surprisingly, a large body of work (see below) has attempted to bridge the gap between this convenient theoretical abstraction and the actual reality. So far, however, most of this work concentrated on studying if *classical* computers can effectively use classical imperfect source of randomness. In this work, we initiate the corresponding study regarding *quantum computation*. To motivate our question, we start by surveying the state of the art in using classical computers, which will demonstrate that such computers are provably incapable of tolerating even "mildly" imperfect random sources.

CLASSICAL APPROACH TO IMPERFECT RANDOMNESS. The most straightforward approach to dealing with an imperfect random source is to *deterministically* (and efficiently) extract nearly-perfect randomness from it. Indeed, many such results were obtained for several classes of imperfect random sources. They include various simple "streaming" sources [von51,Eli72,Blu86,LLS89], different flavors of "bit-fixing" sources [CGH$^+$85,BBR88,AL93,CDH$^+$00,DSS01,KZ03], multiple independent imperfect random sources [SV86,Vaz87a,Vaz87b,CG88,DO03,DEOR04,BIW04] and efficiently samplable sources [TV00]. While these results are interesting and non-trivial, the above "deterministically extractable" sources assume a lot of structure or independence in the way they generate randomness. A less restrictive, and arguably more realistic, assumption on the random source would be to assume only that the source contains *some* entropy. We call such sources *entropy sources*. Entropy sources were first introduced by Santha and Vazirani [SV86], and later generalized by Chor and Goldreich [CG88], and Zuckerman [Zuc96].

The entropy sources of Santha and Vazirani [SV86] are the least imperfect (which means it is the hardest to show impossibility results for such sources) among the entropy sources considered so far (e.g., as compared to [CG88,Zuc96]). SV-sources, as they are called, require *every bit output by the source* to have almost one bit of entropy, even when conditioned on all the previous bits. Unfortunately, already the original work of [SV86] (see also a simpler proof in [RVW04]) showed that deterministic randomness extraction of even a single bit is *not* possible from all SV sources. This can also be considered as impossibility of pseudo-random generators with access to only an SV source. Moreover, this result was later extended by McInnes and Pinkas [MP91], who showed that in the classical setting of computationally unlimited adversaries, one cannot have secure symmetric encryption if the shared key comes from an SV source. Finally and most generally, Dodis et al. [DOPS04] showed that SV sources in fact cannot be used essentially for any interesting classical cryptographic task involving privacy (such as encryption, commitment, zero-knowledge, multiparty computation), even when restricting to computationally bounded adversaries. Thus, even for currently most restrictive en-

tropy sources, classical computation does not seem to suffice for applications inherently requiring randomness (such as extraction and cryptography).[1]

We also mention that the impossibility results no longer hold when the extracting party has a small amount of true randomness (this is the study of so called probabilistic *randomness extractors* [NZ96]), or if several *independent* entropy sources are available [SV86,Vaz87a,Vaz87b,CG88,DO03,DEOR04,BIW04].

QUANTUM COMPUTERS? Given the apparent inadequacy of classical computers to deal with entropy sources — at least for certain important tasks such as cryptography —, it is natural to ask if quantum computers can be of help. More specifically, given that quantum computation is *inherently probabilistic*, can we use quantum computers to generate nearly perfect randomness? (Or maybe just "good enough" randomness for cryptographic tasks like encryption, which, as we know [DS02], do not require *perfect* randomness?) For example, to generate a perfectly random bit from a fixed qubit $|0\rangle$, one can simply apply the Hadamard transform, and then measure the result in the standard basis. Unfortunately, what prevents this simple solution from working in practice is the fact that it is virtually impossible to perform the above transformation (in particular, the measurement) precisely, so the resulting bit is likely to be slightly biased. More generally, noise is a very serious issue in quantum computation, which means that certain error-correction and fault-tolerance must be applied in order to overcome such noise. Indeed, fault-tolerance is one of the major problems in quantum computing (see [NC00]), so we will have to address it as well. Jumping ahead, however, what will differentiate us from all the prior work in the area is the fact that *we do not assume largely independent noise* (which can be dealt with by quantum error-correction).

But, first, let us explain why there are good reasons to hope for quantum computers to be useful despite the noise. When dealing with classical imperfect sources, we usually assume that the source comes from some family of distributions "outside of our control" (e.g., "nature"), so we would like to make as few assumptions about these distributions as we can. For example, this is why the study of imperfect randomness quickly converged to entropy sources as being the most plausible sources one could get from nature. In contrast, by using a quantum computer to generate our random source for us, we are *proactively designing* a source of randomness which is convenient for use, rather than *passively hoping* that nature will give us such a source. Indeed, if not for the noise, it would be trivial to generate ideal randomness in our setting. Moreover, even with noise we have a lot of freedom in *adapting* our quantum computer to generate and measure quantum states *of our choice*, depending on the computation so far.

OUR MODEL. We first define a natural model for using a (realistically noisy) quantum computer for the task of randomness extraction (or, more generally, any probabilistic computation, such as the one needed in classical cryptography). As we will see shortly, we will prove a *negative* result in our model, despite the optimism we expressed in the previous paragraph. Because of this, we will make the noise as small and as restrictive

---

[1] In contrast, a series of celebrated positive results [VV85,SV86,CG88,Zuc96] show that even very weak entropy sources are enough for simulating probabilistic polynomial-time algorithms — namely, the task which does not *inherently* need randomness. This result was extended to interactive protocols by [DOPS04]. [DOPS04] also show that under certain strong, but reasonable computational assumptions, secure signatures seem to be possible with entropy sources.

as we can, even if these restrictions are completely "generous" and unrealistic. Indeed, *we will assume that the actual quantum computation is error-free, and all the errors come in the measurements* (which are necessary to extract some classical result out of the system). Of course, in reality the quantum computation will also be quite noisy, but our assumption will not only allow us to get a stronger result, but also reduce our "quantum" question to a natural "purely classical" question of independent interest.

