$\mathrm{S}_2^p \subseteq \mathrm{ZPP}^{\mathrm{NP}}$

Jin-Yi Cai *

Abstract

We show that the class S_2^p is contained in $\mathrm{ZPP}^{\mathrm{NP}}$. The proof uses universal hashing, approximate counting and witness sampling. As a consequence, a collapse first noticed by Samik Sengupta that the assumption NP has small circuits collapses PH to S_2^p becomes the strongest version to date of the Karp-Lipton Theorem. We also discuss the problem of finding irrefutable proofs for S_2^p in $\mathrm{ZPP}^{\mathrm{NP}}$.

1 Introduction

The class S_2^p was introduced independently by Canetti [C96] and Russell and Sundaram [RS95] in the mid 1990's. Suppose there are two competing all powerful provers Y and Z. A string x is given, Y wishes to convince us that $x \in L$, and Z wishes to convince us the opposite $x \notin L$. We—the verifier—have only deterministic polynomial time computing power. A language L is in S_2^p iff there is a P-time predicate P such that the following holds:

> If $x \in L$ then there exists a y, such that for all z, $P(x, y, z)$ holds;
> If $x \notin L$ then there exists a z, such that for all y, $\neg P(x, y, z)$ holds, where both y and z are polynomially bounded in the length of x.

In other words, if $x \in L$ then Y has irrefutable proof y which can withstand any challenge z from Z; and if $x \notin L$ then Z has irrefutable proof z which can withstand any challenge y from Y.

The motivation by both Canetti [C96] and Russell and Sundaram [RS95] was to provide a refinement of the Sipser-Lautemann Theorem (with contribution by Gacs) that $\mathrm{BPP} \subseteq \Sigma_2^p \cap \Pi_2^p$ [Si83, L83]. Indeed, Canetti [C96] extended Lautemann's proof to show that $\mathrm{BPP} \subseteq \mathrm{S}_2^p$, whereas Russell and Sundaram [RS95] showed further that $\mathrm{MA} \subseteq \mathrm{S}_2^p$. Note that $\mathrm{BPP} \subseteq \mathrm{MA}$ is direct from definition (the two-sided error version) of MA, thus $\mathrm{BPP} \subseteq \mathrm{MA} \subseteq \mathrm{S}_2^p$. Also it is known that $\mathrm{P}^{\mathrm{NP}} \subseteq \mathrm{S}_2^p$ [RS95].

As to upper bound of S_2^p, the only known containment is by definition $\mathrm{S}_2^p \subseteq \Sigma_2^p \cap \Pi_2^p$ (see Section 2). Goldreich and Zuckerman [GZ97] surveyed a number of interesting classes between P and the second level of the Polynomial-time Hierarchy Σ_2^p and Π_2^p. These classes

*Computer Sciences Department, University of Wisconsin, Madison, WI 53706. Research supported in part by NSF CCR-0196197, CCR-0208013 and a Guggenheim Fellowship. A preliminary version appeared in FOCS 2001 [C01]. Email: `jyc@cs.wisc.edu`

include $\mathrm{ZPP}, \mathrm{RP}, \mathrm{BPP}, \mathrm{NP}, \mathrm{P}^{\mathrm{NP}}, \mathrm{MA}, \mathrm{AM}, \mathrm{ZPP}^{\mathrm{NP}}$ and S_2^p. They called the classes listed here up to P^{NP} "Traditional classes—classes of the 1970's", the class Arthur-Merlin "a class of the 1980's", and the class S_2^p "a class of the 1990's", underscoring that not much is yet known about this class S_2^p. In their paper [GZ97] Goldreich and Zuckerman gave a number of elegant proofs of known results with the strikingly sharp amplification technique due to Zuckerman [Z96]. They also prove an interesting result $\mathrm{MA} \subseteq \mathrm{ZPP}^{\mathrm{NP}}$. This last result was new in 1997 when [GZ97] appeared; it was independently obtained by Arvind and Köbler [AK97]. In summarizing the known facts about all these classes between P and Σ_2^p and Π_2^p it was observed that both S_2^p and $\mathrm{ZPP}^{\mathrm{NP}}$ appear to share all the known containment properties both below and above [GZ97]. How these two classes are related was unknown.

The main result of this paper is

Theorem 1 $\mathrm{S}_2^p \subseteq \mathrm{ZPP}^{\mathrm{NP}}$.

The proof uses universal hashing, approximate counting and witness sampling. We also discuss the problem of finding irrefutable proofs in $\mathrm{ZPP}^{\mathrm{NP}}$.

