Randomness Extractors: Complexity and Relaxations

Paul Fermé
Advisers: Mario Berta, Omar Fawzi

Introduction

Randomness extractors are essential for both classical and quantum cryptography: encryption protocols use randomness, and their security relies on the accuracy of that randomness, which must be uniform and unpredictable. However, for the pseudo-random sources we are able to generate, these properties are not fulfilled. The role of randomness extractors is, given an imperfect pseudo-random source, possibly correlated with an adversary, to extract its randomness into a smaller and (almost) perfect random source.

During that internship, we have studied the computation of the error of randomness extractors. First, we have shown that this problem is coNP-complete for extractors. Then, we have studied *quantum-proof* extractors: in that case, the adversary is possibly quantum. These extractors are not yet fully understood, but some differences with the regular ones are known. One open problem about these extractors is the status of random functions, which are known to be good regular extractors. We have studied a semidefinite relaxation of those extractors, which is easier to manipulate: indeed, we were able to show that random functions are bad according to that relaxation, which gives interesting insights about the quantum case.

Contents

Part I
Preliminaries

1 Randomness extractors

Remark. This part is based on [12] (when no explicit reference is given).

1.1 Definition

As presented in the introduction, randomness extractors are functions that, given a source "containing some randomness", output a source which is close to the uniform distribution. Before looking at the definition of randomness extractors, we must define what is the "quantity of randomness" included in a source, also called quantity of *information*, and a distance between sources, ie. random variables.

First, let us define this notion of distance:

Definition 1.1.1 (total variation)**.** The total variation between two random variables X and Y over the same set U is defined by:

$$\Delta(X,Y) := \max_{T \subseteq U} |\Pr(X \in T) - \Pr(Y \in T)| \tag{1}$$

We say that X is ϵ-close to Y if $\Delta(X,Y) \leq \epsilon$.

The set T that maximizes this function can be seen as the best strategy distinguishing the two random variables. Let us give some useful properties of that distance:

Property 1.1.1.

$$\begin{aligned}
&(1) \quad \Delta \text{ is a distance} \\
&(2) \quad 0 \leq \Delta(X,Y) \leq 1 \\
&(3) \quad \Delta(X,Y) = \max_{T \subseteq U} \Pr(X \in T) - \Pr(Y \in T) \\
&(4) \quad \Delta(X,Y) = \frac{1}{2}|X - Y|_1 := \frac{1}{2} \sum_{u \in U} |\Pr(X = u) - \Pr(Y = u)|
\end{aligned} \tag{2}$$

Let us now focus on the quantification of information contained in a random variable, via the notion of *entropy*, essential to define how good is a "pseudo-random" source.

Definition 1.1.2 (min-entropy)**.** Let X be a random variable. The *min-entropy* of X is defined by:

$$H_{\min}(X) := -\log \|X\|_\infty, \text{ where } \|X\|_\infty := \max_x \Pr(X = x) \tag{3}$$

$\|X\|_\infty$ can be interpreted as the winning probability of the best strategy guessing the value of X without any information: simply bet on one value x which happens with the biggest probability for X. Other definitions of entropy exist, but we are interested in that interpretation of information: the bigger $H_{\min}(X)$, the harder it is to guess the value of X.

Now, we have all the necessary concepts to define randomness extractors:

Definition 1.1.3 (randomness extractors)**.** A (k, ϵ)-randomness extractor is a function $\text{Ext} : N \times D \to M$ such that for all X on N with $H_{\min}(X) \geq k$ (called *k-sources*), $\text{Ext}(X, U_D)$ is ϵ-close to U_M (where U_S is a new and independent uniform distribution over S)

Why have we added a seed D? In fact, it can be shown that without a seed, there is no (k, ϵ)-randomness extractor for $k < n$. The point now is to keep D small compared to M. A stronger definition of extractors asks that, even if the seed D is public, the source still looks uniform:

Definition 1.1.4 (strong randomness extractors)**.** A (k, ϵ)-strong randomness extractor is a function $\text{Ext} : N \times D \to M$ such that $\text{Ext'}(X, Y) := (\text{Ext}(X, Y), Y)$ is a (k, ϵ)-randomness extractor.

Equivalently, it must be such that for all X on N with $H_{\min}(X) \geq k$:

$$\frac{1}{2D} \sum_{s \in D, y \in M} \left| \sum_{x \in N : Ext(x,s)=y} \Pr(X = x) - \frac{1}{M} \right| \leq \epsilon \tag{4}$$

In the rest of this report, except when it is explicitly written, we will only consider strong randomness extractors, which we simply call randomness extractors.

Finally, another object that could interest us is classical-proof randomness extractors, ie. extractors which are still good even if an adversary has access to a source C which is correlated to the input source X. This leads to the following definitions:

Definition 1.1.5 (classical-proof randomness extractors)**.** The *conditional min-entropy* of X given C is defined by $H_{\min}(X|C) := -\log(p_{\text{guess}}(X|C))$, where $p_{\text{guess}}(X|C)$ is the best strategy of guessing the value of X given C. We have $p_{\text{guess}}(X|C) = \sum_c \Pr(C = c) \max_x \Pr(X = x|C = c)$.

We then say that $\text{Ext} : N \times D \to M$ is a (k, ϵ)-classical-proof randomness extractor if, for all X and C such that $H_{\min}(X|C) \geq k$:

$$\frac{1}{2D} \sum_{s \in D, y \in M} \sum_c \left| \sum_{x \in N : Ext(x,s)=y} \Pr((X, C) = (x, c)) - \frac{1}{M} \Pr(C = c) \right| \leq \epsilon \tag{5}$$

We can easily see that if C is independent of X, then we get back the usual definition of randomness extractors. So those classical-proof extractors are particular cases of randomness extractors. As shown in [7], there exists a statement of the same kind for the other direction:

Theorem 1.1.2. *If Ext is a (k, ϵ)-randomness extractor, then it is $(k+log\frac{1}{\epsilon}, 2\epsilon)$-classical-proof randomness extractor.*

Thus, it is not really interesting to study those particular extractors, which are almost the same as the regular ones.

1.2 Main properties

A first useful property of randomness extractors is that we don't need to test them on all possible sources, but only on *flat* sources:

Definition 1.2.1 (Flat k-sources)**.** Flat k-sources are uniform distribution on a set $S \subseteq N$, with $|S| = 2^k$ (when $2^k \in \mathbb{N}$).

A consequence of that definition is that every k-source is a convex combination of flat k-sources (see [12]). Since the error $E(X)$ of an extractor on a random If X is such that $\Pr(X = x) = \sum p_i \Pr(X_i = x)$ and $\sum_i p_i = 1$, then the error $E(.)$ of an extractor satisfies $E(X) = \sum_i p_i E(X_i)$: hence, the maximum error is obtained on flat k-sources. Thus, we may restrict our attention to flat sources when testing the error of extractors.

Then, let us give some examples of concrete randomness extractors. In that part, we will consider that $N = \{0,1\}^n$, $M = \{0,1\}^m$ and $D = \{0,1\}^d$.

