Pseudorandom generators and the BQP vs. PH problem

(3 page abstract)

Bill Fefferman* Chris Umans†

1 Introduction

Let U_t denote a random variable uniformly distributed on t-bit strings. A *pseudorandom generator* (PRG) is a function $f : \{0,1\}^t \to \{0,1\}^m$ that stretches a short "seed" into a longer output string, with the property that $f(U_t)$ is *computationally indistinguishable* from the uniform distribution.

There is a vast literature constructing PRGs that achieve computational indistinguishability against a wide variety of computational models (e.g. small circuits, small nondeterministic circuits, small branching programs, small constant-depth circuits). These constructions are typically "hardness vs. randomness" tradeoffs in the sense that they make use of a hard function (either unconditionally hard, or hard conditioned on a complexity assumption), and their proof of correctness takes the form of a reduction that transforms a computationally efficient *distinguisher* into an efficient algorithm for the hard function (thereby deriving a contradiction). This transformation entails the use of the *hybrid argument* [GM84, Yao82] which incurs a loss of a factor $1/m$ in going from a distinguisher (with gap ε) to a *predictor* (with advantage ε/m) and from there to an efficient algorithm (with advantage ε/m).

A question that has been raised in the pseudorandomness literature is whether this loss of a factor of $1/m$ can be avoided (for an explicit framing of this question, and a discussion of its motivation, see [BSW03]). In certain settings, the answer is known to be "yes" – when the notion of "efficient" is small PH circuits, or bounded-width branching programs [BSW03]. In this work (see [FU10] for the full paper), we identify a setting in which this question has surprising connections to a central unresolved question in quantum complexity: whether there exists an oracle relative to which BQP is not in the PH.

Our setting is a familiar one: we will work with the ubiquitous Nisan-Wigderson PRG [NW94], against AC_0 circuits, with MAJORITY as its hard function. We need a precise statement for the discussion below, which can be given via two standard definitions:

Definition 1.1 ([NW94])**.** *A set family $\mathcal{D} = \{S_1, S_2, \ldots, S_m\}$ is an (ℓ, p) design if every set in the family has cardinality ℓ, and for all $i \neq j$, $|S_i \cap S_j| \leqslant p$.*

Definition 1.2 ([NW94])**.** *Given a function $f : \{0,1\}^\ell \to \{0,1\}$ and an (ℓ, p) design $\mathcal{D} = \{S_1, S_2, \ldots, S_m\}$ in a universe of size t, the function $NW^f_{\mathcal{D}} : \{0,1\}^t \to \{0,1\}^m$ is given by*

$$NW^f_{\mathcal{D}}(x) = \left(f_1(x_{|S_1}), f_2(x_{|S_2}), f_3(x_{|S_3}), \ldots, f_m(x_{|S_m})\right),$$

where each f_i is the function f with a fixed set of its inputs negated[1]*, and $x_{|S}$ denotes the projection of x to the coordinates in the set S.*

Generally speaking, the function $NW^f_{\mathcal{D}}$ is a PRG against a class of distinguishers as long as f is hard on average for that class of distinguishers. Recall that the majority function on ℓ bits is known to be hard for AC_0: no polynomial-size (or even quasi-polynomial-size), constant-depth circuit can compute majority correctly on more than a $1/2+\widetilde{O}(1/\sqrt{\ell})$ fraction of the inputs [Smo93, Hås87], and this is essentially tight, since the function that simply outputs the first bit of the input is correct on a random input with probability $1/2+\Theta(1/\sqrt{\ell})$. We make the following quantitative conjecture:

Conjecture 1. *Let $\mathcal{D} = \{S_1, S_2, \ldots, S_m\}$ be an $(\ell, O(1))$-design in a universe of size $t \leqslant \mathrm{poly}(\ell)$, with $m \leqslant \mathrm{poly}(\ell)$. Then for every constant-depth circuit of size at most $\exp(\mathrm{poly}\log m)$,*

$$|\Pr[C(U_{t+m}) = 1] - \Pr[C(U_t, NW^{\mathrm{MAJORITY}}_{\mathcal{D}}(U_t)) = 1]| \leqslant o(1).$$

*Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA 91125. Supported by IQI.

†Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA 91125. Supported by NSF CCF-0846991.

[1]The standard setup has each $f_i = f$; we need the additional freedom in this paper for technical reasons. We know of no settings in which this alteration affects the analysis of the NW generator.

In this work we abuse notation and refer to constant depth circuits of size at most $\exp(\text{poly}\log m)$ as "AC_0."

By the standard argument from [NW94, Nis92], a distinguishing circuit C with gap ε can be converted to a *predictor* with advantage ε/m and then a slightly larger circuit that computes MAJORITY with success rate $1/2 + \varepsilon/m$. Thus the above statement is true for $m \ll \sqrt{\ell}$; if the $1/m$ loss from the hybrid argument can be avoided (or reduced), it would be true for m as large as $\text{poly}(\ell)$ (and even larger) as we conjecture is true. In upcoming work [FSUV10], we are able to show that indeed AC_0 cannot distinguish the uniform distribution from our instantiation of the Nisan-Wigderson distribution relative to a design with *completely disjoint* sets (i.e., $p = 0$), which is encouraging (even this seemingly simple case is not trivial). The general case, with the design consisting of "nearly-disjoint" sets, remains open.

