Texts in Computational Complexity: A proof of Toda's Theorem

Oded Goldreich
Department of Computer Science and Applied Mathematics
Weizmann Institute of Science, Rehovot, ISRAEL.

November 12, 2005

Preliminaries. We denote by $\mathcal{PC}$ the calss of search problems that correspond to $\mathcal{NP}$; that is, $R \in \mathcal{PC}$ if there exists a polynomial p such that for every $(x,y) \in R$ it holds that $|y| \leq p(|x|)$ and membership in R can be decided in polynomial-time. We refer extensively to the standard proof of the hardness of unique solution instances (a.k.a the Valiant-Vazirani Theorem [7]). See further notes at the end of this text.

We will use small-bias generators (see [3, 1] and notes at the end of this text) as well as the following simple characterization of the levels of the Polynomial-time Hierarchy ($\mathcal{PH}$).

Proposition 1 *The set S is in Σ_{k+1} if and only if there exists a polynomial p and a set $S' \in \Pi_k$ such that $S = \{x : \exists y \in \{0,1\}^{p(|x|)} \text{ s.t. } (x,y) \in S'\}$.*

Proving that $\mathcal{PH}$ reduces to $\#\mathcal{P}$

Recall that Toda's Theorem asserts that $\mathcal{PH}$ is Cook-reducible to $\#\mathcal{P}$ (via deterministic reductions). Here we prove a closely related result (also due to Toda [6]), which relaxes the requirement from the reduction (allowing it to be randomized) but uses an oracle to a seemingly weaker class. The latter class is denoted $\oplus\mathcal{P}$ and is the "modulo 2 analogue" of $\#\mathcal{P}$. Specifically, a Boolean function f is in $\oplus\mathcal{P}$ if there exists a function $g \in \#\mathcal{P}$ such that for every x it holds that $f(x) = g(x) \bmod 2$. Equivalently, f is in $\oplus\mathcal{P}$ if there exists a search problem $R \in \mathcal{PC}$ such that $f(x) = |R(x)| \bmod 2$, where $R(x) = \{y : (x,y) \in R\}$. (The $\oplus$ in the notation $\oplus\mathcal{P}$ actually represents parity, which is merely addition modulo 2. Indeed, a notation such as $\#_2\mathcal{P}$ would have been more appropriate.)

Theorem 2 *Every set in $\mathcal{PH}$ is reducible to $\oplus\mathcal{P}$ via a probabilistic polynomial-time reduction. Furthermore, the reduction is many-to-one and fails with negligible error probability.*

The proof follows the underlying ideas of the original proof [6], but the actual presentation is quite different. Alternative proofs of Theorem 2 can be found in [2, 5].

Proof Sketch: The proof uses three main ingredients. The first ingredient is the fact that $\mathcal{NP}$ is reducible to $\oplus\mathcal{P}$ via a probabilistic polynomial-time Karp-reduction, and that this reduction in "highly structured" (see Footnote 2). The second ingredient is the fact that error-reduction is available in the correct context, resulting in reductions that have exponentially vanishing error probability.[1] The third ingredient may be schematically paraphrased by the Boolean equality $\oplus_i(\zeta_i \wedge (\oplus_j X_{i,j})) = \oplus_{i,j}(\zeta_i \wedge X_{i,j})$. These ingredients correspond to the three main steps of the proof.

Rather than presenting the actual proof at an abstract level (while using suitable definitions), we prefer a concrete presentation in which the third step is performed by an extension of the first step. In particular, this allows performing the third step at a level that clarifies what exactly is going on. In addition, it offers the opportunity for revisiting the standard presentations of the first step, while correcting what we consider to be a conceptual error in these presentations. Thus, we begin by dealing with the easy case of $\mathcal{NP}$ (and co$\mathcal{NP}$), and then turn the implementation of error-reduction (in the current context). *Such error-reduction is crucial as a starting point for the third step, which deals with the case of* Σ_2. When completing the third step, we will have all the ingredients needed for the general case (of dealing with Σ_k for any $k \geq 2$), and we will thus conclude with a few comments regarding the latter case. Admittingly, the description of the last part is very sketchy and an actual implementation would be quite cumbersome; however, the ideas are all present in the case of Σ_2. Furthermore, we believe that the case of Σ_2 is of significant interest per se.