Hash functions More precisely, in our setting, we focus on 2-universal functions:

Definition 1.2.2 (2-universal functions)**.** A family of function $\{f_s : N \to M\}_{s \in D}$ is 2-universal if, for all $x \neq x' \in N$, $\Pr_{s \sim \mathcal{U}(D)}[f_s(x) = f_s(x')] \leq \frac{1}{M}$.

With that definition, it is proved in [12] that it is a good extractor:

Theorem 1.2.1 (Leftover Hash Lemma)**.** *If $\{f_s : N \to M\}_{s \in D}$ is 2-universal, with $m = k - 2\log(1/\epsilon)$, then $\{f_s\}$ is a (strong) $(k, \frac{\epsilon}{2})$-extractor.*

However, it was shown in [9] that, in order to have such a 2-universal family, we need $d = \Omega(n)$.

Random functions The issue of the hash functions is that they need a large seed. One can show (for instance in [8]) that taking *random functions*, ie. choosing D functions f_s independently at random among functions from N to M, with the parameters given in equation (6) leads to a (k, ϵ)-extractor with high probability.

$$m = k - 2\log(1/\epsilon) - \mathcal{O}(1) \text{ and } d = \log(n-k) + 2\log(1/\epsilon) + \mathcal{O}(1) \qquad (6)$$

In fact, in [8], it was shown that these parameters cannot be improved by any extractor (except for additive constant).

Trevisan construction The issue of random functions is that it is not an *explicit* construction of extractors: we need a concise and efficient description of that extractor in order to be able to compute it on large sources. Luca Trevisan, in [11], was the first to find an explicit construction of extractors with $d = \mathcal{O}(\log n)$. Now a lot of progress has been done in that domain (for a review, see [12]).

2 Quantum-proof randomness extractors

2.1 Quantum information

Remark. This part is largely taken from the presentation of quantum information in [2]. Furthermore, most of this presentation is not crucial to understand the results of this report, since we will quickly leave the quantum world and focus on the SDP relaxation of quantum-proof randomness extractors.

In quantum theory, a system is described by an inner-product space, that we denote here by letters like N, M, Q. Note that we use the same symbol Q to label the system, the corresponding inner-product space and also the dimension of the space. Let $\mathrm{Mat}_Q(S)$ be the vector space of $Q \times Q$ matrices with entries in S. Whenever S is not specified, it is assumed to be the set of complex numbers $\mathbb{C}$, i.e., we write $\mathrm{Mat}_Q := \mathrm{Mat}_Q(\mathbb{C})$. The state of a system is defined by a positive semidefinite operator ρ_Q with trace 1 acting on Q, which is also called a *density matrix*. The set of states on system Q is denoted by $S(Q) \subseteq \mathrm{Mat}_Q(\mathbb{C})$. The inner-product space of a composite system QN is given by the tensor product of the inner-product spaces $QN := Q \otimes N$. From a joint state $\rho_{QN} \in S(QN)$, we can obtain marginals on the system Q by performing a partial trace of the N system $\rho_Q := \mathrm{Tr}_N[\rho_{QN}]$. The state ρ_{QN} of a system QN is called quantum-classical (with respect to some basis) if it can be written as $\rho_{QN} = \sum_x \rho(x) \otimes \Pr(X = x) |x\rangle \langle x|$ for some basis $\{|x\rangle\}$ of N, some random variable X on N, and some density matrices $\rho(x)$ acting on Q.

To measure the distance between two states, we use the trace norm $\|A\|_1 := \mathrm{Tr}[\sqrt{A^* A}]$, where A^* is the conjugate transpose of A. In the special case when A is diagonal, $\|A\|_1$ becomes the familiar l_1 norm of the diagonal entries. Moreover, the Hilbert-Schmidt norm is defined as $\|A\|_2 := \sqrt{\mathrm{Tr}[A^* A]}$, and when A is diagonal this becomes the usual l_2 norm. Another important norm we use is the operator norm, or the largest singular value of A, denoted by $\|A\|_\infty$. When A is diagonal, this corresponds to the familiar l_∞ norm of the diagonal entries. The conditional min-entropy of X given Q is used to quantify the uncertainty of the source X given the system Q. The conditional min-entropy is defined as:

$$H_{\min}(X|Q)_\rho := -\log \min_{\sigma_Q \in S(Q)} \|(\sigma_Q^{-1/2} \otimes I_N)\rho_{QN}(\sigma_Q^{-1/2} \otimes I_N)\|_\infty \qquad (7)$$

2.2 Definition and main properties

Intuitively, a quantum-proof extractor is the same thing as a classical-proof extractor, but instead of having our source X correlated to a random variable C, instead we speak of a quantum classical system $\rho_{QN} = \sum_x \rho(x) \otimes \Pr(X = x) |x\rangle \langle x|$, with $\rho(x)$ being a density matrix on Q, ie. describing the quantum state correlated to x. We won't explain in more details the definition of the entropy, which can be interpreted as being the logarithm of the inverse of the success probability of guessing the value of X (see [6]), as it was done before for regular and classical-proof extractors. By analogy with classical-proof extractors, let us give now the definition of quantum-proof ones:

Definition 2.2.1 (quantum-proof randomness extractors)**.** We say that $\text{Ext} : N \times D \to M$ is a (k, ϵ)-quantum-proof randomness extractor if, for all quantum classical system $\rho_{QN} = \sum_x \rho(x) \otimes \Pr(X = x) |x\rangle \langle x|$ such that $H_{\min}(X|Q)_\rho \geq k$:

$$\frac{1}{2D} \sum_{s \in D, y \in M} \left\| \sum_{x \in N : Ext(x,s) = y} \Pr(X = x)\rho(x) - \frac{1}{M} \sum_x \Pr(X = x)\rho(x) \right\|_1 \leq \epsilon \quad (8)$$

If all $\rho(x)$ are diagonal, ie. they represent only classical correlations, then we get the definition of classical-proof extractors, $\rho(x)$ being the vector $(\Pr(Q = q|X = x))_q$.

The first question that comes now is the existence of a clear gap between this class of extractors and the regular ones (which doesn't exist when we just add classical correlation): this was solved in [5], where it is shown that there exist a regular extractor of error ϵ, with the quantum error being at least $\Omega(m\epsilon)$. This implies that it is crucial to study quantum-proof extractors directly, in order to understand what are precisely the constructions that still work in the presence of a quantum adversary.

In [10], it was shown that 2-universal functions are quantum-proof extractors with the same parameters as regular extractors. In [4], it was shown that Trevisan extractors are quantum-proof, thus giving an example of short-seeded quantum-proof extractors. However, the status of random functions is still open, and will be the center of our focus on quantum extractors.

3 Semidefinite programming

Semidefinite programs (SDP) are a large class of optimization problems that can be solved in polynomial time. Semidefinite programming has been extensively used in various contexts in quantum information: we could say semidefinite programming is to quantum computational problems what linear programming is to combinatorial problems. Like with linear programming, it can be used to relax some complex problems, and some good properties of that class helps to understand the original problem.