Moreover, we will further assume that all the measurement errors are not only very small, but also *detectable*: namely, all that can happen is that with a small probability $p_\perp \leq \delta$ the (perfectly performed) measurement will result in some distinguished symbol $\perp$ (indicating an "erasure"). Specifically, we assume that if an element x was supposed to be observed with probability p_x, in reality it might be observed with probability $p'_x \in [(1-\delta)p_x, p_x]$, for some small $\delta > 0$ (so that $p_\perp = 1 - \sum_x p'_x \leq \delta$). Thus, the adversary is not allowed to completely "remove" undesirable events of small probability, or increase the probability of some event. Moreover, as compared to the classical SV model, in our model the state to be measured can be prepared arbitrarily, irrespective of the computational complexity of preparing this state. Further, such quantum states can even be generated adaptively and based on the measurements so far. For comparison, in the SV model the "ideal" measurement would always correspond to an unbiased bit; additionally, SV model allows for "errors" while we only allow "erasures".

OUR RESULT. Unfortunately, our main result will show that even in this *extremely restrictive* noise model, one cannot extract even a *single* nearly uniform bit. In other words, if the measurement errors could be *correlated*, quantum computers do not help to extract classical randomness. More generally, we extend the technique of [DOPS04] to our model and show that one cannot generate two (classical) computationally indistinguishable distributions which are not nearly identical to begin with. This can be used to show the impossibility of classical encryption, commitment, zero-knowledge and other tasks exactly as in [DOPS04]. We notice, however, that our result does *not* exclude the possibility of generating perfect entanglement, which might be used to encrypt a message into a *quantum state*. Nevertheless, our result implies that, even with the help of such perfect entanglement, the user will not be able to generate a (shared) classical key that can be used for cryptographic tasks. To summarize, we only rule out the possibility of *classical* cryptography with *quantumly generated randomness*, leaving open the question of (even modeling!) *quantum* cryptography with noise.

Of independent interest, we reduce our "quantum" problem to the study of a new classical source, which is considerably more restrictive than the SV source (and this restriction can really be enforced in our model). We then show a classical impossibility result for our new source, which gives a non-trivial generalization of the corresponding impossibility result for the SV sources [DOPS04,SV86].[2] To summarize, our main results can be viewed in three areas:

1. A model of using noisy quantum computers for classical probabilistic computation.
2. A reduction from a "quantum" question to the classical question concerning a much more restrictive variant of the SV source.
3. The non-trivial impossibility result for the classical source we define.

---

[2] Indeed, our proof is *considerably* more involved than the above mentioned results, and we were unable to find any simplification.

RELATION TO QUANTUM ERROR-CORRECTION. As we stated, what differentiates us from the usual model of quantum computation with noise is the fact that our errors are not assumed independent. In particular, conventional results on fault-tolerant quantum computation (such as the *threshold theorem*; see [NC00] for more details) do not apply in our model (as is apparent from our negative results). From another perspective, our impossibility result is not just a trivial application of the principle that one can always and without loss of generality postpone all the measurements until the end (a useful observation true in the "perfect measurement" case). For example, if all the measurements are postponed to the end, then the adversary can cause us to observe a single "useless" $\perp$ symbol with non-trivial probability δ, while with many measurements we are bound to observe a lot of "useful" non-$\perp$ symbols which exponentially high probability.

Nevertheless, in our model one can trivially simulate probabilistic algorithms computing deterministic outputs, just as was the case for the classical computation, where this was true even for much weaker sources than SV sources [CG88,Zuc96]. For example, here we actually *can* postpone all the measurements until the end, and then either obtain an error (with probability at most δ in which case the computation can be repeated), or the desired result (with probability arbitrarily close to $1 - \delta$). Of course, this "positive" result only holds because our noise model was made unrealistically restrictive (since we proved an *impossibility result*). Thus, it would be interesting to define a less restrictive (and more realistic!) error model — for example where the actual quantum computation is not error-free — and see if this feasibility result would still hold.

Finally, the problem of detection errors has been studied in the context of non-locality testing [CH74,MSS83,Mas02], which tries to experimentally prove the intriguing phenomenon that the behavior of certain distant but entangled particles cannot be explained by classical randomness. These results are of the same flavor as our impossibility result. Indeed, they show that, if certain detection probability is too low, then the outputs might be chosen in a malicious way such that the resulting statistics does not imply non-locality. To our knowledge, this is the only result where some impossibility is proved, based on the assumption that certain errors occur.

2 Definition of the source

A source with n outputs $X_1, X_2, \ldots, X_n$ is specified by a joint probability distribution $P_{X_1 \cdots X_n}$. However, for most realistic sources, the actual distribution $P_{X_1 \cdots X_n}$ can usually not be fully determined. Instead, only a few characteristics of the source are known, e.g., that the conditional probability distributions[3] $P_{X_i|X^{i-1}}$ have certain properties. A well-known example for such a characterization are the Santha-Vazirani sources.

Definition 1 ([SV86]). *A probability distribution $P_{X_1 \cdots X_n}$ on $\{0,1\}^n$ is an α-SV source if*[4]

$$P_{X_i|X^{i-1}=x^{i-1}}(0) \in [\alpha, 1 - \alpha] \,,$$

for all $i \in \{1, \ldots, n\}$ and $x^{i-1} \in \{0,1\}^{i-1}$.

---

[3] We write X^k to denote the k-tuple $(X_1, \ldots, X_k)$.

[4] $P_{X_i|X^{i-1}=x^{i-1}}$ denotes the probability distribution of X_i conditioned on the event that the $(i-1)$-tuple $X^{i-1} = (X_1, \ldots, X_{i-1})$ takes the value $x^{i-1} = (x_1, \ldots, x_{i-1})$.

We will define a more general class of sources which, in some sense, includes the SV sources (cf. Section A). The main motivation for our definition is to capture any kind of randomness that can be generated using imperfect (quantum) physical devices. Indeed, we will show in Section 3 that the randomness generated by any imperfect physical device cannot be more useful than the randomness obtained from a source as defined below.

Intuitively, a source can be seen as a box which sequentially outputs symbols $X_1, \ldots, X_n$ from some alphabet $\mathcal{X}$. Each output X_i is chosen according to some fixed probability distribution which might depend on all previous outputs $X_1, \ldots, X_{i-1}$. The "imperfectness" of the source is then modeled as follows. Each output X_i is "erased" with some probability p, i.e., it is replaced by some fixed symbol $\perp$. This erasure probability might depend on the actual output X_i as well as on all previous outputs $X_1, \ldots, X_{i-1}$, but is upper bounded by some fixed parameter δ.