Let us first prove that every set in $\mathcal{NP}$ is reducible to $\oplus\mathcal{P}$ via a probabilistic polynomial-time Karp-reduction. Indeed, this follows immediately from the NP-hardness of deciding unique solution for some relations $R \in \mathcal{PC}$ (i.e., Theorem 3), because the corresponding modulo 2 counter (i.e., $\#R \bmod 2$) solves the unique solution problem associated with this relation (i.e., deciding the existence of unique solutions for R). Specifically, Theorem 3 asserts that, for some complete problems $R \in \mathcal{PC}$, deciding membership in any NP-set is reducible in probabilistic polynomial-time to the promise problem $(\mathtt{US}_R, \overline{S}_R)$, where $\mathtt{US}_R = \{x : |R(x)| = 1\}$ and $\overline{S}_R = \{x : |R(x)| = 0\}$. The point is that the function $\oplus R(x) \stackrel{\rm def}{=} |R(x)| \bmod 2$ solves the latter promise problem; that is, $(\mathtt{US}_R, \overline{S}_R)$ is reducible to $\oplus R$ by the identity mapping. Thus, any reduction to the promise problem $(\mathtt{US}_R, \overline{S}_R)$ constitutes a reduction to $\oplus R$. Still, for the sake of self-containment and concreteness, let us consider an alternative proof.[2]

Step 1: a direct proof for the case of $\mathcal{NP}$. As in the proof of Theorem 3, we start with any $R \in \mathcal{PC}$ and our goal is reducing $S_R = \{x : |R(x)| \geq 1\}$ to $\oplus\mathcal{P}$ by a randomized Karp-reduction.[3] The standard way of obtaining such a reduction (e.g., in [2, 4, 5, 6]) consists of just using the reduction presented in the proof of Theorem 3, but we believe that this way is conceptually wrong. Recall that the proof of Theorem 3 consists of implementing a randomized sieve that has the following property. For any $x \in S_R$, with noticeable probability, a single element of $R(x)$ passes the sieve (and this event can be detected by an oracle to a unique solution problem). Indeed, an oracle in $\oplus\mathcal{P}$ correctly detects the case in which a single element of $R(x)$ passes the sieve. However, by definition, an oracle in $\oplus\mathcal{P}$ correctly detects the more general case in which any odd number of elements of $R(x)$ pass the sieve. Thus, insisting on a random sieve that allows the passing of a single

[1] We comment that such an error-reduction is not available in the context of reductions to unique solution problems. This comment is made in view of the similarity between the reduction of $\mathcal{NP}$ to $\oplus\mathcal{P}$ and the reduction of $\mathcal{NP}$ to problems of unique solution.

[2] Indeed, the presentation can be modified such that the following direct proof is omitted. In this case, we shall only use the fact that each set in $\mathcal{NP}$ is reducible to $\oplus\mathcal{P}$ by a randomized Karp-reduction. Actually, we will have to rely on the fact that the reduction is "highly structured" in the sense that for any polynomially bounded relation R it reduces S_R to $\oplus R_2$ such that x is mapped to $\langle x, s\rangle$ and $y \in R_2(\langle x, s\rangle)$ if and only if $y \in R(x) \wedge \psi(x, s, y)$, where ψ is some polynomial-time computable predicate.

[3] As in Theorem 3, if any search problem in $\mathcal{PC}$ is reducible to R via a parsimonious reduction, then we can reduce S_R to $\oplus R$. Specifically, we shall show that S_R is randomly reducible to $\oplus R_2$, for some $R_2 \in \mathcal{PC}$, and a reduction of S_R to $\oplus R$ follows (by using the parsimonious reduction of R_2 to R).