We use a formulation of semidefinite programs called vector programs. For some fixed values $\alpha_{x,x'}, \beta_{x,x',k}$ and γ_k, the optimization program can be written as follows:

$$\begin{aligned} \text{maximize} \quad & \sum_{x,x'} \alpha_{x,x'} \vec{a}_x \cdot \vec{a}_{x'} \\ \text{subject to} \quad & \sum_{x,x'} \beta_{x,x',k} \vec{a}_x \cdot \vec{a}_{x'} \leq \gamma_k \text{ for all } k \end{aligned} \tag{9}$$

The optimization is made over all vectors $\vec{a}_x$ of finite dimension satisfying the constraints stated before.

Part II
Results

4 SRE is coNP-complete

We show in this section that the following problem is coNP-complete:

Definition 4.0.2 (SRE (Strong Randomness Extractor))**.** Given a function $\text{Ext} : N \times D \to M$, $k \in \mathbb{R}$, and $\epsilon \in]0,1[$, decide if Ext is a strong (k,ϵ)-extractor:

YES : The error of Ext is lower or equal than ϵ for all k-sources

NO : The error of Ext is greater than ϵ on a particular k-source

To achieve this, we show that its complementary $\overline{\text{SRE}}$ is NP-complete: from now on, we only focus on this problem.

The outline of the proof is the following: we'll first translate this problem into an equivalent graph problem (SEG), and look at some particular instances of that problem, which are the incidence graphs of a bipartite graph problem (SIG), which can be shown to be NP-hard via a reduction from the bipartite densest subgraph problem (BDS).

In this proof of NP-hardness, we only use instances with ϵ close to 1. So, this only tells us that checking that extractors are not *really* bad is NP-hard, but the status of the problem when ϵ is small (for instance, $\epsilon < \frac{1}{2}$) remains unknown.

4.1 Graph formulation of $\overline{\text{SRE}}$

Definition 4.1.1 (FG (Functional Graph))**.** A bipartite graph $G = (N \mid M \times D, E)$ is a functional graph if $\forall (n,d) \in N \times D, \{(n,(m,d)) : m \in M\} \cap E$ is a singleton.

Definition 4.1.2 (Graph associated to an extractor)**.** The associated graph of $\text{Ext} : N \times D \to M$ is $G_{\text{Ext}} = (N \mid M \times D, E)$ with $E = \{(n,(m,d)) \mid \text{Ext}(n,d) = m\}$.

Property 4.1.1. *$\phi : Ext \mapsto G_{Ext}$ is a bijection between $M^{N \times D}$ and the set of FGs.*

Definition 4.1.3 (SEG (Strong Extractor Graph))**.** SEG is the set of functional graph $G = (N \mid M \times D, E)$, $k \in \mathbb{N}$ and $\epsilon \in]0,1[$ such that:

$$\boxed{\exists S \subseteq N \text{ with } S = k, \exists T \subseteq M \times D,\ \frac{E(S,T)}{SD} - \frac{T}{MD} > \epsilon} \tag{10}$$

where $E(S,T)$ is the set of edges between S and T.

Remark. In a slight abuse of notations, we denote by A both the set A and its size $card(A)$.

Lemma 4.1.2. $(G, k, \epsilon) \in SEG \Leftrightarrow (\phi^{-1}(G), \log(k), \epsilon) \in \overline{SRE}$

Proof. We take the definition of total variation without absolute value; then, the value of the extractor can be written as:

$$\max\left\{\frac{1}{SD}\sum_{x\in S,(y,s)\in T}\delta_{f_s(x)=y} - \frac{T}{MD} : S \subseteq N, S = k, T \subseteq M \times D\right\} \tag{11}$$

Indeed, the value of our extractor is maximized by flat $\log(k)$-sources, ie. sources that are uniform over sets of size $2^{\log(k)} = k$, which leads to our previous statement.

Thus, by definition of the extractor $\phi^{-1}(G)$, $\sum_{x\in S,(y,s)\in T}\delta_{f_s(x)=y} = E(S,T)$, which concludes the proof. □

4.2 Incidence graph point of view

Definition 4.2.1 (SIG (Strong Incidence Graph))**.** SIG is the set of balanced bipartite graph $G = (V_0 \mid V_1, E)$, $k \in \mathbb{N}$ and $\epsilon \in]0, 1[$ such that:

$$\boxed{\exists S \subseteq E \text{ with } S = k, \exists T \subseteq V_0 \cup V_1,\ \frac{\sum_{v\in T} deg_S(v)}{2k} - \frac{T}{2M} > \epsilon} \tag{12}$$

where $deg_S(v)$ is the degree of the vertex v in $(V_0 \mid V_1, S)$ and $M = |V_0| = |V_1|$.

In order to state the link between SIG and SEG, we need to define what is an incidence graph:

Definition 4.2.2 (Incidence Graph)**.** The incidence of the graph $G = (V, E)$ is $G_{\text{inc}} := (V \cup E, \{(v, e) \in V \times E : v \in e\})$.

An example of this transformation is represented in figure 1. Let us now state precisely the link between SIG and SGE.

Lemma 4.2.1. $(G, k, \epsilon) \in SIG \Leftrightarrow (G_{inc}, k, \epsilon) \in SGE$.

Proof. First, what does G_{inc} look like? Its vertices are the edges and the vertices of G (ie. $E \cup V_0 \cup V_1 \cong E \mid V \times \{0, 1\}$) and its edges are $\{(v, e) \in (V_0 \cup V_1) \times E : v \in e\}$. A consequence of this definition is that G_{inc} is a functional graph with $D = \{0, 1\}$: each $e \in E$ is exactly linked with one $u \in V_0$ and one $v \in V_1$, so it is bipartite and satisfies the constraint of functional graphs, as we can see in the example in figure 1.

Then, notice that the sets S and T chosen in G can directly be interpreted in G_{inc} the way we would want to. Since $D = 2$ here, we only have to show that $\sum_{v\in T} deg_S(v) = E(S, T)$. This follows from the fact that $deg_S(v)$, being the number of edges in S linked to v, can be interpreted in G_{inc} as the number of edges between S and v. □

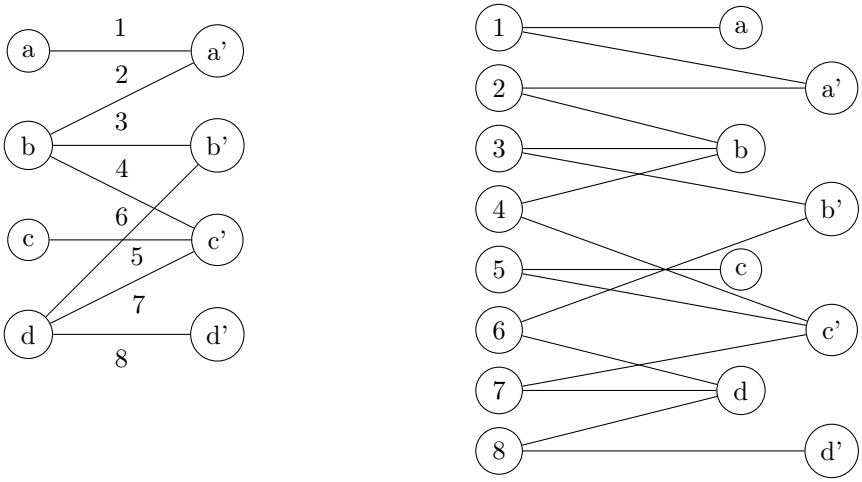

Figure 1: On the left, the original balanced bipartite graph; on the right, its incidence graph, which is a functional graph with $D = 2$.