There is an interesting consequence of this result with respect to the well known Karp-Lipton Theorem concerning sparse sets (with contribution by Sipser) [KL80]. This theorem says, if NP is Cook-reducible ($\leq_T^p$) to sparse sets, or equivalently, if **SAT** has polynomial size circuits, then the Polynomial-time Hierarchy collapses to its second level: $\mathrm{PH} = \Sigma_2^p \cap \Pi_2^p$. Many researchers have since tried to improve on this signature theorem—To simplify the proof and to strengthen the collapse. On the one hand, there emerged what I consider to be the "book" proof (as Erdös would say) of the theorem (as far as I know John Hopcroft [H81] was the first to give essentially this proof):

> To simulate Π_2^p by Σ_2^p, guess a poly-size circuit C for **SAT**, modify C via self-reducibility so that whenever $C(\phi) = 1$ it also produces a satisfying assignment to ϕ, then check all universal paths of the Π_2^p computation lead to a satisfiable formula.

Samik Sengupta [Se00] first noticed that this "book" proof actually gave the collapse to S_2^p. (See Section 6.)

While the proof of Karp-Lipton Theorem becomes extremely transparent, more research effort went into trying to extend this beautiful result. Much work was done on the general theme (we mention some in Section 6). Over the years there have been steady improvements on the exact level of collapse of PH, assuming **SAT** has small circuits. In this regard, the best result so far is due to Bshouty et. al. [BCGKT94] and Köbler and Watanabe [KW95]. Their result states that if NP has polynomial size circuits, then the Polynomial-time Hierarchy collapses to $\mathrm{ZPP}^{\mathrm{NP}}$. Admittedly the proofs of the theorem of Bshouty et. al. and Köbler-Watanabe are more involved than the "book" proof of the basic version of the Karp-Lipton Theorem and depend on previous interesting results by Jerrum, Valiant and V. Vazirani [JVV86] and others.

By the new theorem $\mathrm{S}_2^p \subseteq \mathrm{ZPP}^{\mathrm{NP}}$ (unconditionally), the (currently) strongest Karp-Lipton Theorem becomes the following Theorem 2. (See Section 6.)

Theorem 2 (Sengupta) *If* **SAT** *has polynomial size circuits, then the Polynomial-time Hierarchy collapses to* S_2^p.

Theorem 1 also subsumes the result $\mathrm{MA} \subseteq \mathrm{ZPP}^{\mathrm{NP}}$ by Goldreich-Zuckerman [GZ97] and Arvind-Köbler [AK97], as we know from Russell and Sundaram [RS95] that $\mathrm{MA} \subseteq \mathrm{S}_2^p$.

2 Preliminaries

The class S_2^p was defined by Russell and Sundaram [RS95] as follows: $L \in \mathrm{S}_2^p$ iff there is a P-time computable 0-1 function P on three arguments, such that

$$x \in L \implies (\exists^p y)(\forall^p z)[P(x,y,z) = 1] \tag{1}$$
$$x \notin L \implies (\exists^p z)(\forall^p y)[P(x,y,z) = 0] \tag{2}$$

where as usual "$\exists^p y$" stands for "$\exists y \in \{0,1\}^{p_1(|x|)}$" for some polynomial $p_1(\cdot)$. Similarly "$\forall^p z$" stands for "$\forall z \in \{0,1\}^{p_2(|x|)}$" for some polynomial $p_2(\cdot)$. By padding we can suitably extend the length of both y and z, and henceforth we can assume they both vary over the same length n which is a power of 2, and n is polynomially bounded in the length of x.

Given x, for convenience, for a pair (y, z) we say y beats z if $P(x, y, z) = 1$, and z beats y if $P(x, y, z) = 0$.

It is immediately clear that both implications "$\Longrightarrow$" can be replaced by the if and only if relation "$\Longleftrightarrow$" without changing the class S_2^p. For instance, suppose $(\exists^p y)(\forall^p z)[P(x,y,z) = 1]$, let y_0 be such a y. Then certainly $x \in L$, else we would have a z_0 such that $(\forall^p y)[P(x,y,z_0) = 0]$, which is clearly a contradiction to $P(x, y_0, z_0) = 1$. Similarly $(\exists^p z)(\forall^p y)[P(x,y,z) = 0]$ implies $x \notin L$. Thus

$$x \in L \iff (\exists^p y)(\forall^p z)[P(x,y,z) = 1]$$
$$x \notin L \iff (\exists^p z)(\forall^p y)[P(x,y,z) = 0]$$

It follows from this if and only if condition that $\mathrm{S}_2^p \subseteq \Sigma_2^p \cap \Pi_2^p$. In fact S_2^p consists of precisely those languages in $\Sigma_2^p \cap \Pi_2^p$ where membership in both Σ_2^p and Π_2^p is demonstrated by the same predicate P.