Before stating the formal definition, let us introduce some notation to be used in the sequel. For any set $\mathcal{X}$, we denote by $\bar{\mathcal{X}}$ the set $\bar{\mathcal{X}} := \mathcal{X} \cup \{\perp\}$ which contains an extra symbol $\perp$. For a probability distribution P_X on $\mathcal{X}$ and $\delta \geq 0$, let $\mathcal{P}^{\delta}(P_X)$ be the set of probability distributions $\bar{P}_X$ on $\bar{\mathcal{X}}$ such that

$$(1 - \delta)P_X(x) \leq \bar{P}_X(x) \leq P_X(x) \,,$$

for all $x \in \mathcal{X}$. In particular, the probability of the symbol $\perp$ is bounded by δ, that is, $\bar{P}_X(\perp) \leq \delta$.

Definition 2. *Let $\delta \geq 0$ and let, for any $i \in \{1, \ldots, n\}$, $Q_{X_i|X_{i-1}}$ be a channel from $\bar{\mathcal{X}}^{i-1}$ to $\mathcal{X}$. A probability distribution $P_{X_1 \cdots X_n}$ on $\bar{\mathcal{X}}^n$ is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source if*

$$P_{X_i|X^{i-1}=x^{i-1}} \in \mathcal{P}^{\delta}(Q_{X_i|X^{i-1}=x^{i-1}}) \,,$$

for all $i \in \{1, \ldots, n\}$ and $x^{i-1} = (x_1, \ldots, x_{i-1}) \in \bar{\mathcal{X}}^{i-1}$.

In Appendix A, we show that $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources can be used to simulate α-SV sources, for some appropriately chosen α. This means that $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources are at least as useful as SV sources. The other direction is, however, not true. That is, $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources have a strictly less "malicious" behavior than SV sources (which makes our impossibility proofs stronger).

3 The quantum model

In this section, we propose a model that describes the extraction of classical information from imperfect quantum physical devices. Clearly, our considerations also include purely classical systems as a special case.

First, in Section 3.1, we review the situation where the quantum device is perfect. In this case, the process of extracting randomness can most generally be seen as a sequence of perfect quantum operations and perfect measurements. Then, in Section 3.2, we consider the general case where the quantum device is subject to (malicious) noise. We will argue that, in order to state strong impossibility results, it is sufficient to generalize the standard notion of measurements such as to include the possibility of detectable failures in the measurement process.

3.1 The perfect case

Let us briefly review some basic facts about quantum mechanics. The *state* of a quantum system is specified by a projector $P_{|\psi\rangle}$ onto a vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. More generally, if a system is prepared by choosing a state from some family $\{|\psi_z\rangle\}_{z\in\mathcal{Z}}$ according to a probability distribution P_Z on $\mathcal{Z}$, then the behavior of the system is fully described by the *density operator* $\rho := \sum_{z\in\mathcal{Z}} P_Z(z)P_{|\psi_z\rangle}$. The most general *operation* that can be applied on a quantum system is specified by a family $\mathcal{E} = \{E_x\}_{x\in\mathcal{X}}$ of operators on $\mathcal{H}$ such that $\sum_{x\in\mathcal{X}} E_x^\dagger E_x = \mathrm{id}_\mathcal{H}$ (see, e.g., [NC00]). When $\mathcal{E}$ is applied to a system which is in state ρ, then, with probability $P_X(x) := \mathrm{tr}(E_x\rho E_x^\dagger)$, the classical output $x \in \mathcal{X}$ is produced and the final state ρ_x of the system is $\rho_x := \frac{1}{P_X(x)}E_x\rho E_x^\dagger$. Hence, when ignoring the classical output x, the state $\mathcal{E}(\rho)$ of the system after applying the operation $\mathcal{E}$ is the average of the states ρ_x, that is, $\mathcal{E}(\rho) := \sum_x P_X(x)\rho_x = \sum_x E_x\rho E_x^\dagger$.

It is important to note that also the action of preparing a quantum system to be in a certain state ρ_0 can be described by a quantum operation $\mathcal{E}$. To see this, let ρ_0 be given by $\rho_0 = \sum_{z\in\mathcal{Z}} P_Z(z)P_{|\psi_z\rangle}$, for some family of vectors $\{|\psi_z\rangle\}_{z\in\mathcal{Z}}$ and a probability distribution P_Z on $\mathcal{Z}$. Additionally, let $\{|i\rangle\}_{i\in\{1,\dots,d\}}$ be an orthonormal basis of $\mathcal{H}$. It is easy to verify that the quantum operation $\mathcal{E} = \{E_{z,i}\}_{z\in\mathcal{Z},i\in\{1,\dots,d\}}$ defined by the operators

$$E_{z,i} := \sqrt{P_X(z)}|\psi_z\rangle\langle i|$$

maps any arbitrary state ρ to ρ_0, that is, $\mathcal{E}(\rho) = \rho_0$.

We are now ready to describe the process of randomness extraction from a quantum system. Consider a classical user with access to a quantum physical device. The most general thing he can do is to subsequently apply quantum operations, where each of these operations provides him with classical information which he might use to select the next operation. To describe this on a formal level, let $\mathcal{H}$ be a Hilbert space and let $\mathcal{X}$ be a set. The strategy of the user in each step i is then defined by the quantum operation $\mathcal{E}^{x^{i-1}} = \{E_x^{x^{i-1}}\}_{x\in\mathcal{X}}$ he applies depending on the classical outputs $x^{i-1} \in \mathcal{X}^{i-1}$ obtained in the previous steps. Note that, according to the above discussion, this description also includes the action of preparing (parts of) the quantum system in a certain state. We can thus assume without loss of generality that the initial state of the system is given by some fixed projector $P_{|\psi_0\rangle}$. The probability distribution $P_{X_i|X^{i-1}=x^{i-1}}$ of the classical outcomes in the ith step conditioned on the previous outputs x^{i-1} as well as the quantum state ρ_{x^i} after the ith step given the outputs x^i is then recursively defined by $\rho_{x^0} := P_{|\psi_0\rangle}$ and

$$P_{X_i|X^{i-1}=x^{i-1}}(x) := \mathrm{tr}\big(E_x^{x^{i-1}}\rho_{x^{i-1}}E_x^{x^{i-1}\dagger}\big) \tag{1}$$

$$\rho_{x^i} = \rho_{(x^{i-1},x)} := \frac{1}{P_{X_i|X^{i-1}=x^{i-1}}(x)}E_x^{x^{i-1}}\rho_{x^{i-1}}E_x^{x^{i-1}\dagger}\,. \tag{2}$$