Now, let us simplify two details of that definition: the strict inequality and the choice of T.

Concerning the inequality, since $\frac{\sum_{v\in T} deg_S(v)}{2k} - \frac{T}{2M} > \epsilon \Leftrightarrow M\sum_{v\in T} deg_S(v) - kT > 2kM\epsilon$, and the left part of the latter inequality being an integer, we thus can replace, in SIG, $>$ by $\geq$ and keeping an equivalent problem.

Concerning the choice of T, the problem is solved by the following property:

Property 4.2.2. *For a given S of size k, obj$(T) := \frac{\sum_{v\in T} deg_S(v)}{2k} - \frac{T}{2M}$ is maximized for $T = T_S := \{v : deg_S(v) > \frac{k}{M}\}$.*

Proof. Since obj is linear (ie. $\text{obj}(T \sqcup U) = \text{obj}(T) + \text{obj}(U)$), $T = \{v : \text{obj}(\{v\}) > 0\}$ maximizes obj. But, $\text{obj}(\{v\}) > 0 \Leftrightarrow deg_S(v) > \frac{k}{M}$, so $T_S = T$. □

This leads to the definition of the problem we will use in the reduction:

Definition 4.2.3 (SIG*)**.** Given a bipartite graph $G = (V_0 \mid V_1, E)$ with $|V_0| = |V_1| = M$, $k \in \mathbb{N}$, $\epsilon \in]0, 1[$, decide if the following statement is true:

$$\boxed{\exists S \subseteq E \text{ with } S = k,\ \frac{\sum_{v\in T_S} deg_S(v)}{2k} - \frac{T_S}{2M} \geq \epsilon} \tag{13}$$

Lemma 4.2.3. *SIG is equivalent to SIG*.*

4.3 Reduction from the bipartite densest subgraph problem

With what has been done in the previous section, in order to show the coNP-completeness of SRE, we only have to show the NP-hardness of SIG* (SRE $\in$

coNP follows from the definition). In order to achieve that, we do a reduction from the Bipartite Densest Subgraph, which has been proven to be NP-complete in [3].

Definition 4.3.1 (BDS (Bipartite Densest Subgraph))**.** Given a bipartite graph $G = (V_0 \mid V_1, E)$, K , $N \in \mathbb{N}$, is there $V \subseteq V_0 \cup V_1$ with $V = K$ and $E_V \geq N$? (where $E_V = \{(u, v) \in E : u \in V \text{ and } v \in V\}$)

Theorem 4.3.1. *SIG* is NP-hard.*

Corollary 4.3.2. *SRE is coNP-complete.*

Proof. Let $(G = (V_0 \mid V_1, E), K, N \in \mathbb{N})$ an instance of BDS. We choose as instance of SIG*: $(G' = (V_0' \mid V_1', E), k = N, \epsilon = 1 - \frac{K}{2M})$ where V_b' is such that $V_b \subseteq V_b'$ and $|V_0'| = |V_1'| > |E|(\geq N)$, simply by adding isolated vertices to the original bipartite graph. Thus, we have a correct instance (it is a balanced bipartite graph) which is of polynomial size in the size of the instance of BDS.

Let us show that $(G, K, N) \in \text{BDS} \Leftrightarrow (G', k, \epsilon) \in \text{SIG*}$.

First of all, for any S of size k, $T_S = \{v : deg_S(v) > \frac{k}{M}\} = \{v : deg_S(v) > 0\}$ since $\frac{k}{M} < 1$. Thus $\sum_{v \in T_S} deg_S(v) = 2S = 2k$, since all edges in S are taken in T_S. Thus, in that setting, $obj(T_S) = 1 - \frac{T_S}{2M}$.

Now, let us prove the necessary condition. Assume there exists $V \subseteq V_0 \cup V_1$ with $V = K$ and $E_V \geq N$. Then, in particular, there exists $S \subseteq E_V$ of exact cardinality $N = k$. Furthermore, vertices in T_S are selected among neighbor vertices of edges in $S \subseteq E_V$, but $E_V \subseteq V^2$, so $T_S \subseteq V$. Thus:

$$obj(T_S) = 1 - \frac{T_S}{2M} \geq 1 - \frac{V}{2M} = \epsilon$$

Let us now prove the sufficient condition. Assume there exists $S \subseteq E$ of cardinality k such that $obj(T_S) \geq \epsilon$. Since $T_S = \{v : deg_S(v) > 0\}$ and elements of $V_b' - V_b$ are isolated, we have $T_S \subseteq V_0 \cup V_1$. But $obj(T_S) \geq \epsilon$ and $S = k$, so:

$$1 - \frac{T_S}{2M} \geq 1 - \frac{K}{2M}$$

hence $T_S \leq K$. We take thus V by adding some vertices to T_S in order to have $V = K$. Then, $S \subseteq N_{T_S} \subseteq N_V$. Thus, $N_V \geq S = k = N$ which concludes the proof. □

5 SDP relaxations

5.1 Quadratic programs

As shown in [2], we can see the definitions of both classical and quantum randomness extractors as optimization programs. We take the notation Ext $:= \{f_s : N \to M, s \in D\}$. First, for classical randomness extractors, we get the following program:

$$\begin{aligned} \mathrm{C}(\mathrm{Ext}, k) := \text{maximize} \quad & \frac{1}{2D} \sum_{s,y,x} \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) p(x) \beta_{s,y} \\ \text{subject to} \quad & 0 \leq p(x) \leq 2^{-k} \\ & \sum_x p(x) = 1 \\ & -1 \leq \beta_{s,y} \leq 1 \end{aligned} \tag{14}$$

Property 5.1.1. *Ext is a (k, ϵ)-extractor if and only if C(Ext, k) $\leq \epsilon$.*

Proof. Indeed, $p := (p(x))_x$ is ranging over all k-sources, and the optimal choice for $\beta_{s,y}$ given p leads to the value $\frac{1}{2D} \sum_{s,y} |\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right) p(x)|$, which is the usual definition of randomness extractors. □

We have shown in the previous section that the problem of deciding if $\mathrm{C}(\mathrm{Ext}, k) \leq \epsilon$ is coNP-complete.