Canetti [C96] defined the class S_2^p as follows: $L \in \mathrm{S}_2^p$ iff there is a P-time computable 0-1 function P on three arguments, such that for all x,

$$(\exists^p y)(\forall^p z)[P(x,y,z) = \chi_L(x)]$$

and

$$(\exists^p z)(\forall^p y)[P(x,y,z) = \chi_L(x)],$$

where χ_L is the characteristic function of L.

Clearly the Canetti definition implies the Russell-Sundaram definition. The reverse implication also holds. For completeness we sketch a simple proof (see [RS95, C96] for more details.) Suppose a predicate P is given in the Russell-Sundaram definition. We define an extended predicate $\hat{P}$ to satisfy the Canetti definition. For x, suppose y and z vary over $\{0,1\}^n$. Then $\hat{P}$ is defined over $\{0,1\}^{|x|} \times \{0,1\}^{n+1} \times \{0,1\}^{n+1}$:

$$\begin{aligned}
\hat{P}(x, 1y, 1z) &= 1 \\
\hat{P}(x, 1y, 0z) &= P(x, y, z) \\
\hat{P}(x, 0y, 1z) &= P(x, z, y) \\
\hat{P}(x, 0y, 0z) &= 0
\end{aligned}$$

This can be rephrased in the language of boolean matrices. Thus, for the Russell-Sundaram definition, the predicate P, for a given x, corresponds to a boolean matrix M whose rows and columns are indexed by y and $z \in \{0,1\}^n$ respectively. When $x \in L$, there exists an all-1 row; and when $x \notin L$, there exists an all-0 column. In this language, the Canetti definition requires that, when $x \in L$, there exist both an all-1 row as well as an all-1 column; and when $x \notin L$, there exist both an all-0 row as well as an all-0 column.

To go from the Russell-Sundaram definition to the Canetti definition, we simply take the matrix M from the Russell-Sundaram definition, and form the new matrix

$$\begin{pmatrix} 0 & M^T \\ M & J \end{pmatrix},$$

where J denotes the all-1 matrix, and M^T denotes the transpose of M.

ZPP denotes zero-error probabilistic polynomial time. ZPP^{NP} is the class accepted by zero-error probabilistic polynomial time oracle Turing machines using an NP oracle. By Cook's Theorem, we can assume without loss of generality that this oracle is the set of satisfiable boolean formulae **SAT**.

3 Main Theorem

To prove the main Theorem 1, we proceed as follows. Let x be given. Let $\{0,1\}^n$ be the witness sets for both provers Y and Z. Here n is polynomially bounded by $|x|$, and is a power of 2.

We will grow a list $Y_k \subset \{0,1\}^n$ of y's, where $|Y_k| = k$, and $k = 1, 2, \ldots, n^{O(1)}$; initially Y_1 can be arbitrarily given, for example $Y_1 = \{0^n\}$. In the k-th stage, with Y_k in hand, we ask the **SAT** oracle whether there exists a $z \in \{0,1\}^n$ such that $P(x, y, z) = 0$ for every $y \in Y_k$, i.e., a z that beats every $y \in Y_k$. Let

$$Z(Y_k) = \{z \in \{0,1\}^n \mid (\forall y \in Y_k)[P(x, y, z) = 0]\}.$$

Then the question we ask the **SAT** oracle is whether $Z(Y_k) \neq \emptyset$.

Since $|Y_k| = k$ is polynomially bounded, this is clearly a **SAT** query by Cook's Theorem. If the answer is No, i.e., $Z(Y_k) = \emptyset$, then we can already conclude that $x \in L$ and halt. This is because if it were the case that $x \notin L$, by definition it is guaranteed that some z_0 exists beats *all* y, which certainly include all $y \in Y_k$. Note that in this case we concluded $x \in L$, even though we may not have found a witness y_0 which beats every z as promised in the definition.

Hence let's assume the answer to the **SAT** query is Yes, i.e., $Z(Y_k) \neq \emptyset$.

Next we would like to append Y_k to Y_{k+1}. Our goal is, either to find conclusively that $x \notin L$, or to find a new y^* to be appended to the list Y_k so that the corresponding $Z(Y_{k+1})$ is shrunk significantly.