This paper contains three main results, which together make Conjecture 1 interesting and worthy of further study:

- We show that our conjecture implies the existence of an oracle relative to which BQP is not in the PH, and would thus resolve a major question in quantum complexity. We are encouraged by the fact that our conjecture is recognizable as a natural question in pseudorandomness that has been previously and independently studied (e.g., in [BSW03]).

  The crucial component in showing that our conjecture is sufficient for the existence of an oracle relative to which BQP is not in the PH, is an explicit construction of unitary matrices whose row-supports form an (ℓ, p)-design. We give such a construction and show how to realize these matrices with small quantum circuits. This is the technical core of the paper.

- We generalize the setting of [Aar10b] (which proposed a so-called *forrelated* distribution as one that is easy to distinguish from uniform by a quantum computer, but possibly hard for AC_0) to a simple framework in which any efficiently quantumly computable unitary U gives rise to a distribution that can be distinguished from uniform by a quantum computer (and Aaronson's setup is recovered by choosing U to be a DFT matrix).

  Together with our construction of explicit unitaries whose row-supports form an (ℓ, p)-design, this framework has the following interesting interpretation: it gives an instantiation of the Nisan-Wigderson generator that can be broken by quantum computers, but not by the relevant modes of classical computation, if Conjecture 1 is true.

  Also of independent interest is the fact the unitaries that form the basis of our quantum algorithms don't seem to resemble the DFT matrices for problems in the Hidden Subgroup framework, or even the few other unitaries used in known quantum algorithms. But they possess natural extremal combinatorial (as opposed to algebraic) properties, and we wonder if they can be useful elsewhere in the quantum realm.

- We show that the "Nisan-Wigderson" distribution $(U_t, NW_{\mathcal{D}}^{\text{MAJORITY}}(U_t))$ is ε-almost k-wise independent, in the sense of Aaronson [Aar10b], whose "GLN conjecture" asserted that all such distributions fool AC_0; a depth-3 counterexample was later found [Aar10a]. Whether all such distributions fool depth-2 AC_0 remains open. A distribution in our general framework (thus efficiently quantumly distinguishable from uniform) that fools depth-2 AC_0 would imply an oracle relative to which BQP is not in AM, a weaker (and still unresolved) version of the BQP vs. PH problem. Thus there are two potential routes to resolving this weaker version of the main problem (the depth-2 version of our conjecture, and the depth-2 version of the GLN conjecture); ours is formally easier, and arguably conceptually easier because its connection to the pseudorandomness literature suggests initial lines of attack.

Finally, since [Aar10b] has shown that the classes SZK and BPP_{path} require exponentially many queries to distinguish ε-almost k-wise independent distributions from uniform, our constructions *unconditionally* yield oracles relative to which BQP does not lie in either of these classes (and MA as well, since $MA \subseteq BPP_{\text{path}}$), just as Aaronson's construction does.

2 Techniques

In this section we briefly discuss the techniques we use for each of the main results listed above.

Showing that our NW distribution is ε-almost k-wise independent. We prove that whenever $\mathcal{D}$ is an (ℓ, p) design in a universe of size t, the random variable $(U_t, NW_{\mathcal{D}}^{\text{MAJORITY}}(U_t))$ is $O(pk^2/\sqrt{\ell})$-almost k-wise independent, for $k < o(\ell^{1/4}p^{-1/2})$. The relevant definition of almost-k-wise independence is inherited from [Aar10b]. Recall that this

property of our distribution is the technical basis of the SZK and BPP_{path} results, as well as the connections to the depth-2 GLN conjecture.

This statement amounts to the assertion that after conditioning on the value of up to $k-1$ coordinates, the bias (away from $1/2$) of any specified k-th coordinate is at most $O(pk/\sqrt{\ell})$. This is an easy calculation when the conditioned coordinates all lie among the first t coordinates (since the k-th coordinate is either completely independent, if it lies among the first t coordinates, or else it is MAJORITY applied to a subset of ℓ of the first t coordinates, of which up to $k-1$ may be fixed). In the actual proof, when some conditioned coordinates lie *outside* the first t coordinates (which would otherwise be difficult to analyze), we use the following simple trick to reduce to the easy case: we replace conditioning on coordinate $t+i$ with conditioning on *all* of the coordinates in set S_i of the (ℓ, p)-design (which determine it). Since at most p of these can affect the bias of the k-th coordinate, we are back in the easy case with up to $p(k-1)$ bits fixed instead of $(k-1)$.