In a similar way, we can write an optimization program for the quantum version of randomness extractors:

$$\begin{aligned} \mathrm{Q}(\mathrm{Ext}, k) := \text{maximize} \quad & \frac{1}{2D} \sum_{s,y,x} \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \mathrm{Tr}[\rho(x) B_{s,y}] \\ \text{subject to} \quad & 0 \leq \rho(x) \leq 2^{-k} \sigma \\ & \sum_x \mathrm{Tr}[\rho(x)] = 1 \\ & \mathrm{Tr}[\sigma] = 1 \\ & \|B_{s,y}\|_\infty \leq 1 \end{aligned} \tag{15}$$

Property 5.1.2. *Ext is a quantum-proof (k, ϵ)-extractor if and only if Q(Ext, k) $\leq \epsilon$.*

Proof. The condition on $\rho(x)$ and σ ensures that $\sum_x \rho(x) \otimes |x\rangle \langle x|$ has conditional min-entropy at least k. Furthermore, given $\rho(x)$, the optimal choice for $B_{s,y}$ leads to the value of $\frac{1}{2D} \sum_{s,y} \| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right) \rho(x) \|_1$. □

Here, since the maximization is made over $\rho(x)$ of arbitrary (though finite) dimension, we do not even know if $\mathrm{Q}(\mathrm{Ext}, k)$ is computable. This was one of the motivations that lead to look at semidefinite relaxations of this problem.

5.2 An improved semidefinite relaxation

A first *ad hoc* relaxation of the optimization programs (15) (and (14)), was proposed in [2]:

$$\begin{aligned} \mathrm{SDP}_{\mathrm{old}}(\mathrm{Ext},k) := \text{maximize} \quad & \frac{1}{2D}\sum_{s,y,x}\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\cdot\vec{b}_{s,y} \\ \text{subject to} \quad & 0\le \vec{a}_x\cdot\vec{a}_{x'}\le 2^{-k}q(x) \\ & 0\le q(x)\le 2^{-k} \\ & \sum_x q(x)=1 \\ & \sum_{x,x'}\vec{a}_x\cdot\vec{a}_{x'}\le 1 \\ & \|\vec{b}_{s,y}\|_2\le 1 \end{aligned} \tag{16}$$

Given $\vec{a}_x$, the best choice of $\vec{b}_{s,y}$ here leads to the value $\frac{1}{2D}\sum_{s,y}\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\|_2$

Theorem 5.2.1. *For any Ext and k, we have:*

$$C(Ext,\ k)\le SDP_{old}(Ext,k)$$

$$Q(Ext,\ k)\le\sqrt{2}SDP_{old}(Ext,k)$$

Although it had some good properties, as shown in [2], the value of this relaxation is unbounded on extractors satisfying some property typically satisfied by function taken at random, whereas they are good classical extractors.

Later on, taking a larger point of view on general quantum bilinear optimization problems, a systematic way to find relaxations was proposed in [1], which then leads (more or less) to the SDP we have been focusing on:

$$\begin{aligned} \mathrm{SDP}(\mathrm{Ext},k) := \text{maximize} \quad & \frac{1}{2D}\sum_{s,y,x}\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\cdot\vec{b}_{s,y} \\ \text{subject to} \quad & 0\le \vec{a}_x\cdot\vec{a}_{x'}\le 2^{-k}\vec{a}_x\cdot\vec{\mathbb{1}} \\ & 0\le \vec{a}_x\cdot\vec{\mathbb{1}}\le 2^{-k} \\ & \sum_x \vec{a}_x\cdot\vec{\mathbb{1}}=1 \\ & \sum_{x'}\vec{a}_x\cdot\vec{a}_{x'}\le\vec{a}_x\cdot\vec{\mathbb{1}} \\ & \vec{b}_{s,y}\cdot\vec{b}_{s,y}\le\vec{\mathbb{1}}\cdot\vec{\mathbb{1}} \\ & |\vec{a}_x\cdot\vec{b}_{s,y}|\le\vec{a}_x\cdot\vec{\mathbb{1}} \\ & \vec{\mathbb{1}}\cdot\vec{\mathbb{1}}=2 \end{aligned} \tag{17}$$

Theorem 5.2.2. *For any Ext and k, Q(Ext, k) ≤ SDP(Ext, k) ≤ $\sqrt{2}$SDP$_{old}$(Ext, k)*

Proof. First, let us focus on the left inequality. Let $\rho(x)$, σ and $B_{s,y}$ a solution maximizing Q(Ext, k). Let us define $\overline{\rho} := \sum_x \rho(x)$ and $\omega := \overline{\rho} + \sigma$. Thus we take $\vec{a}_x$ as the list of entries of the matrix $\omega^{-1/4}\rho(x)\omega^{-1/4}$, $\vec{b}_{s,y}$ as the list of entries of the matrix $\omega^{1/4}B_{s,y}\omega^{1/4}$ and $\vec{\mathbb{1}}$ as the list of entries of the matrix $\omega^{1/2}$. Thus:

$$\begin{aligned}
&\vec{a}_x \cdot \vec{a}_{x'} = \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')] \\
&\vec{a}_x \cdot \vec{b}_{s,y} = \mathrm{Tr}[\rho(x)B_{s,y}] \\
&\vec{a}_x \cdot \vec{\mathbb{1}} = \mathrm{Tr}[\rho(x)] \\
&\vec{b}_{s,y} \cdot \vec{b}_{s,y} = \mathrm{Tr}[\omega^{1/2}B_{s,y}\omega^{1/2}B_{s,y}] \\
&\vec{\mathbb{1}} \cdot \vec{\mathbb{1}} = \mathrm{Tr}[\omega] = 2
\end{aligned} \tag{18}$$

As a consequence of this, we already see that the objective function $\sum_{s,y,x} \left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x \cdot \vec{b}_{s,y} = \sum_{s,y,x} \left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\mathrm{Tr}[\rho(x)B_{s,y}] = \mathrm{Q(Ext}, k)$. We now only have to show that those vectors satisfy all the constraints.

$$\begin{aligned}
\vec{a}_x \cdot \vec{a}_{x'} &= \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')] \leq \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}2^{-k}\sigma] \\
&\leq 2^{-k}\mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}2^{-k}\omega] \leq 2^{-k}\mathrm{Tr}[\rho(x)] = 2^{-k}\vec{a}_x \cdot \vec{\mathbb{1}}
\end{aligned} \tag{19}$$

Furthermore, since $\omega^{-1/4}\rho(x)\omega^{-1/4}$ is positive semidefinite, then $\vec{a}_x \cdot \vec{a}_{x'} = \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')] \geq 0$. The inequalities depending only on $\vec{a}_x \cdot \vec{\mathbb{1}}$ follow from the fact that it is equal to $\mathrm{Tr}[\rho(x)]$. Then:

$$\begin{aligned}
\sum_{x'} \vec{a}_x \cdot \vec{a}_{x'} &= \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\overline{\rho}] \leq \mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\omega] \\
&\leq \vec{a}_x \cdot \vec{\mathbb{1}}
\end{aligned} \tag{20}$$

Now, since $\|B_{s,y}\|_\infty \leq 1$, which equivalently means that $-I \leq B_{s,y} \leq I$, we have that:

$$\vec{b}_{s,y} \cdot \vec{b}_{s,y} = \mathrm{Tr}[\omega^{1/2}B_{s,y}\omega^{1/2}B_{s,y}] \leq \mathrm{Tr}[\omega] = \vec{\mathbb{1}} \cdot \vec{\mathbb{1}} \tag{21}$$

and also that:

$$\pm\vec{a}_x \cdot \vec{b}_{s,y} = \pm\mathrm{Tr}[\rho(x)B_{s,y}] \leq \mathrm{Tr}[\rho(x)] = \vec{a}_x \cdot \vec{\mathbb{1}} \tag{22}$$

which ends the first part of the proof.