3.2 Quantum measurements with malicious noise

We will now extend the model of the previous section to include situations where the quantum operations are subject to noise. As we are interested in proving the *impossibility* of certain tasks in the presence of noise, our results are stronger if we assume that

only parts of the quantum operation are noisy. In particular, we will restrict to systems where only the classical measurements are subject to perturbations.[5]

Formally, we define an imperfect quantum device by its behavior when applying any operation $\mathcal{E}$. Let $\delta \geq 0$ and let $\mathcal{E} = \{E_{x,u}\}_{x \in \mathcal{X}, u \in \mathcal{U}}$ be a quantum operation which produces two classical outcomes x and u, where x is the part of the output that is observed by the user. The operation $\mathcal{E}$ acts on the imperfect device as it would in the perfect case, except that each output x is, with some probability $\lambda_x \leq \delta$, replaced by a symbol $\perp$, indicating that something went wrong. Additionally, we assume that, whenever such an error occurs, the state of the system remains unchanged.[6] The resulting probability distribution P_X of the outputs when applying $\mathcal{E}$ to an imperfect device in state ρ is thus given by

$$P_X(x) := \sum_u (1 - \lambda_x) \operatorname{tr}(E_{x,u} \rho E_{x,u}^\dagger) \, ,$$

for any $x \in \mathcal{X}$. Note that the probability of the symbol $\perp$ is $P_X(\perp) = 1 - \sum_{x \in \mathcal{X}} P_X(x) \leq \delta$.

Let us now consider the interaction of a user with such an imperfect quantum device. In each step i, he either observes the correct outcome or he gets the output $\perp$, indicating that something went wrong. The user might want to use this information to choose the subsequent operations. His strategy is thus defined by a family $\{\mathcal{E}^{x^{i-1}}\}_{x^{i-1} \in \bar{\mathcal{X}}^{i-1}}$ of quantum operations $\mathcal{E}^{x^{i-1}} = \{E_{x,u}^{x^{i-1}}\}_{x \in \mathcal{X}, u \in \mathcal{U}}$.[7] The conditional probability distributions $P_{X_i | X^{i-1} = x^{i-1}}$ of the observed outputs in the ith step, for $x^{i-1} \in \bar{\mathcal{X}}^{i-1}$, and the states ρ_{x^i} after the ith step are recursively defined, analogously to (1) and (2), by

$$P_{X_i | X^{i-1} = x^{i-1}}(x) := (1 - \lambda_{x^{i-1}, x}) Q_{X_i | X^{i-1} = x^{i-1}}(x) \qquad \text{for } x \in \mathcal{X}$$

$$\rho_{x^i} = \rho_{(x^{i-1}, x)} := \begin{cases} \frac{1}{Q_{X_i | X^{i-1} = x^{i-1}}(x)} \sum_{u \in \mathcal{U}} E_{x,u}^{x^{i-1}} \rho_{x^{i-1}} E_{x,u}^{x^{i-1}\dagger} & \text{if } x \in \mathcal{X} \\ \rho_{x^{i-1}} & \text{if } x = \perp \, . \end{cases}$$

---

[5] To see that our model leads to strong impossibility results, consider for example an adversary who is allowed to transform the quantum state ρ of the device into a state ρ' which has at most trace distance δ to the original state ρ. Let $\mathcal{M}$ be a fixed measurement and let P be the distribution resulting from applying $\mathcal{M}$ to ρ. It is easy to see that, for any given probability distribution P' which is δ-close to P, the adversary can set the device into a state ρ' such that a measurement $\mathcal{M}$ of ρ' gives raise to the distribution P'. Consequently, such an adversary is at least as powerful as an adversary who can only modify the distribution of the measurement outcomes, as proposed in our model. In particular, our impossibility results also apply to this case.

[6] This means that, even if a measurement error occurs, the state of the quantum system is not destroyed. (Recall that, as we are interested in impossibility results, we would like to have a model for imperfect devices which is as close as possible to perfect systems.)

[7] Note that, unlike in the perfect case, the measurements cannot be postponed to the end of the protocol. For example, if the user performs many measurements during the protocol, it is very unlikely that *all* the outcomes are wrong, i.e., he still gets some useful information with probability almost one. On the other hand, if the user replaces all his measurements by *one* single overall measurement (at the end of the protocol) it might fail with probability δ.

where $Q_{X_i|X^{i-1}=x^{i-1}}$ are the conditional probability distributions given by

$$Q_{X_i|X^{i-1}=x^{i-1}}(x) := \sum_{u \in \mathcal{U}} \text{tr}(E^{x^{i-1}}_{x,u} \rho_{x^{i-1}} E^{x^{i-1}}_{x,u}{}^\dagger), \quad \text{for } x \in \mathcal{X},$$

and where $\lambda_{x^{i-1},x} \in [0,\delta]$.

Let $P_{X^n} = P_{X_1 \cdots X_n}$ be the probability distribution of the observed outcomes after n steps. It follows directly from the above formulas that P_{X^n} is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source. On the other hand, if P_{X^n} is a $(\delta, \{Q_{X_i|X^{i-1}})$-source, then there exist weights $\lambda_{x^{i-1},x} \in [0,\delta]$ such that the conditional probabilities are given by the above formulas. This reduces our "quantum" problem to a totally classical problem for an imperfect source considerably more restrictive than an SV source (see Lemma 2). The corresponding impossibility result is given in the next section.

4 Main technical lemma

Our main technical result can be seen as an extension of a result proved for SV sources (cf. Lemma 3.5 of [DOPS04]). Roughly speaking, Lemma 1 below states that a task g which requires perfect random bits can generally not be replaced by another task f which only uses imperfect bits. Note that this impossibility is particularly interesting for cryptography where many tasks do in fact use randomness.