Showing that our conjecture is sufficient to resolve the BQP vs. PH question. We will find it convenient to speak almost exclusively about the "scaled down" version of the problem, which is equivalent via the well-known connection between PH and AC_0. In it, the goal is to design a promise problem (rather than an oracle) that lies in (promise)-BQLOGTIME but not (promise)-AC_0.

In order to show that our conjecture is sufficient to imply an oracle relative to which BQP is not in the PH, we need to discuss the quantum part of the argument. Conjecture 1 implies that the NW generator with certain parameters fools AC_0, which is one part of the overall argument. The other part is to exhibit a BQLOGTIME algorithm that "breaks" this instantiation of the NW generator. Generalizing [Aar10b], our quantum algorithm[2] will receive a random string $x \in \{+1,-1\}^t$ (which should be thought of as the input to the NW generator) as the first half of its input, and as the second half of its input, *either*

1. a second random string in $\{+1,-1\}^t$, *or*
2. a string containing the *signs* of a unitary U (with entries in $\{0,1,-1\}$) applied to x.

The algorithm distinguishes the two cases (roughly) by querying x into the phases, applying U, multiplying the second string into the phases, and measuring in the Hadamard basis.

Note that in case (2), each coordinate of the second string is the sign of a $+1/-1$ weighted sum of certain coordinates of x; i.e., it computes MAJORITY (with a fixed pattern of inputs negated) on this subset of the coordinate of x. Thus, if we can construct a unitary U whose row-supports form an (ℓ, p) design $\mathcal{D}$ in a universe of size t, then case (2) will be the distribution $(U_t, NW_{\mathcal{D}}^{\text{MAJORITY}}(U_t))$, and case (1) will be the uniform distribution. The parameters of this instantiation of the NW generator will be such that Conjecture 1 implies that it fools AC_0. Our task becomes to construct such a unitary U.

Note that it is *not* possible to simply take an existing (ℓ, p) design (random, or other explicit constructions that appear in the literature [NW94, HR03]) and attach $+/-$ signs to the elements of the sets so as to make their characteristic vectors pairwise orthogonal, which is what is needed for them to come from the rows of a unitary U. On the other hand we have a different setting of the parameters in mind than usual: we want p to be unusually small (a constant), but the number of sets in the design is also unusually small (only $\mathrm{poly}(\ell)$ instead of $\exp(\ell)$). For these parameters we manage to obtain the required (ℓ, p) design using a geometric construction, in which the sets are the characteristic vectors of pairs of lines in an affine plane. The strong symmetries in this construction allow us to assign $+/-$ signs to the elements of each set to achieve pairwise orthogonality of their characteristic vectors. In fact these set systems have only $t/2$ (rather than t) sets in them, so the resulting unitaries will have the required properties only among half of their rows, but a small modification of the distribution given to the quantum algorithm in case (2) above can handle this without difficulty.

We give a *local decomposition* of these unitaries, which is necessary to have an *efficient* quantum algorithm. This is the most technically involved part of the paper. We also describe a modification of our construction that is *extremal* in the sense that it optimizes all relevant parameters simultaneously: *all* rows of the unitary participate, we have $p \leqslant 2$, and $t \leqslant \ell^2$. This is not required for our results, but it is aesthetically pleasing. We have been unable to find a local decomposition that would enable us to actually use this construction as the basis of an efficient quantum algorithm, and we leave finding such a decomposition as an intriguing open problem.

[2] We ignore normalization factors in this discussion.

References

[Aar10a] S. Aaronson. A counterexample to the Generalized Linial-Nisan Conjecture. In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, number 109, 2010.

[Aar10b] Scott Aaronson. BQP and the polynomial hierarchy. In Leonard J. Schulman, editor, *STOC*, pages 141–150. ACM, 2010.

[BSW03] B. Barak, R. Shaltiel, and A. Wigderson. Computational analogues of entropy. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *RANDOM-APPROX*, volume 2764 of *Lecture Notes in Computer Science*, pages 200–215. Springer, 2003.

[FSUV10] B. Fefferman, R. Shaltiel, C. Umans, and E. Viola. On beating the hybrid argument. Submitted, 2010.

[FU10] B. Fefferman and C. Umans. Pseudorandom generators and the BQP vs. PH problem. Available at `http://www.cs.caltech.edu/~umans/papers/FU10.pdf`, 2010.

[GM84] S. Goldwasser and S. Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[Hås87] Johan Håstad. *Computational limitations of small-depth circuits*. MIT Press, 1987.

[HR03] T. Hartman and R. Raz. On the distribution of the number of roots of polynomials and explicit weak designs. *Random Struct. Algorithms*, 23(3):235–263, 2003.

[Nis92] N. Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

[NW94] N. Nisan and A. Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.

[Smo93] R. Smolensky. On representations by low-degree polynomials. In *FOCS*, pages 130–138. IEEE, 1993.

[Yao82] A. Yao. Theory and applications of trapdoor functions (extended abstract). In *FOCS*, pages 80–91. IEEE, 1982.