For the right inequality, just by keeping the same $\vec{a}_x$, taking $q(x) = \vec{a}_x \cdot \vec{\mathbb{1}}$ and dividing $\vec{b}_{s,y}$ by $\sqrt{2}$, we get a solution of the old SDP up to a factor $\sqrt{2}$. □

One of the strange parts of this relaxation is that $\vec{\mathbb{1}} \cdot \vec{\mathbb{1}} = 2$. This is caused by the trick of taking $\omega := \overline{\rho} + \sigma$ in the proof. If we have taken simply $\omega := \sigma$, then $\vec{\mathbb{1}} \cdot \vec{\mathbb{1}} = 1$ and all the constraints except $\sum_{x'} \vec{a}_x \cdot \vec{a}_{x'} \leq \vec{a}_x \cdot \vec{\mathbb{1}}$ would be satisfied.

Let us call this smoothed relaxation SDP*, which is exactly the relaxation SDP where we have removed the constraint $\sum_{x'} \vec{a}_x \cdot \vec{a}_{x'} \leq \vec{a}_x \cdot \vec{\mathbb{1}}$ and have $\vec{\mathbb{1}} \cdot \vec{\mathbb{1}} = 1$. It turns out that this particular relaxation has some good properties concerning hash functions, since it gives the best known bound on the error for both quantum and classical extractors:

Theorem 5.2.3. *If $\{f_s\}$ is 2-universal and $M = \epsilon^2 2^k$, then SDP*$(\{f_s\}, k) \leq \frac{\epsilon}{2}$*

Proof.

$$\begin{aligned}
\text{SDP*}(\{f_s\}, k) &= \frac{1}{2D} \sum_{s,y,x} \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \vec{a}_x \cdot \vec{b}_{s,y} \\
&\leq \frac{1}{2D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \vec{a}_x \right\|_2 \text{ since } \|\vec{b}_{s,y}\|_2 = \vec{b}_{s,y} \cdot \vec{b}_{s,y} \leq 1 \\
&\leq \frac{1}{2} \sqrt{M} \sqrt{\frac{1}{D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \vec{a}_x \right\|_2^2} \text{ by Cauchy-Schwarz}
\end{aligned} \tag{23}$$

We will now show that the term under the last square root is smaller than 2^{-k}, which is enough to conclude.

$$\begin{aligned}
&\frac{1}{D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \vec{a}_x \right\|_2^2 \\
&= \frac{1}{D} \sum_{s,y} \left(\sum_x \delta_{f_s(x)=y} \vec{a}_x \right) \cdot \left(\sum_{x'} \delta_{f_s(x')=y} \vec{a}_{x'} \right) - \frac{1}{DM^2} \sum_{s,y} \left(\sum_x \vec{a}_x \right) \cdot \left(\sum_{x'} \vec{a}_{x'} \right) \\
&= \frac{1}{D} \sum_s \sum_{x,x'} \delta_{f_s(x)=f_s(x')} \vec{a}_x \cdot \vec{a}_{x'} - \frac{1}{M} \sum_{x,x'} \vec{a}_x \cdot \vec{a}_{x'}
\end{aligned} \tag{24}$$

The first term, when $x \neq x'$, satisfies $\frac{1}{D} \sum_s \delta_{f_s(x)=f_s(x')} \leq \frac{1}{M}$, since $\{f_s\}$ is 2-universal, so that part is canceled by the second term. Thus, we have that it is smaller than $\sum_x \vec{a}_x \cdot \vec{a}_x \leq \sum_x 2^{-k} \vec{a}_x \cdot \vec{\mathbb{1}} \leq 2^{-k}$.

□

A new property fulfilled by this relaxation is that its value is bound:

Property 5.2.4. *For any Ext and k, SDP(Ext, k) $\leq \frac{M-1}{M}$*

Remark. It is exactly the same bound as usual extractors: it is the value obtained for constant functions. See the appendix for a proof.

The natural question that comes now is the status of typical random functions, ie. extractors having some property fulfilled by random functions w.h.p.. Their SDP value is still large, whereas they are known to be good classical extractors:

Theorem 5.2.5. *If Ext is taken uniformly at random with $D = \Omega(\log(N)) \leq M = \mathcal{O}(N/\log(N))$ and $k \leq \log(2N/M)$, then w.h.p.*

$$SDP(Ext, k) \geq \frac{M-1}{2M} \tag{25}$$

For instance, if $N = \{0,1\}^n$, $M = \{0,1\}^m$ and $D = \{0,1\}^d$ with $m = n/4, k = n/2, d = \mathcal{O}(\log n)$ and Ext is chosen at random, then w.h.p. we have (as told in [2]):

$$\mathrm{C}(\mathrm{Ext}, k) \leq \frac{1}{n} \text{ and } \mathrm{SDP}(\mathrm{Ext}, k) \geq \frac{M-1}{2M} \geq \frac{1}{4} \tag{26}$$

In the proof, we will use the following lemma (proven in the appendix):

Lemma 5.2.6. *If $\{f_s\}$ is taken uniformly over $M^{N \times D}$ with $D = \Omega(\log(N))$ and $M = \mathcal{O}(N/\log(N))$, then w.h.p.*

$$\forall x, \sum_{s,x'} \delta_{f_s(x)=f_s(x')} \leq 2\frac{DN}{M} \tag{27}$$

Proof. We take:

$$\begin{aligned}
\vec{a}_x &= a^{-1/2} \sum_{s,y} \delta_{f_s(x)=y} \left|s\right\rangle \left|y\right\rangle \text{ where } a = 2\frac{DN^2}{M} \\
\vec{b}_{s,y} &= \sqrt{\frac{2D}{M}} \left|s\right\rangle \left|y\right\rangle \\
\vec{\mathbb{1}} &= \sqrt{\frac{2}{DM}} \sum_{s,y} \left|s\right\rangle \left|y\right\rangle
\end{aligned} \tag{28}$$

We can then see that all constraints are satisfied and the value of the SDP for those vectors is $\frac{M-1}{2M}$.

