Random numbers certified by Bell's theorem

Reading group summary
Thomas Vidick

Setup. Our goal is to generate classical random bits in a *device-independent* way. This means we have access to an (untrusted) box $\mathcal{B}$ which takes classical inputs $x, y, \ldots$ and produces classical outputs $a, b, \ldots$. We would like to devise a (classical) procedure which has access to $\mathcal{B}$ and is such that

1. *(completeness)* There is an "ideal" box $\mathcal{B}_0$ such that, if our procedure is given access to $\mathcal{B}_0$, then it outputs random bits which are ε-close to uniform (in statistical distance).

2. *(soundness)* While making as few assumptions as possible on the inner workings of a generic box $\mathcal{B}$ that is given to us, we can guarantee that either i) the procedure with access to $\mathcal{B}$ produces random bits which are ε-close to uniform, or ii) the procedure aborts, rejecting $\mathcal{B}$ as malfunctioning.

The precise requirements on the output bits may vary. The weakest requirement is that the random variable Y describing them should be statistically close to uniform. A stronger requirement is that this holds even conditioned on an adversary E who has prepared the box $\mathcal{B}$ used by the procedure. In that case, we will ask for a bound of the type[1] $H_\infty(Y|E) \geq \alpha\, n$ to hold with some confidence[2] $1-\delta$. The adversary E may have classical information (a description of the box's workings) or even quantum information (he may hold a state that is entangled with the box).

Assumptions. We make the following assumptions.

1. The inner workings of the box $\mathcal{B}$ are described by Quantum Mechanics.

2. We have access to a short perfectly random and secure string X, independent of $\mathcal{B}$, to be used as a *seed*. This is a necessary assumption, as if our procedure was deterministic it would be easy to come up with a deterministic $\mathcal{B}$ that is always accepted but creates no randomness at all.

3. We have some way of enforcing a no-signaling condition between the outputs of different boxes (and also between the boxes and the environment). This also seems somewhat necessary (under complexity assumptions), in the sense that, for every input, any local procedure will simply produce a string drawn according to some distribution, and there is no known efficient way to decide whether that string is actually random, or is just pseudo-random.

[1] A bound on the conditional min-entropy suffices to produce bits which are ε-close to uniform: it suffices to keep a side a small (poly-logarithmic) private and secure random string to use it as a seed for an extractor applied on Y. One should use an extractor secure against quantum side information [TRSS10, DPRV09].

[2] The confidence is there to quantify the fact that there is always some small chance that the procedure will not reject a fully deterministic box, and hence output a string which has no entropy at all.

Result. The main result of [PAM+10] (see also [CK11] for some previous work describing similar ideas) is that, for any $\varepsilon, \delta = \Omega(1)$,[3] there exists a procedure which satisfies the completeness and soundness properties described above, requires an initial random string X of $\sqrt{n}$ random bits, and outputs $\Omega(n)$ bits Y which are within statistical distance ε of uniform, with confidence $1 - \delta$.

Proof idea. The general idea is the following. The box $\mathcal{B}$ will be made of two parts $\mathcal{B}_1$ and $\mathcal{B}_2$, in-between which we enforce no-signaling (this is possible by assumption 3.). Each part will take as input a bit $x \in \{0,1\}$ (resp. $y \in \{0,1\}$) and output a single bit $a \in \{0,1\}$ (resp. $b \in \{0,1\}$). Let $p(a,b|x,y)$ be the probability distribution on $\mathcal{B}$'s outputs, for every input x, y. Define

$$I = I(p) = \sum_{xy} (-1)^{x \cdot y} \big(p(0,0|x,y) + p(1,1|x,y) - p(0,1|x,y) - p(1,0|x,y) \big)$$

I is a quantity which can be associated to any probability distribution, and hence by extension to any box $\mathcal{B}$. We will ask for a $\mathcal{B}$ that has as large an I as possible. Within quantum mechanics, the largest I is $2\sqrt{2}$ (this is just Tsirelson's bound on the CHSH inequality), while any deterministic $\mathcal{B}$ has a corresponding $I \leq 2$ (this *is* the CHSH inequality). The proof proceeds in two parts.

1. Show that any box with a $I > 2$ necessarily produces answers with some amount of min-entropy (depending on I), for *any* input (x, y) that may be provided to it. This achieves our goal of randomness expansion: concatenating the box's outputs with its inputs produces a string with strictly larger entropy than what we started with (the inputs).

2. Show that one can produce an accurate estimate of I by using the box many times repeatedly. This lets us reject boxes with too small I.[4] We will be able to produce such an estimate by making queries (x, y) taken according to a biased distribution q, with very low entropy (q associates probability $1/\sqrt{n}$ to three of the possible pairs of inputs, and the remaining probability to the last). Given n accesses to the box, this will be sufficient to estimate I, and hence the output's min-entropy, accurately. The advantage of using such a biased distribution is that generating inputs for n uses of the box only require approximately $\sqrt{n}$ random bits. Moreover, from the point of view of the actual entropy produced, as quantified in 1., the input distribution does *not* matter — only the fact that it lets us compute an accurate estimate of I does.