element of $R(x)$ seems an over-kill (or at least is conceptually wrong). Instead, *we should just apply a less stringent random sieve that, with noticeable probability, allows the passing of an odd number of elements of $R(x)$*. The adequate tool for this sieve is a small-bias generator (see notes at the end of this text). Specifically, we use a strongly efficient generator that given a seed s and index i produces the adequate bit, denoted $G(s,i)$, in the $\ell(|s|)$-bit generator sequence $G(s)$, where $G(U_k)$ has small bias and $\ell(k) = \exp(\Omega(k))$. Assuming, without loss of generality, that $R(x) \subseteq \{0,1\}^{p(|x|)}$ for some polynomial p, we consider the relation

$$R_2 = \{(\langle x,s\rangle,y) : (x,y)\in R \wedge G(s,y)=1\} \tag{1}$$

where $y \in \{0,1\}^{p(|x|)} \equiv [2^{p(|x|)}]$ and $s \in \{0,1\}^{O(|y|)}$ such that $\ell(|s|) = 2^{|y|}$. Then, for every $x \in S_R$, with probability at least $1/3$, a uniformly selected $s \in \{0,1\}^{O(|y|)}$ satisfies $|R_2(\langle x,s\rangle)| \equiv 1 \pmod 2$, whereas for every $x \notin S_R$ and every $s \in \{0,1\}^{O(|y|)}$ it holds that $|R_2(\langle x,s\rangle)| = 0$. A key observation is that $R_2 \in \mathcal{PC}$ (and thus $\oplus R_2$ is in $\oplus\mathcal{P}$). Thus, deciding membership in S_R is randomly reducible to $\oplus R_2$ (by the many-to-one randomized mapping of x to $\langle x,s\rangle$, where s is uniformly selected in $\{0,1\}^{O(|y|)}$). Since the foregoing holds for any $R \in \mathcal{PC}$, it follows that $\mathcal{NP}$ is reducible to $\oplus\mathcal{P}$ via randomized Karp-reductions.

Dealing with co$\mathcal{NP}$. We may Cook-reduce co$\mathcal{NP}$ to $\mathcal{NP}$ and thus prove that co$\mathcal{NP}$ is randomly reducible to $\oplus\mathcal{P}$, but we wish to highlight the fact that a randomized Karp-reduction will also do. Starting with the reduction present for the case of sets in $\mathcal{NP}$, we note that for $S \in$ co$\mathcal{NP}$ we obtain a relation R_2 such that $x \in S$ is indicated by $|R_2(\langle x,\cdot\rangle)| \equiv 0 \pmod 2$. We wish to flip the parity such that $x \in S$ will be indicated by $|R_2(\langle x,\cdot\rangle)| \equiv 1 \pmod 2$, and this can be done by augmenting the relation R_2 with a single dummy solution per each x. For example, we may redefine $R_2(\langle x,s\rangle)$ as $\{0y : y\in R_2(\langle x,s\rangle)\} \cup \{10^{p(|x|)}\}$. Indeed, we have demonstrated and used the fact that $\oplus\mathcal{P}$ is closed under complementation.

We note that dealing with the cases of $\mathcal{NP}$ and co$\mathcal{NP}$ is of interest only because we reduced these classes to $\oplus\mathcal{P}$ rather than to $\#\mathcal{P}$. In contrast, even a reduction of Σ_2 to $\#\mathcal{P}$ is of interest, and thus the reduction of Σ_2 to $\oplus\mathcal{P}$ (presented in Step 3) is interesting. This reduction relies heavily on the fact that error-reduction is applicable in the context of randomized Karp-reductions to $\oplus\mathcal{P}$.

Step 2: error reduction. An important observation, towards the core of the proof, is that it is possible to drastically reduce the (one-sided) error probability in randomized Karp-reductions to $\oplus\mathcal{P}$. Specifically, let R_2 be as in Eq. (1) and t be any polynomial. Then, a binary relation R_2' that satisfies

$$|R_2'(\langle x,s_1,...,s_{t(|x|)}\rangle)| \;=\; 1 + \prod_{i=1}^{t(|x|)} (1 + |R_2(\langle x,s_i\rangle)|) \tag{2}$$

offers such an error reduction, because $|R_2'(\langle x,s_1,...,s_{t(|x|)}\rangle)|$ is odd if and only if for some $i \in [t(|x|)]$ it holds that $|R_2(\langle x,s_i\rangle)|$ is odd. Thus,

$$\begin{aligned}&\mathsf{Pr}_{s_1,\dots,s_{t(|x|)}}[|R_2'(\langle x,s_1,...,s_{t(|x|)}\rangle)| \equiv 0 \pmod 2]\\ &\qquad = \mathsf{Pr}_s[|R_2(\langle x,s\rangle)| \equiv 0 \pmod 2]^{t(|x|)}\end{aligned}$$