Indeed, we have:

$$\begin{aligned}
\vec{a}_x \cdot \vec{a}_{x'} &= \frac{1}{a} \sum_s \delta_{f_s(x)=f_s(x')} \\
\vec{a}_x \cdot \vec{b}_{s,y} &= \frac{1}{N} \delta_{f_s(x)=y} \\
\vec{a}_x \cdot \vec{\mathbb{1}} &= \frac{1}{DN} \sum_{s,y} \delta_{f_s(x)=y} = \frac{1}{N} \\
\vec{b}_{s,y} \cdot \vec{b}_{s,y} &= 2\frac{D}{M} \leq 2 = \vec{\mathbb{1}} \cdot \vec{\mathbb{1}}
\end{aligned} \tag{29}$$

and thus

$$
\begin{aligned}
&0 \leq \vec{a}_x \cdot \vec{a}_{x'} \leq \frac{D}{a} = \frac{M}{2N}\frac{1}{N} \leq 2^{-k}\frac{1}{N} \\
&0 \leq \vec{a}_x \cdot \vec{\mathbb{1}} = \frac{1}{N} \leq 2^{-k} \\
&\sum_x \vec{a}_x \cdot \vec{\mathbb{1}} = 1 \\
&|\vec{a}_x \cdot \vec{b}_{s,y}| = \frac{1}{N}\delta_{f_s(x)=y} \leq \frac{1}{N} = \vec{a}_x \cdot \vec{\mathbb{1}}
\end{aligned}
\tag{30}
$$

In order to prove the last constraint, we use the lemma 5.2.6:

$$
\sum_{x'} \vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{a}\sum_{s,x'} \delta_{f_s(x)=f_s(x')} \leq 2\frac{DN}{aM} = \vec{a}_x \cdot \vec{\mathbb{1}} \tag{31}
$$

Finally, the value of the SDP for that admissible solution is:

$$
\begin{aligned}
\frac{1}{2D}\sum_{s,y,x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x \cdot \vec{b}_{s,y} \quad &= \frac{1}{2DN}\sum_{s,y,x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\delta_{f_s(x)=y} \\
&= \frac{1}{2DN}\sum_{s,x}\frac{M-1}{M} = \frac{M-1}{2M}
\end{aligned}
\tag{32}
$$

□

5.3 Some open questions on this relaxation

An important remark on the previous relaxation is the following: if we want only to relax the program C(Ext, k), then we can have $\vec{\mathbb{1}} \cdot \vec{\mathbb{1}} = 1$. To show that it is indeed a relaxation, we only have to choose:

$$
\begin{aligned}
&\vec{a}_x = p(x) \\
&\vec{b}_{s,y} = \beta_{s,y} \\
&\vec{\mathbb{1}} = 1
\end{aligned}
\tag{33}
$$

Then, our particular counter-example showing that random functions have a big relaxation value, if we normalize $\vec{\mathbb{1}}$ and $\vec{b}_{s,y}$ and we take, in $\vec{a}_x, a = \frac{DN^2}{M}$ (the simplest way to normalize $\vec{\mathbb{1}}$ and keep the same value for $\vec{a}_x \cdot \vec{\mathbb{1}}$) does not work. Indeed, we have now $\vec{a}_x \cdot \vec{\mathbb{1}} = \frac{DN}{aM}$, but:

$$
\sum_{x'} \vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{a}\sum_{s,x'} \delta_{f_s(x)=f_s(x')} \tag{34}
$$

and

$$
\mathbb{E}\left[\frac{1}{a}\sum_{s,x'} \delta_{f_s(x)=f_s(x')}\right] = \frac{D(N+M-1)}{aM} > \frac{DN}{aM} \tag{35}
$$

In lemma 5.2.6, since we assumed that $M = \mathcal{O}(N/\log N)$, we could hence get a bound with a factor 2, whereas here it is not possible.

First, note that the constraint that fails here is exactly the constraint we were able to get by taking $\vec{\mathbb{1}}\cdot\vec{\mathbb{1}} = 2$ instead of $\vec{\mathbb{1}}\cdot\vec{\mathbb{1}} = 1$ in the proof of theorem 5.2.2. Indeed, this could be a clue that, random functions may have a small value with the relaxation of $\mathrm{C}(\mathrm{Ext}, k)$, or even that there is a fundamental difference between classical and quantum extractors (remind that it is unknown if random functions are good quantum extractors).

An argument against that conjecture is the following: this norm issue can be solved if we allow $\sum_x \vec{a}_x\cdot\vec{\mathbb{1}} \leq 1$ instead of $\sum_x \vec{a}_x\cdot\vec{\mathbb{1}} = 1$. Indeed, if we divide every vector by $\sqrt{2}$, we still get a solution and $\vec{\mathbb{1}}\cdot\vec{\mathbb{1}} = 1$ (the objective value is then twice smaller only). Furthermore, if we call $\mathrm{C}_{\leq}(\mathrm{Ext}, k)$ (resp. $\mathrm{Q}_{\leq}(\mathrm{Ext}, k)$) the version of $\mathrm{C}(\mathrm{Ext}, k)$ (resp. $\mathrm{Q}(\mathrm{Ext}, k)$) where $\sum_x p(x) = 1$ is replaced by $\sum_x p(x) \leq 1$ (resp. $\sum_x \mathrm{Tr}[\rho(x)] = 1$ is replaced by $\sum_x \mathrm{Tr}[\rho(x)] \leq 1$), then the following property holds:

Property 5.3.1. *$C_{\leq}(Ext, k) \leq 2C(Ext, k-1)$ and $Q_{\leq}(Ext, k) \leq 2Q(Ext, k-1)$*

However, extending the proof of that property in a natural way does not work for the relaxation. So the question of bounding $\mathrm{SDP}_{\leq}(\mathrm{Ext}, k)$ (where $\sum_x \vec{a}_x\cdot\vec{\mathbb{1}} \leq 1$ instead of $\sum_x \vec{a}_x\cdot\vec{\mathbb{1}} = 1$) by $\mathrm{SDP}(\mathrm{Ext}, l)$, with l close to k remains open.

Proof. First, let $p(x)$ a solution maximizing $\mathrm{C}_{\leq}(\mathrm{Ext}, k)$. Let us call $S := \sum_x p(x) \leq 1$ take $p'(x) := p(x) + \frac{1-S}{N}$. Then, since $\frac{1-S}{N} \leq \frac{1}{N} \leq 2^{-k}$, we have $p'(x) \leq 2^{-(k-1)}$.

Furthermore, $\sum_x p'(x) = S + N\frac{1-S}{N} = 1$. Finally, we have:

$$\begin{aligned}
\mathrm{C}_{\leq}(\mathrm{Ext}, k) \quad &= \frac{1}{2D}\sum_{s,y}\left|\sum_x\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)p(x)\right| \\
&= \frac{1}{2D}\sum_{s,y}\left|\sum_x\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\left(p'(x) - \frac{1-S}{N}\right)\right| \\
&\leq \frac{1}{2D}\sum_{s,y}\left|\sum_x\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)p'(x)\right| + |1-S|\frac{1}{2D}\sum_{s,y}\left|\sum_x\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\frac{1}{N}\right| \\
&\leq (2-S)\mathrm{C}(\mathrm{Ext}, k-1) \text{ since both } p'(x) \text{ and } \frac{1}{N} \text{ are admissible solutions} \\
&\leq 2\mathrm{C}(\mathrm{Ext}, k-1)
\end{aligned} \tag{36}$$

which concludes the proof for classical extractors.

The proof is the same for the quantum case, where we take $\rho'(x) := \rho(x) + \frac{1-\sum_x \mathrm{Tr}[\rho(x)]}{N}\sigma$. □

However, the natural extension to this proof in the relaxation would be to take $\vec{a'}_x := \vec{a}_x + \frac{1-\sum_x \vec{a}_x\cdot\mathbb{1}}{N}\frac{\vec{\mathbb{1}}}{\|\vec{\mathbb{1}}\|}$ but this choice fails on the constraint on

$\sum_{x'} \vec{a'}_x \cdot \vec{a'}_{x'}$. Note that this failure implies $\vec{a'}_x \cdot \vec{a'}_{x'}$, which was not a constraint present in the usual programs: this is essentially the part where we are clearly in the relaxed world of vectors. Hence, there may be a real difference here between the relaxation and the programs, implying that the norm of $\vec{\mathbb{1}}$ could matter.