More precisely, let g be an arbitrary strategy which uses imperfect randomness X^n and, in addition, some perfect randomness Y (whose probability distribution might be chosen depending on the values of X^n). Let f be another strategy which only uses imperfect randomness X^n. Furthermore, assume that, for any $(\delta, \{Q_{X_i|X^{i-1}}\})$-source $P_{X_1 \cdots X_n}$, the output distributions of the strategies g and f are (almost) identical. Then the strategy g is (roughly) the same as f, that is, it (virtually) does not use the randomness Y.

Lemma 1. *Let f be a function from $\bar{\mathcal{X}}^n$ to $\mathcal{Z}$ and let g be a function from $\bar{\mathcal{X}}^n \times \mathcal{Y}$ to $\mathcal{Z}$. For any $i \in \{1, \ldots, n\}$, let $Q_{X_i|X^{i-1}}$ be a channel from $\bar{\mathcal{X}}^{i-1}$ to $\mathcal{X}$, let $Q_{Y|X^n}$ be a channel from $\bar{\mathcal{X}}^n$ to $\mathcal{Y}$, and let $\delta \geq 0$. Let Γ be the set of all probability distributions $P_{X^n Y}$ on $\bar{\mathcal{X}}^n \times \mathcal{Y}$ such that P_{X^n} is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source[8] and $P_{Y|X^n} = Q_{Y|X^n}$. If, for all $P_{X^n Y} \in \Gamma$,*

$$|P_{f(X^n)} - P_{g(X^n,Y)}|_1 < \varepsilon,$$

then there exists $P_{\tilde{X}^n \tilde{Y}} \in \Gamma$ such that

$$\Pr_{(x^n,y) \leftarrow P_{\tilde{X}^n \tilde{Y}}}[f(x^n) \neq g(x^n,y)] < 5\varepsilon m \delta^{-1},$$

where $m = \lceil \log_2(|\mathcal{Z}|) \rceil$.

---

[8] Similarly to the argument in [DOPS04], the proof can easily be extended to a statement which holds for an even stronger type of sources, where the conditional probability distributions of each X_i given all other source outputs, and not only the previous ones X^{i-1}, is contained in a certain set $\mathcal{P}^\delta$.

Proof. Assume first that the functions f and g are binary, i.e., $\mathcal{Z} = \{0, 1\}$. The idea is to define two probability distributions $P_{V^n Y}, P_{W^n Y} \in \Gamma$ such that the output distributions of the function f, $f(V^n)$ and $f(W^n)$, are "maximally different". Then, by assumption, the output distributions of $g(V^n, Y)$ and $g(W^n, Y)$ must be different as well. This will then be used to conclude that the outputs of f and g are actually equal for most inputs.

In order to define the distributions $P_{V^n Y}$ and $P_{W^n Y}$, we first consider some "intermediate distribution" $P_{\tilde{X}^n \tilde{Y}}$. It is defined as the unique probability distribution on $\bar{\mathcal{X}}^n \times \mathcal{Y}$ such that $P_{\tilde{Y}|\tilde{X}^n} = Q_{Y|X^n}$ and, for any $i \in \{1, \ldots, n\}$ and $x^{i-1} \in \bar{\mathcal{X}}^{i-1}$,

$$P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}}(x) := \begin{cases} (1 - \frac{\delta}{2})Q_{X_i|X^{i-1}=x^{i-1}}(x) & \text{if } x \in \mathcal{X} \\ \frac{\delta}{2} & \text{if } x = \perp \, . \end{cases}$$

Note that $P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}}$ is indeed a probability distribution because

$$\sum_{x \in \bar{\mathcal{X}}} P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}}(x) = (1 - \tfrac{\delta}{2}) \sum_{x \in \mathcal{X}} Q_{X_i|X^{i-1}=x^{i-1}}(x) + \tfrac{\delta}{2} = 1 \, .$$

Moreover, $P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}} \in \mathcal{P}^\delta(Q_{X_i|X^{i-1}=x^{i-1}})$, i.e., $P_{\tilde{X}^n}$ is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source, and thus $P_{\tilde{X}^n \tilde{Y}} \in \Gamma$.

The distribution P_{V^n} is now defined from $P_{\tilde{X}^n}$ by raising the probabilities of all values[9] $x^n \in f^{-1}(0)$ that f maps to 0 and lowering the probabilities of all $x^n \in f^{-1}(1)$. Similarly, P_{W^n} is defined by changing the probabilities of $P_{\tilde{X}^n}$ in the other direction. For the formal definition, we assume without loss of generality that $P_{f(\tilde{X}^n)}(0) \leq \frac{1}{2}$ and set

$$\alpha := \frac{P_{f(\tilde{X}^n)}(0)}{P_{f(\tilde{X}^n)}(1)} \, ,$$

i.e., $\alpha \leq 1$. P_{V^n} and P_{W^n} are then given by

$$P_{V^n}(x^n) := \begin{cases} P_{\tilde{X}^n}(x^n)(1 + \tau) & \text{if } x^n \in f^{-1}(0) \\ P_{\tilde{X}^n}(x^n)(1 - \alpha\tau) & \text{if } x^n \in f^{-1}(1) \end{cases}$$

$$P_{W^n}(x^n) := \begin{cases} P_{\tilde{X}^n}(x^n)(1 - \tau) & \text{if } x^n \in f^{-1}(0) \\ P_{\tilde{X}^n}(x^n)(1 + \alpha\tau) & \text{if } x^n \in f^{-1}(1) \, , \end{cases}$$

where $\tau := \frac{\delta}{4}$. Because

$$\sum_{x^n \in \bar{\mathcal{X}}^n} P_{V^n}(x^n) = \sum_{x^n \in f^{-1}(0)} P_{X^n}(x^n)(1 + \tau) + \sum_{x^n \in f^{-1}(1)} P_{X^n}(x^n)(1 - \alpha\tau)$$

$$= P_{f(X^n)}(0)(1 + \tau) + P_{f(X^n)}(1)(1 - \alpha\tau) = 1 \, ,$$

P_{V^n} and, similarly, P_{W^n}, is indeed a probability distribution.

---

[9] For $z \in \{0, 1\}$, $f^{-1}(z) := \{x \in \bar{\mathcal{X}}^n : f(x) = z\}$ denotes the preimage of z under the mapping f.