where $s, s_1, ..., s_{t(|x|)}$ are uniformly and independently distributed in $\{0,1\}^{O(p(|x|))}$ (and p is such that $R(x) \subseteq \{0,1\}^{p(|x|)}$). This means that the one-sided error probability of a randomized reduction of S_R to $\oplus R_2$ (which maps x to $\langle x,s\rangle$) can be reduced by reducing S_R to $\oplus R_2'$, where the reduction maps x to $\langle x,s_1,...,s_{t(|x|)}\rangle$. Specifically (for $S_R \in \mathcal{NP}$), error probability ε (e.g., $\varepsilon = 2/3$) in the case that we desire an "odd outcome" (i.e., $x \in S_R$) is reduced to error probability ε^t, whereas zero error probability in the case of a desired "even outcome" (i.e., $x \in \overline{S}_R$) is preserved. A key question is whether this yields error-reduction for reductions to $\oplus\mathcal{P}$; that is, whether R_2' (as postulated in Eq. (2)) can be implemented in $\mathcal{PC}$ (and so imply $\oplus R_2' \in \oplus\mathcal{P}$). The answer is positive, and this can be shown by using a Cartesian product construction (and adding some dummy solutions). For example, let $R_2'(\langle x, s_1, ..., s_{t(|x|)}\rangle)$ consists of tuples $\langle \sigma_0, y_1, ..., y_{t(|x|)}\rangle$ such that either $\sigma_0 = 1$ and $y_1 = \cdots = y_{t(|x|)} = 0^{p(|x|)+1}$ or $\sigma_0 = 0$ and for every $i \in [t(|x|)]$ it holds that $y_i \in (\{0\} \times R_2(\langle x, s_i\rangle)) \cup \{10^{p(|x|)}\}$.

We wish to stress that, when starting with R_2 as in Eq. (1), the forgoing process of error-reduction can be used for obtaining error probability that is upper-bounded by $\exp(-q(|x|))$ for any desired polynomial q. The importance of this comment will become clear shortly.

Step 3: the case of Σ_2. With the foregoing preliminaries, we are now ready to handle the case of $S \in \Sigma_2$. By Proposition 1, there exists a polynomial p and a set $S' \in \Pi_1 = \mathrm{co}\mathcal{NP}$ such that $S = \{x : \exists y \in \{0,1\}^{p(|x|)} \text{ s.t. } (x, y) \in S'\}$. Using $S' \in \mathrm{co}\mathcal{NP}$, we apply the forgoing reduction of S' to $\oplus\mathcal{P}$ as well as an adequate error-reduction that yields an upper-bound of $\varepsilon \cdot 2^{-p(|x|)}$ on the error probability, where $\varepsilon \le 1/7$ is unspecified at this point. (For the case of Σ_2 the setting $\varepsilon = 1/7$ will do, but for the dealing with Σ_k we will need a much smaller value of $\varepsilon > 0$.) Thus, we obtain a relation $R_2' \in \mathcal{PC}$ such that the following holds: *for every x and $y \in \{0,1\}^{p(|x|)}$, with probability at least $1 - \varepsilon \cdot 2^{-p(|x|)}$ over the random choice of $s' \in \{0,1\}^{O(p(|x|))^2}$, it holds that $x' \stackrel{\text{def}}{=} (x, y) \in S'$ if and only if $|R_2'(\langle x', s'\rangle)|$ is odd.*[4] Using a union bound (over all possible $y \in \{0,1\}^{p(|x|)}$), it follows that, with probability at least $1 - \varepsilon$ over the choices of s', it holds that $x \in S$ if and only if there exists a y such that $|R_2'(\langle (x, y), s'\rangle)|$ is odd. Now, as in the treatment of $\mathcal{NP}$, we wish to reduce the latter "existential problem" to $\oplus\mathcal{P}$. That is, we wish to define a relation $R_3 \in \mathcal{PC}$ such that for a randomly selected s the value $|R_3(\langle x, s, s'\rangle)| \bmod 2$ provides an indication to whether or not $x \in S$ (by indicating whether or not there exists a y such that $|R_2'(\langle (x, y), s'\rangle)|$ is odd). Analogously to Eq. (1), consider the binary relation

$$I_3 = \{(\langle x, s, s'\rangle, y) : |R_2'(\langle (x, y), s'\rangle| \equiv 1 (\mathrm{mod}\ 2) \ \wedge G(s, y) = 1\}. \tag{3}$$