More precisely, we would like either to find conclusively $x \notin L$, or to find with high probability a new y^* such that $|Z(Y_{k+1})| \leq |Z(Y_k)|/2$, where $Y_{k+1} = Y_k \cup \{y^*\}$. If so, we would guarantee that the size $|Z(Y_k)|$ shrinks geometrically every step by a constant fraction with high probability, and thus in polynomial time with high probability we either find out $x \notin L$, or we end up in the case with $Z(Y_k) = \emptyset$, in which case we can conclude that $x \in L$ as discussed earlier.

Lemma 1 *For every set S in P, there is a probabilistic sampling procedure A using a* **SAT** *oracle, such that for every n, and for every $0 < \varepsilon < 1$, $A(n, \varepsilon)$ samples at most $O(n/\varepsilon)$ elements $S' \subseteq S^{=n} = S \cap \{0,1\}^n$ in such a way that, for every subset $T \subseteq S^{=n}$, with $|T| > \varepsilon |S^{=n}|$,*

$$\Pr[S' \cap T = \emptyset] \leq \frac{1}{2^{2n}}.$$

The algorithm runs in time $(n/\varepsilon)^{O(1)}$.

We will discuss Lemma 1 in the next section. For now we assume Lemma 1.

For any witness $y' \in \{0,1\}^n$, consider the set

$$T_{y'} := Z(Y_k \cup \{y'\}) = \{z \in Z(Y_k) \mid P(x, y', z) = 0\}.$$

We say that a $y' \in \{0,1\}^n$ is a "bad witness" with respect to $Z(Y_k)$ if

$$|T_{y'}| = |\{z \in Z(Y_k) \mid P(x, y', z) = 0\}| > \frac{|Z(Y_k)|}{2}.$$

That is, y' is a "bad witness" iff more than $1/2$ of $Z(Y_k)$ beat this y'. Thus for any fixed bad witness y', by Lemma 1 with $\varepsilon = 1/2$, we can sample a polynomial number of $z \in Z(Y_k)$, call the set Z', such that the probability

$$\Pr[Z' \cap T_{y'} = \emptyset] \leq \frac{1}{2^{2n}}.$$

Since there are at most 2^n bad witnesses,

$$\Pr[(\exists \text{ a bad witness } y' \in \{0,1\}^n)[Z' \cap T_{y'} = \emptyset]] \leq \frac{1}{2^n}.$$

Suppose now for every bad witness $y' \in \{0,1\}^n$, the sample set Z' has a non-empty intersection with $T_{y'} = Z(Y_k \cup \{y'\})$. That means that for every bad witness y', y' cannot beat all of Z'. With the polynomial sized set Z' in hand, we ask the **SAT** oracle once again whether there is a y which beats all these $z \in Z'$. Again this is a **SAT** query by Cook's Theorem. If the answer is No, then we know $x \notin L$ since otherwise there is a y which beats all $z \in \{0,1\}^n$, and certainly y beats all these $z \in Z'$. So we reject x and halt.

If the answer is Yes, we use self-reducibility of the **SAT** oracle to obtain one such y^*. Notice that by now there is no bad witness y' which can beat all of Z'. Thus this y^* is not a bad witness. This is true with probability $\geq 1 - 1/2^n$. We then define $Y_{k+1} = Y_k \cup \{y^*\}$. Then with high probability we have

$$|Z(Y_{k+1})| \leq \frac{|Z(Y_k)|}{2}.$$

As remarked earlier this gives our $\mathrm{ZPP}^{\mathrm{NP}}$ algorithm.

4 A Sampling Lemma

The Sampling Lemma 1 follows from the work of Jerrum, Valiant and V. Vazirani [JVV86]. However Lemma 1 has a relatively simple proof based on universal hashing. We give a self-contained account in this section using the notion of *isolation* of Sipser [Si83] (see also [St83]).

Consider a family of hash functions:

$$\{h_s : \{0,1\}^n \to \{0,1\}^k\}_{s \in \mathcal{S}}$$

Recall that a family of hash functions is 2-universal if for every pair of distinct $x \neq y$ in $\{0,1\}^n$, and for every $\alpha, \beta \in \{0,1\}^k$, $\Pr_{s \in \mathcal{S}}[h_s(x) = \alpha \wedge h_s(y) = \beta] = 1/2^{2k}$, i.e., $h_s(x)$ and $h_s(y)$ are pair-wise independent and uniformly distributed when $s \in_R \mathcal{S}$. It is well known such a family of 2-universal hash functions exists and can be easily constructed with small sample space, e.g., $h_{a,b}(x) = ax + b$ and then truncate to k bits, where a, b and x range over a finite field $\mathbf{GF}[2^n]$.