We claim that P_{V^n} and P_{W^n} are $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources. To see this, note first that, for any $i \in \{1, \ldots n\}$ and $x^i \in \bar{\mathcal{X}}^i$,

$$(1 - \alpha\tau)P_{\tilde{X}^i}(x^i) \leq P_{V^i}(x^i) \leq (1 + \tau)P_{\tilde{X}^i}(x^i) .$$

Hence, for any $x \in \mathcal{X}$ and $x^{i-1} \in \bar{\mathcal{X}}^{i-1}$,

$$
\begin{aligned}
P_{V_i|V^{i-1}=x^{i-1}}(x) &= \frac{P_{V_i V^{i-1}}(x, x^{i-1})}{P_{V^{i-1}}(x^{i-1})} \\
&\geq \frac{(1 - \alpha\tau)P_{\tilde{X}_i \tilde{X}^{i-1}}(x, x^{i-1})}{(1 + \tau)P_{\tilde{X}^{i-1}}(x^{i-1})} \\
&= \frac{1 - \alpha\tau}{1 + \tau}P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}}(x) \\
&= \frac{1 - \alpha\tau}{1 + \tau}(1 - \tfrac{\delta}{2})Q_{X_i|X^{i-1}=x^{i-1}}(x) .
\end{aligned}
$$

Because $\alpha \leq 1$, we have $P_{V_i|V^{i-1}=x^{i-1}}(x) \geq (1 - \delta)Q_{X_i|X^{i-1}=x^{i-1}}(x)$. Similarly,

$$P_{V_i|V^{i-1}=x^{i-1}}(x) \leq \frac{1 + \tau}{1 - \alpha\tau}P_{\tilde{X}_i|\tilde{X}^{i-1}=x^{i-1}}(x) = \frac{1 + \tau}{1 - \alpha\tau}(1 - \tfrac{\delta}{2})Q_{X_i|X^{i-1}=x^{i-1}}(x)$$

which implies $P_{V_i|V^{i-1}=x^{i-1}}(x) \leq Q_{X_i|X^{i-1}=x^{i-1}}(x)$. Combining these inequalities, we conclude $P_{V_i|V^{i-1}=x^{i-1}} \in \mathcal{P}^\delta(Q_{X_i|X^{i-1}=x^{i-1}})$, i.e., P_{V^n} is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source. A similar computation shows that also the distribution P_{W^n} is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source. Consequently, the probability distributions $P_{V^n Y}$ and $P_{W^n Y}$ defined by $P_{Y|V^n} = Q_{Y|X^n}$ and $P_{Y|W^n} = Q_{Y|X^n}$, respectively, are contained in the set Γ.

Next, we will analyze the behavior of the function g for inputs chosen according to $P_{V^n Y}$ and $P_{W^n Y}$, respectively, and compare it to f. For this, let

$$q_{x^n} := \Pr_{y \leftarrow Q_{Y|X^n=x^n}}[g(x^n, y) = 0] .$$

be the probability that, given some fixed $x^n \in \bar{\mathcal{X}}^n$, the output of g is zero. Because, by definition, $P_{\tilde{Y}|\tilde{X}^n} = P_{Y|V^n} = P_{Y|W^n} = Q_{Y|X^n}$, this quantity does not depend on the distribution we consider, that is,

$$q_{x^n} = P_{g(\tilde{X}^n, \tilde{Y})|\tilde{X}^n=x^n}(0) = P_{g(V^n, Y)|V^n=x^n}(0) = P_{g(W^n, Y)|W^n=x^n}(0) .$$

The probability that the output of f is zero for the distributions P_{V^n} and P_{W^n} can then, obviously, be written as

$$
\begin{aligned}
P_{f(V^n)}(0) &= \sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)(1 + \tau) \\
P_{f(W^n)}(0) &= \sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)(1 - \tau)
\end{aligned}
$$

Similarly, for g, we have

$$P_{g(V^n,Y)}(0) = \sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)(1+\tau)q_{x^n} + \sum_{x^n \in f^{-1}(1)} P_{\tilde{X}^n}(x^n)(1-\alpha\tau)q_{x^n}$$

$$P_{g(W^n,Y)}(0) = \sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)(1-\tau)q_{x^n} + \sum_{x^n \in f^{-1}(1)} P_{\tilde{X}^n}(x^n)(1+\alpha\tau)q_{x^n} .$$

By assumption of the lemma, because, $P_{V^n Y}$ and $P_{W^n Y}$ are contained in the set Γ, the output distributions of f and g must be close, that is,

$$|P_{f(V^n)}(0) - P_{g(V^n,Y)}(0)| < \frac{\varepsilon}{2}$$

$$|P_{f(W^n)}(0) - P_{g(W^n,Y)}(0)| < \frac{\varepsilon}{2} ,$$

and hence

$$(P_{f(V^n)}(0) - P_{g(V^n,Y)}(0)) - (P_{f(W^n)}(0) - P_{g(W^n,Y)}(0)) < \varepsilon .$$

Replacing these probabilities by the above expressions leads to

$$\sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)2\tau(1-q_{x^n}) + \sum_{x^n \in f^{-1}(1)} P_{\tilde{X}^n}(x^n)2\alpha\tau q_{x^n} < \varepsilon . \qquad (3)$$

Note that this imposes some restrictions on the possible values of q_{x^n}. Roughly speaking, if f maps a certain input x^n to 0, then the probability $1 - q_{x^n}$ that g maps x^n to 1 must be small. In fact, as we shall see, (3) implies a bound on the probability that the outputs of f and g are different.