Indeed, if $x \in S$ then, with probability at least $1 - \varepsilon$ over the random choice of s' and probability at least $1/3$ over the random choice of s, it holds that $|I_3(\langle x, s, s'\rangle)|$ is odd, whereas for every $x \notin S$ and every choice of s it holds that $\mathsf{Pr}_{s'}[|I_3(\langle x, s, s'\rangle)| = 0] \ge 1 - \varepsilon$. (For $\varepsilon \le 1/7$, it follows for every $x \in S$ we have $\mathsf{Pr}_{s,s'}[|I_3(\langle x, s, s'\rangle)| \equiv 1 \ (\mathrm{mod}\ 2)] \ge (1 - \varepsilon)/3 \ge 2/7$, whereas for every $x \notin S$ we have $\mathsf{Pr}_{s,s'}[|I_3(\langle x, s, s'\rangle)| \equiv 1 \ (\mathrm{mod}\ 2)] \le \varepsilon \le 1/7$.) Thus, $|I_3(\langle x, \cdot, \cdot\rangle)| \bmod 2$ provides a randomized indication to whether or not $x \in S$, but it is not clear whether I_3 is in $\mathcal{PC}$ (and in fact I_3 is likely not to be in $\mathcal{PC}$). The key observation is that

$$\begin{aligned} |R_3(\langle x, s, s'\rangle)| \quad &\equiv \quad |I_3(\langle x, s, s'\rangle)| \quad (\mathrm{mod}\ 2) \\ \text{where} \qquad & R_3(\langle x, s, s'\rangle) \stackrel{\text{def}}{=} \{\langle y, z\rangle : (\langle (x, y), s'\rangle, z) \in R_2' \wedge G(s, y) = 1\} \end{aligned} \tag{4}$$

(with $\langle y, z\rangle \in \{0,1\}^{p(|x|)} \times \{0,1\}^{p'(|x|)}$), where Eq. (4) is justified by letting $\chi_{y,z} = 1$ (resp., ξ_y) indicate the event $(\langle (x, y), s'\rangle, z) \in R_2'$ (resp., the event $G(s, y) = 1$), and noting that $\oplus_{y,z}\chi_{y,z} \wedge \xi_y$ equals $\oplus_y(\oplus_z \chi_{y,z}) \wedge \xi_y$. The punch-line is that $R_3 \in \mathcal{PC}$. It follows that S is randomly Karp-reducible to $\oplus\mathcal{P}$ (by the many-to-one randomized mapping of x to $\langle x, s, s'\rangle$, where (s, s') is uniformly selected in $\{0,1\}^{O(p(|x|))} \times \{0,1\}^{O(p'(|x|))}$).

[4] Note that $R_2' \subseteq \{0,1\}^{|x|+p(|x|)+O(p(|x|)^2)} \times \{0,1\}^{p'(|x|)}$, where p' is some polynomial that may depend on p. In particular, the specific implementation of R_2', which uses $t = O(p)$, yields $p' = O(p^2)$.

Again, error-reduction may be applied to this reduction (of Σ_2 to $\oplus\mathcal{P}$) such that it can be used for dealing with Σ_3. A technical difficulty arises since the foregoing reduction has two-sided error probability, where one type (or "side") of error is due to the error in the reduction of $S' \in \mathrm{co}\mathcal{NP}$ to $\oplus R'_2$ (which occurs on no-instances of S') and the second type (or "side") of error is due to the (new) reduction of S to $\oplus R_3$ (and occurs on the yes-instances of S). However, the error probability in the first reduction is (or can be made) very small and can be ignored when applying error-reduction to the second reduction. See following comments.