Conclusion

Although the question of the complexity of the computation of the efficiency of randomness extractors is solved, the status of that problem with small ϵ or k constant is unknown. Also, the question of the hardness of approximation of that problem remains opened.

The relaxation of the quantum problem, giving some good properties, is a framework where random functions fail to be efficient. However, for the higher levels of the hierarchy of relaxations of that problem, as presented in [1], this question has not been studied yet. The answer to that question could help determining the status of random functions as quantum-proof extractors.

This internship allowed me to study a cross-disciplinary topic, between physics and computer science. This was an interesting experience, especially trying to understand a language I was not used to, and fundamental as modern research becomes more and more interdisciplinary.

References

[1] Mario Berta, Omar Fawzi, and Volkher B. Scholz. Quantum bilinear optimization. Preprint, `http://arxiv.org/abs/1506.08810`, 2015.

[2] Mario Berta, Omar Fawzi, and Volkher B. Scholz. Semidefinite programs for randomness extractors. In *Leibniz International Proceedings in Informatics*, 2015.

[3] Derek G Corneil and Yehoshua Perl. Clustering and domination in perfect graphs. *Discrete Applied Mathematics*, 9(1):27–39, 1984.

[4] Anindya De, Christopher Portmann, Thomas Vidick, and Renato Renner. Trevisan's extractor in the presence of quantum side information. *SIAM Journal on Computing*, 41(4):915–940, 2012.

[5] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 516–525, New York, NY, USA, 2007. ACM.

[6] Robert Koenig, Renato Renner, and Christian Schaffner. The operational meaning of min-and max-entropy. *Arxiv preprint ArXiv:0807.1338*, 2008.

[7] Robert T König and Barbara M Terhal. The bounded-storage model in the presence of a quantum adversary. *Information Theory, IEEE Transactions on*, 54(2):749–762, 2008.

[8] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM journal on discrete mathematics*, 13:2000, 2000.

[9] D.R. Stinson. Universal hashing and authentication codes. *Designs, Codes and Cryptography*, 4(3):369–380, 1994.

[10] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing against quantum side information. *Information Theory, IEEE Transactions on*, 57(8):5524–5535, 2011.

[11] Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48:2001, 1999.

[12] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1 3):1–336, 2011.

A Caltech

This internship was performed at Caltech in Pasadena, CA in the United States of America. I was working inside the Quantum Information and Processing group, which was mainly composed of physicists either doing research in quantum computer science topics: quantum algorithms, quantum cryptography, quantum complexity, ..., or some more physical topics like condensed matter. Each week, I have attended a group meeting, where everyone could talk about his progress. There were also a lot of weekly conferences I would try to attend when it was understandable with my computer science background.

I have discussed mainly with Mario Berta and Omar Fawzi, authors of the paper I was working on, but also with other graduate students (with a computer science background) from MIT, ENS Cachan, ...

All along this internship, I studied the basis of quantum mechanics needed for computer science, as well as semidefinite programming and hierarchies of relaxation (like the sum of squares hierarchies).

B Missing proofs

- Property 5.2.4 page 16:

  Proof.

$$\begin{aligned}
\mathrm{SDP}(\mathrm{Ext}, k) =& \frac{1}{2D}\sum_{s,y,x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x\cdot\vec{b}_{s,y} \\
\leq& \frac{1}{2D}\sum_{s,y,x}\left|\delta_{f_s(x)=y} - \frac{1}{M}\right|\vec{a}_x\cdot\vec{\mathbb{1}} \text{ since } |\vec{a}_x\cdot\vec{b}_{s,y}| \leq \vec{a}_x\cdot\vec{\mathbb{1}} \\
=& \frac{1}{2D}\sum_{s,y,x:f_s(x)=y}\left(1-\frac{1}{M}\right)\vec{a}_x\cdot\vec{\mathbb{1}} + \frac{1}{2D}\sum_{s,y,x:f_s(x)\neq y}\frac{1}{M}\vec{a}_x\cdot\vec{\mathbb{1}} \\
=& \frac{1}{2D}\sum_{s,x}\left(1-\frac{1}{M}\right)\vec{a}_x\cdot\vec{\mathbb{1}} + \frac{1}{2D}\sum_{s,x}\frac{M-1}{M}\vec{a}_x\cdot\vec{\mathbb{1}} \\
=& 2\frac{M-1}{2MD}\sum_{s,x}\vec{a}_x\cdot\vec{\mathbb{1}} = \frac{M-1}{M}
\end{aligned} \tag{37}$$

  □

- Lemma 5.2.6 page 17:

  Proof. For $x' \neq x$, $f_s(x)$ and $f_s(x')$ are independently chosen. Thus, for a given x, the boolean random variables $(\delta_{f_s(x)=f_s(x')})_{s,x':x'\neq x}$ are independent, of same expectancy $\frac{1}{M}$. We can apply Chernoff bound on their sum:

$$\Pr\left(\sum_{s,x':x'\neq x}\delta_{f_s(x)=f_s(x')} > (1+\epsilon)\frac{D(N-1)}{M}\right) < \left(\frac{e^\epsilon}{(1+\epsilon)^{(1+\epsilon)}}\right)^{\frac{D(N-1)}{M}} \tag{38}$$

so

$$\Pr\left(\sum_{s,x'}\delta_{f_s(x)=f_s(x')} > \frac{D(N+M-1)}{M}+\epsilon\frac{D(N-1)}{M}\right) < \left(\frac{e^\epsilon}{(1+\epsilon)^{(1+\epsilon)}}\right)^{\frac{D(N-1)}{M}} \tag{39}$$

If we take $\epsilon = \frac{N-M+1}{N-1}$, we get

$$\Pr\left(\sum_{s,x'}\delta_{f_s(x)=f_s(x')} > 2\frac{DN}{M}\right) < \left(\frac{e^\epsilon}{(1+\epsilon)^{(1+\epsilon)}}\right)^{\frac{D(N-1)}{M}} \tag{40}$$

So, taking the union bound for all x, and then taking the complementary, we finally get

$$\Pr\left(\forall x, \sum_{s,x'}\delta_{f_s(x)=f_s(x')} \leq 2\frac{D(N-1)}{M}\right) \geq 1 - N\left(\frac{e^\epsilon}{(1+\epsilon)^{(1+\epsilon)}}\right)^{\frac{D(N-1)}{M}} \tag{41}$$

Thus, since we have $M = \mathcal{O}(N/\log N)$ and $D = \Omega(\log N)$, then:

$$N\left(\frac{e^\epsilon}{(1+\epsilon)^{(1+\epsilon)}}\right)^{\frac{D(N-1)}{M}} \sim N(e/4)^{\frac{D(N-1)}{M}} = N(e/4)^{\Omega(\log^2 N)} \to 0 \text{ when } N \to \infty \tag{42}$$

□