With the definition $p_{z,w} := P_{f(\tilde{X}^n)g(\tilde{X}^n,Y)}(z,w)$, for $(z,w) \in \{0,1\}^2$ and using again the assumption of the lemma,

$$|p_{0,1} - p_{1,0}| = |(p_{0,0} + p_{0,1}) - (p_{0,0} + p_{1,0})|$$

$$= |P_{f(\tilde{X}^n)}(0) - P_{g(\tilde{X}^n,\tilde{Y})}(0)| < \frac{\varepsilon}{2} ,$$

hence,

$$\Pr_{(x^n,y) \leftarrow P_{\tilde{X}^n Y}}[f(x^n) \neq g(x^n,y)] = p_{0,1}+p_{1,0} \leq p_{0,1}+p_{0,1}+|p_{1,0}-p_{0,1}| < 2p_{0,1}+\frac{\varepsilon}{2} . \qquad (4)$$

Using (3) and the fact that the second sum is nonnegative, we get an upper bound for $p_{0,1}$, that is,

$$p_{0,1} = \sum_{x^n \in \bar{\mathcal{X}}^n} P_{\tilde{X}^n}(x^n)P_{f(\tilde{X}^n)|\tilde{X}^n=x^n}(0)P_{g(\tilde{X}^n,\tilde{Y})|\tilde{X}^n=x^n}(1)$$

$$= \sum_{x^n \in f^{-1}(0)} P_{\tilde{X}^n}(x^n)(1-q_{x^n}) < \frac{\varepsilon}{2\tau} = \frac{2\varepsilon}{\delta} .$$

Combining this with (4), we conclude

$$\Pr_{(x^n,y)\leftarrow P_{\tilde{X}^n Y}}[f(x^n) \neq g(x^n,y)] < \frac{4\varepsilon}{\delta} + \frac{\varepsilon}{2} \leq \frac{5\varepsilon}{\delta}\ ,$$

which proves the lemma for the binary case where $\mathcal{Z} = \{0,1\}$.

To deduce the statement for arbitrary sets $\mathcal{Z}$, consider an (injective) encoding function c which maps each element $z \in \mathcal{Z}$ to an m-tuple $(c_1(z),\ldots,c_m(z))$. Since the norm $|\cdot|_1$ can only decrease when applying a function, the assumption of the lemma implies that, for all probability distributions $P_{X^n Y} \in \Gamma$,

$$|P_{f_k(X^n)} - P_{g_k(X^n,Y)}|_1 < \varepsilon\ ,$$

where, $f_k := c_k \circ f$ and $g_k := c_k \circ g$, for any $k \in \{1,\ldots,m\}$. The assertion then follows from the binary version of the lemma and the union bound.

As an immediate corollary, we deduce the impossibility of randomness extraction. The impossibility of other cryptographic tasks such as encryption follows exactly the same argument as in [DOPS04].

Corollary 1. *Let f be a function from $\bar{\mathcal{X}}^n$ to $\{0,1\}$. For any $i \in \{1,\ldots,n\}$, let $Q_{X_i|X^{i-1}}$ be a channel from $\bar{\mathcal{X}}^{i-1}$ to $\mathcal{X}$ and let $\delta \geq 0$. Then there exists a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source P_{X^n} such that*

$$|P_{f(X^n)} - P_U|_1 \geq \frac{\delta}{10}\ ,$$

where P_U is the uniform distribution on $\{0,1\}$.

Proof. Assume by contradiction that, for any $(\delta, \{Q_{X_i|X^{i-1}}\})$-source P_{X^n},

$$|P_{f(X^n)} - P_U|_1 < \frac{\delta}{10}\ .$$

Let g be the function on $\mathcal{X}^n \times \{0,1\}$ defined by $g(x^n, u) := u$. Then, for any probability distribution $P_{X^n U} = P_{X^n} \times P_U$, where P_{X^n} is a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source, we have

$$|P_{f(X^n)} - P_{g(X^n,U)}|_1 < \frac{\delta}{10}\ .$$

Lemma 1 thus implies that there exists a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source $P_{\tilde{X}^n}$ with

$$\Pr_{(x^n,u)\leftarrow P_{\tilde{X}^n} \times P_U}[f(x^n) \neq g(x^n,u)] < \frac{1}{2}\ ,$$

that is,

$$\Pr_{(x^n,u)\leftarrow P_{\tilde{X}^n} \times P_U}[f(x^n) \neq u] < \frac{1}{2}\ .$$

This is a contradiction because P_U is the uniform distribution on $\{0,1\}$.

References

[AL93]   Miklós Ajtai and Nathal Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[BBR88]   Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM J. Comput.*, 17(2):210–229, 1988.

[BIW04]   Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness from few independent sources. In *Proc. 45th FOCS*, 2004.

[Blu86]   M. Blum. Independent unbiased coin flips from a correlated biased source—a finite state Markov chain. *Combinatorica*, 6(2):97–108, 1986.

[CDH+00]   Ran Canetti, Yevgeniy Dodis, Shai Halevi, Eyal Kushilevitz, and Amit Sahai. Exposure-resilient functions and all-or-nothing transforms. In *Proc. EUROCRYPT '00*, pages 453–469, 2000.

[CG88]   Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, 1988.

[CGH+85]   Benny Chor, Oded Goldreich, Johan Håstad, Joel Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem of t-resilient functions. In *Proc. 26th FOCS*, pages 396–407. IEEE, 1985.

[CH74]   John F. Clauser and Michael A. Horne, Experimental consequences of objective local theories, Phys. Rev. D 10, 526–535, 1974.

[DEOR04]   Yevgeniy Dodis, Ariel Elbaz, Roberto Oliveira, and Ran Raz. Improved randomness extraction from two independent sources. In *Proc. RANDOM '04*, 2004.

[DO03]   Yevgeniy Dodis and Roberto Oliveira. On extracting private randomness over a public channel. In *Proc. RANDOM '03*, pages 252–263, 2003.

[DOPS04]   Yevgeniy Dodis, Shien Jin Ong, Manoj Prabhakaran, and Amit Sahai. On the (im)possibility of cryptography with imperfect randomness. In *Proc. FOCS '04*, pages 196–205, 2004.

[DSS01]   Yevgeniy Dodis, Amit Sahai, and Adam Smith. On perfect and adaptive security in exposure-resilient cryptography. In *Proc. EUROCRYPT '01*, pages 301–324, 2001.

[DS02]   Yevgeniy Dodis and Joel Spencer. On the (non)Universality of the One-Time Pad. In *Proc. FOCS '02*, pages 376–385, 2002.

[Eli72]   Peter Elias. The efficient construction of an unbiased random sequence. *Ann. Math. Stat.*, 43(2):865–870, 1972.

[KZ03]   Jess Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. In *Proc. 35th FOCS*, pages 92–101, 2003.

[LLS89]   D. Lichtenstein, N. Linial, and M. Saks. Some extremal problems arising from discrete control processes. *Combinatorica*, 9(3):269–287, 1989.