The general case. First note that, as in the case of $\mathrm{co}\mathcal{NP}$, we can obtain a similar reduction for $\Pi_2 = \mathrm{co}\Sigma_2$. It remains to extend the treatment of Σ_2 to Σ_k, for every $k \geq 2$. Indeed, $S \in \Sigma_k$ is treated by considering a polynomial p and a set $S' \in \Pi_{k-1}$ such that $S = \{x : \exists y \in \{0,1\}^{p(|x|)} \text{ s.t. } (x,y) \in S'\}$. Next, we use a relation R'_k such that, with overwhelmingly high probability over the choice of s' the value $|R'_k(\langle (x,y), s'\rangle)| \bmod 2$ indicates whether or not $(x,y) \in S'$. Using the ideas underlies the treatment of $\mathcal{NP}$ (and Σ_2) we check whether for some y it holds that $|R'_k(\langle (x,y), s'\rangle)| \equiv 1 \pmod 2$. This yields a relation R_{k+1} such that for random s, s' the value $|R_{k+1}(\langle x, s, s'\rangle)| \bmod 2$ indicates whether or not $x \in S$. Finally, we apply error reduction, while ignoring the probability that s' is bad, and obtain the desired relation R'_{k+1}. This means that if we wish to upper-bound the error probability in the reduction (of S) to $\oplus R'_{k+1}$ by ε_{k+1}, then the error probability in the reduction (of S') to $\oplus R'_k$ should be upper-bounded by $\varepsilon_k = \varepsilon_{k+1} \cdot 2^{-p(|x|)}$. Thus, the proof that $\mathcal{PH}$ is randomly reducible to $\oplus\mathcal{P}$ actually proceed "top down" (at least partially); that is, starting with an arbitrary $S \in \Sigma_k$, we first determine the auxiliary sets (as per Proposition 1) as well as the error-bounds that should be proved for the reductions of these sets (which reside in lower levels of $\mathcal{PH}$), and only then we establish the existence of such reductions. Indeed, this latter (and main) step is done "bottom up" using the reduction (to $\oplus\mathcal{P}$) of the set in the i^{th} level when reducing (to $\oplus\mathcal{P}$) the set in the $i+1^{\text{st}}$ level. $\square$

Notes

In the main text, we refer to a version of the Valiant-Vazirani Theorem, which is stated below. For a binary relation R, we denote $R(x) = \{y : (x,y) \in R\}$, and say that x has a unique solution $|R(x)| = 1$. We say that a many-to-one reduction f of R' to R is parsimonious if for every x it holds that $|R(x)| = |R'(f(x))|$.

Theorem 3 *Let $R \in \mathcal{PC}$ and suppose that every search problem in $\mathcal{PC}$ is parsimoniously reducible to R. Then solving the search problem of R* (resp., deciding membership in $S_R = \{x : |R(x)| \geq 1\}$) *is reducible in probabilistic polynomial-time to finding unique solutions for R* (resp., the promise problem $(\mathtt{US}_R, \overline{S}_R)$, where $\mathtt{US}_R = \{x : |R(x)| = 1\}$ and $\overline{S}_R = \{x : |R(x)| = 0\}$). *Furthermore, there exists a probabilistic polynomial-time computable mapping M such that for every $x \in \overline{S}_R$ it holds that $M(x) \in \overline{S}_R$, whereas for every $x \in S_R$ it holds that* $\mathsf{Pr}[M(x) \in \mathtt{US}_R] \geq 1/\mathrm{poly}(|x|)$.

The proof of Theorem 3 uses a mapping of x to $\langle x, i, h\rangle$, where i is uniformly selected in $\{1, ..., \mathrm{poly}(|x|)\}$ and h is a pairwise independent hashing function mapping $\mathrm{poly}(|x|)$-bit long strings to i-bit long strings. This mapping reduces S_R to the promise problem $(\mathtt{US}_{R'}, \overline{S}_{R'})$, where $R' = \{(\langle x, i, h\rangle, y) : (x,y) \in R \wedge h(y) = 0^i\}$ is clearly in $\mathcal{PC}$. Note that every $x \in \overline{S}_R$ is mapped to $\overline{S}_{R'}$, whereas for every $x \in S_R$ it holds that $\mathsf{Pr}_{i,h}[\langle x, i, h\rangle \in \mathtt{US}_{R'}] > 1/\mathrm{poly}(|x|)$. The desired reduction to $(\mathtt{US}_R, \overline{S}_R)$ is obtained by composing the foregoing reduction with parsimonious reduction of R' to R.