First step: bounding the output min-entropy as a function of the violation I. By using the fact that for the CHSH inequality (and indeed any inequality with two settings and two outcomes) one can always assume without loss of generality that the underlying state is a convex combination of 2-qubit states (note that here we use our assumptions 1. and 3. to model the box's behavior as applying a tensor product measurement on an underlying entangled state), one can show the following.

Claim 1. *For any distribution p that arises from a box $\mathcal{B}$ as described above, with associated violation I, and for any $x, y \in \{0,1\}$, we have*

$$H_\infty(AB|xy) \geq f(I)$$

where $f(I) = \frac{1}{2}\big(1 + \sqrt{2 - I^2/4}\big)$ for $-4 \leq I \leq 4$.

[3] One can make ε, δ smaller at the cost of worse expansion properties.

[4] Because of the memory property, one has to associate a different value I_i to the box as used in each round i; we will get back to this later.

The strength of this claim is that it states that, *whatever the inputs*, the box's outputs have some min-entropy, which is strictly positive as soon as $I > 2$. Hence, *provided* the box we have access to corresponds to an $I > 2$ (each time it is being used), we are guaranteed that the outputs have some randomness.

Second step: producing an accurate estimate of I. The catch so far is that we do not know what the value of I is, or how to check that it is indeed strictly larger than 2, by making a single query. Moreover, I may vary across the rounds (the device changes with its memory), and we let I_i be the quantity associated with the device at the i-th round. Using the chain rule and the fact that f is convex, one can show that the total outputs produced by n uses of the box, A^n and B^n, have min-entropy

$$H_\infty(A^n B^n | x^n y^n) \geq n\, f\Big(\frac{1}{n}\sum_{i=1}^{n} I_i\Big)$$

and this holds for any choice of questions $x^n, y^n \in \{0,1\}^n$. Hence it remains to get an estimate for the *average* violation $I := \frac{1}{n}\sum_{i=1}^n I_i$. We will use the following unbiased estimator for I_i:

$$\hat{I}_i := \sum_{x,y} (-1)^{xy} \frac{1_{\{(a,b)=(0,0)|x,y\}} + 1_{\{(a,b)=(1,1)|x,y\}} - 1_{\{(a,b)=(0,1)|x,y\}} - 1_{\{(a,b)=(1,0)|x,y\}}}{q(x,y)}$$

where $q(x,y)$ is the probability distribution that we use on questions: we simply count $+1$ for a valid answer, -1 for a wrong answer, and scale by $q(x,y)$ in order to make $\hat{I}_i$ an unbiased estimator: $\mathrm{E}_{(x,y)\sim q,\,(a,b)}\big[\hat{I}_i\big] = I_i$. Note that our specific choice of q implies that $\hat{I}_i$ has large variance, as it can take values as large as $\pm\sqrt{n}$.

To wrap up the argument it is sufficient to show that $\frac{1}{n}\sum_i \hat{I}_i$ is within[5] η of $\frac{1}{n}\sum_i I_i$ with probability at least $1-\delta$. A proof that roughly n uses of the boxes are sufficient to guarantee such an estimate can be found in Appendix A.2 of [PAM+10], and is based on a simple Martingale argument and the use of the Azuma-Hoeffding inequality (the key property being that the box at step i only depends on previous history, but not on anything else). Note that, if the rounds were independent, we'd have that $\hat{I}_i$ has standard deviation $\sqrt{n}$, so we do expect about n repetitions to be necessary.

Open questions. There are two big open questions. The first consists in proving that the procedure described (or some variant of it) can produce random bits even from the point of view of an adversary holding a quantum memory which is entangled with the devices themselves. This is would lead to a universally composable proof, and open the way to the use of this procedure as a building block in any cryptographic protocol. The second is to improve the rate of generation of randomness. Could n bits be produced from only $\log(n)$ bits? Can one "bootstrap" the procedure to produce an infinitely long string of bits from a finite amount of initial random bits?

References

[CK11] Roger Colbeck and Adrian Kent. Private randomness expansion with untrusted devices. *Journal of Physics A: Mathematical and Theoretical*, 44(9):095305, March 2011.

[DPRV09] Anindya De, Christopher Portmann, Renato Renner, and Thomas Vidick. Trevisan's extractor in the presence of quantum side information. *CoRR*, abs/0912.5, 2009.

[5] Here η can be any constant so that, if $I = 2\sqrt{2}$ then $\hat{I}$ is strictly more than 2 with high probability, to make sure we do not reject the "honest" box.

[PAM+10] S Pironio, A Acín, S Massar, A Boyer de la Giroday, D N Matsukevich, P Maunz, S Olmschenk, D Hayes, L Luo, T A Manning, and C Monroe. Random numbers certified by Bell's theorem. *Nature*, 464(7291):1021–4, April 2010.

[TRSS10] Marco Tomamichel, Renato Renner, Christian Schaffner, and Adam Smith. *Leftover Hashing against quantum side information*. IEEE, June 2010.