[MSS83]   Trevor W. Marshall, Emilio Santos, and Franco Selleri, Local realism has not been refuted by atomic-cascade experiments, Phys. Lett. A 98, 5–9, 1983.

[Mas02]   Serge Massar, Non locality, closing the detection loophole and communication complexity, Phys. Rev. A 65, 032121, 2002.

[MP91]   J. L. McInnes and B. Pinkas. On the impossibility of private key cryptography with weakly random keys. In *Proc. CRYPTO '90*, pages 421–436, 1991.

[NC00]   M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[NZ96]   Noam Nisan and David Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43–52, 1996.

[RVW04]   Omer Reingold, Salil Vadhan, and Avi Wigderson. A note on extracting randomness from Santha-Vazirani sources. Unpublished manuscript, 2004.

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

[TV00]   Luca Trevisan and Salil Vadhan. Extracting randomness from samplable distributions. In *Proc. 41st FOCS*, pages 32–42, 2000.

[Vaz87a]   U. V. Vazirani. Strong communication complexity or generating quasi-random sequences from two communicating semi-random sources. *Combinatorica*, 7(4):375–392, 1987.

[Vaz87b]   Umesh V. Vazirani. Efficiency considerations in using semi-random sources. In *Proc. 19th STOC*, pages 160–168, 1987.

[von51]   John von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards, Applied Mathematics Series*, 12:36–38, 1951.

[VV85]   Umesh V. Vazirani and Vijay V. Vazirani. Random polynomial time is equal to slightly-random polynomial time. In *Proc. 26th FOCS*, pages 417–428, 1985.

[Zuc96]   David Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4/5):367–391, 1996.

A   Relation to Santha-Vazirani sources

Let P_{X^n} be a $(\delta, \{Q_{X_i|X^{i-1}}\})$-source, for some $\delta \geq 0$ and channels $Q_{X_i|X^{i-1}}$. It is easy to verify that, if $\delta \leq \frac{1}{|\mathcal{X}|}$ then the entropy of the ith output X_i conditioned on any value of the previous outputs $X_1, \ldots, X_{i-1}$ is lower bounded by the entropy of $Q_{X_i|X^{i-1}=x^{i-1}}$, i.e.,

$$H(X_i|X^{i-1} = x^{i-1}) = H(P_{X_i|X^{i-1}=x^{i-1}}) \geq H(Q_{X_i|X^{i-1}=x^{i-1}}), \qquad (5)$$

for any $x^{i-1} \in \bar{\mathcal{X}}^{i-1}$. This holds with respect to any "reasonable" entropy measure H, as, for instance, the Shannon entropy, the min-entropy, or, more generally, the Rényi entropy of order α, for any $\alpha \in [0, \infty]$.

It is thus not surprising that $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources are at least as useful as Santha-Vazirani sources. More precisely, Lemma 2 below states that, for any α, there exist channels $Q_{X_i|X^{i-1}}$ and a deterministic[10] strategy γ which allows to simulate an α-SV source from any $(\delta, \{Q_{X_i|X^{i-1}}\})$-source, for $\delta = 1 - 2\alpha$. Hence, any impossibility result for $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources also holds for α-SV sources.

Lemma 2. *For any $\delta \geq 0$, there exist channels $Q_{X_i|X^{i-1}}$, for $i \in \{1, \ldots, n\}$, and a function γ such that the following holds: Let P_{X^n} be an arbitrary $(\delta, \{Q_{X_i|X^{i-1}}\})$-source. Then the probability distribution P_{Y^n} defined by $Y_i := \gamma(X_i)$, for $i \in \{1, \ldots n\}$, is an α-SV source, for $\alpha = \frac{1-\delta}{2}$.*

Proof. Let P_δ be the binary probability distribution with $P_\delta(0) = \frac{1+\delta}{2}$. For any $i \in \{1, \ldots, n\}$, let the channel $Q_{X_i|X^{i-1}}$ be defined by $Q_{X_i|X^{i-1}=x^{i-1}} := P_\delta$. Additionally, let γ be the function on $\{0, 1, \bot\}$ defined by

$$\gamma(x) := \begin{cases} x & \text{if } x \in \{0,1\} \\ 1 & \text{if } x = \bot . \end{cases}$$

---

[10] Note that any probabilistic strategy would require additional (perfect) randomness.

It is easy to verify that, for any $i \in \{1, \ldots, n\}$ and $x^{i-1} \in \{0, 1, \perp\}^{i-1}$,

$$P_{\gamma(X_i)|X^{i-1}=x^{i-1}}(0) \leq P_\delta(0) = \frac{1+\delta}{2} = 1 - \alpha$$

$$P_{\gamma(X_i)|X^{i-1}=x^{i-1}}(0) \geq P_\delta(0)(1-\delta) = \frac{1+\delta}{2}(1-\delta) \geq \alpha ,$$

i.e., $P_{\gamma(X_i)|X^{i-1}=x^{i-1}}(0) \in [\alpha, 1-\alpha]$. By convexity, it follows that $P_{Y_i|Y^{i-1}=y^{i-1}}(0) \in [\alpha, 1-\alpha]$, for any $y^{i-1} \in \{0,1\}^{i-1}$, which concludes the proof.

Note that the converse of Lemma 2 is not true, i.e., Santha-Vazirani sources are generally weaker than $(\delta, \{Q_{X_i|X^{i-1}}\})$-sources. To see this, let, e.g., for any $i \in \{1, \ldots, n\}$, $Q_{X_i|X^{i-1}}$ be the channel defined by the uniform distribution over $\mathcal{X} := \{0, 1\}$, i.e., $Q_{X_i|X^{i-1}=x^{i-1}}(0) = \frac{1}{2}$, for all $x^{i-1} \in \bar{\mathcal{X}}^{i-1}$. It follows from (5) that the entropy of any $(\delta, \{Q_{X_i|X^{i-1}}\})$-source $P_{X_1 \cdots X_n}$ is at least n, for any small enough $\delta \geq 0$. On the other hand, the entropy of an α-SV source $P_{Y_1 \cdots Y_n}$, for any $\alpha \neq \frac{1}{2}$, is generally smaller than n. As the entropy of a random variable can only decrease when applying a (deterministic) function, the values $(Y_1, \ldots, Y_n)$ cannot be used to simulate $(X_1, \ldots, X_n)$.