Small bias generators. For $\varepsilon : \mathbb{N} \to [0,1]$, an ε-bias generator with stretch function ℓ is an efficient deterministic algorithm (e.g., working in poly($\ell(k)$) time) that expands a k-bit long random seed into a sequence of $\ell(k)$ bits such that for any fixed non-empty set $S \subseteq \{1, ..., \ell(k)\}$ the bias of the output sequence over S is at most $\varepsilon(k)$. The bias of a sequence of n (possibly dependent) Boolean random variables $\zeta_1, ..., \zeta_n \in \{0,1\}$ over a set $S \subseteq \{1, .., n\}$ is defined as

$$2 \cdot \left| \Pr[\oplus_{i \in S} \zeta_i = 1] - \frac{1}{2} \right| \;=\; \left| \Pr[\oplus_{i \in S} \zeta_i = 1] - \Pr[\oplus_{i \in S} \zeta_i = 0] \right|. \qquad (5)$$

The factor of 2 is introduced so to make these biases correspond to the Fourier coefficients of the distribution (viewed as a function from $\{0,1\}^n$ to the reals). Efficient small-bias generators with exponential stretch and exponentially vanishing bias are know.

Theorem 4 (small-bias generators [3]): *For some universal constant $c > 0$, let $\ell : \mathbb{N} \to \mathbb{N}$ and $\varepsilon : \mathbb{N} \to [0,1]$ such that $\ell(k) \leq \varepsilon(k) \cdot \exp(k/c)$. Then, there exists an ε-bias generator with stretch function ℓ operating in time polynomial in the length of its output.*

Three simple constructions of small-bias generators that satisfy Theorem 4 are known (see [1]). One of these constructions is based on Linear Feedback Shift Registers. Loosely speaking, the first half of the seed, denoted $f_0 f_1 \cdots f_{(k/2)-1}$, is interpreted as a (non-degenerate) feedback rule[5], the other half, denoted $s_0 s_1 \cdots s_{(k/2)-1}$, is interpreted as "the start sequence", and the output sequence, denoted $r_0 r_1 \cdots r_{\ell(k)-1}$, is obtained by setting $r_i = s_i$ for $i < k/2$ and $r_i = \sum_{j=0}^{(k/2)-1} f_j \cdot r_{i-(k/2)+j}$ for $i \geq k/2$. We highlight the fact that *the aforementioned constructions satisfy a stronger notion of efficient generation, which is use in the main text: there exists a polynomial-time algorithm that given a seed and a bit location $i \in [\ell(k)]$* (in bianry)*, outputs the i^{th} bit of the corresponding output.*

References

[1] N. Alon, O. Goldreich, J. Håstad, R. Peralta. Simple Constructions of Almost k-wise Independent Random Variables. *Journal of Random structures and Algorithms*, Vol. 3, No. 3, (1992), pages 289–304.

[2] R. Kannan, H. Venkateswaran, V. Vinay, and A.C. Yao. A Circuit-based Proof of Toda's Theorem. *Information and Computation*, Vol. 104 (2), pages 271–276, 1993.

[3] J. Naor and M. Naor. Small-bias Probability Spaces: Efficient Constructions and Applications. *SIAM Journal on Computing*, Vol 22, 1993, pages 838–856.

[4] C.H. Papadimitriou. *Computational Complexity.* Addison Wesley, 1994.

[5] D.A. Spielman. *Advanced Complexity Theory*, Lectures 10 and 11. Notes (by D. Lewin and S. Vadhan), March 1997. Available from `http://www.cs.yale.edu/homes/spielman/AdvComplexity/1998/` as `lect10.ps` and `lect11.ps`.

[6] S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, Vol. 20 (5), pages 865–877, 1991.

[7] L.G. Valiant and V.V. Vazirani. NP Is as Easy as Detecting Unique Solutions. *Theoretical Computer Science*, Vol. 47 (1), pages 85–93, 1986.

[5] That is, $f_0 = 1$ and $f(z) \stackrel{\text{def}}{=} z^{k/2} + \sum_{j=0}^{(k/2)-1} f_j \cdot z^j$ is an irreducible polynomial over GF(2).