Here is an outline of the proof of Lemma 1. First we will use hash functions and the **SAT** oracle to get an approximate count of the subset $S^{=n}$. We will use the notion of *isolation* of Sipser [Si83] for this. Using the **SAT** oracle we can decide if $S^{=n} = \emptyset$. If so then Lemma 1 is vacuously true (no subset T exists with $|T| > \varepsilon|S^{=n}|$). Suppose $S^{=n} \neq \emptyset$. Then we will devise a simple sampling strategy based on an estimate of the number of points with unique inverse images from $S^{=n}$ under a random hash function. The details follow.

Given $x \neq y$, we say x **collides** with y under h_s if $h_s(x) = h_s(y)$. For a subset $E \subseteq \{0,1\}^n$, we say that h_s **isolates** $x \in E$ iff x does not collide under h_s with any other element of E. The following lemma of Sipser is well known and follows from a simple probability estimate [Si83].

Lemma 2 *Let $E \subseteq \{0,1\}^n$, and let $\{h_s : \{0,1\}^n \to \{0,1\}^k\}_{s \in \mathcal{S}}$ be a family of 2-universal hash functions of cardinality 2^{2n} with $1 \leq k \leq n$. Then for all $m \geq k$,*

1. *if $|E| \leq 2^{k-1}$ then*

$$\Pr_{s_1,\ldots,s_m \in_R \mathcal{S}}[\forall x \in E \textit{ some } h_{s_i} \textit{ isolates } x\,] \geq 1 - \frac{1}{2^{m-k+1}}$$

2. *if $|E| > m2^k$ then*

$$\Pr_{s_1,\ldots,s_m \in_R \mathcal{S}}[\forall x \in E \textit{ some } h_{s_i} \textit{ isolates } x\,] = 0.$$

For our set $E = S^{=n}$, there is some k_e, where $1 \le k_e \le n$, such that $2^{k_e - 1} \le |E| \le 2^{k_e}$. If we take every k in the range $1 \le k \le n+1$, and randomly pick $m = 4n$ hash functions $h_{s_1}, \ldots, h_{s_m} : \{0,1\}^n \to \{0,1\}^k$, with probability $\ge 1 - \frac{1}{2^{3n}}$, at least for $k = k_e + 1$, we would get *isolation*. For each k we ask the **SAT** oracle, whether the chosen set of $h_{s_1}, \ldots, h_{s_m}$ has the property that "$\forall x \in E$, one of h_i isolates x". Since there are only $m = 4n$ hash functions this is a **SAT** query. We pick the least k_0 such that the oracle confirms *isolation*. We abort if for no k the chosen hash functions achieve *isolation*. With probability $\ge 1 - \frac{1}{2^{3n}}$ we do not abort, and we get $k_0 \le k_e + 1$. Also by the second part of the Lemma 2, we know definitely $|E| \le 4n2^{k_0}$. Denote by $U = 4n2^{k_0}$. This is defined with probability $\ge 1 - \frac{1}{2^{3n}}$. Whenever k_0 is defined, U is an upper bound of $|E|$. Also, with probability $\ge 1 - \frac{1}{2^{3n}}$, U is defined and it is not too far from a lower bound of $|E|$,

$$\frac{U}{16n} \le |E| \le U.$$

Let $r = 2^{\lceil \log_2 1/\varepsilon \rceil}$, so that $1/\varepsilon \le r < 2/\varepsilon$. Also $r \ge 2$ as $\varepsilon < 1$. Let $R = \{0,1\}^{k_0 + \log_2 n + \lceil \log_2 1/\varepsilon \rceil + 4}$. Then $|R| = 4rU$.

The sampling procedure can be summarized as follows: First we get an estimate U as described above. Then, for each $1 \le i \le 3n$, uniformly and independently choose a hash function $h_i : \{0,1\}^n \to R$. Now repeat the following $2^{10} r^2 n^2$ times for each h_i: Uniformly and independently pick a target $\alpha \in R$. Ask the **SAT** oracle whether it has an inverse image from the set $E = S^{=n}$. Since S is in P, this is a **SAT** query. If $\alpha \in h_i(E)$, we use self-reducibility to get one inverse image. This inverse image is a sample point. We exit the "repeat" loop as soon as we obtain $4rn$ samples.

```
1.  Get estimate U = 4n2^{k_0}
2.  For i = 1,...,3n
3.      Randomly pick h_{s_i} : {0,1}^n -> R with |R| = 4rU
4.      Repeat 2^{10}r^2n^2 times steps 5 and 6
5.          Randomly pick α ∈ R
6.          Try to find an x ∈ E s.t. h_{s_i}(x) = α using SAT
7.        if found 4rn points, Goto 3 with i := i + 1.
```

Consider $3n$ hash functions $h_1, h_2, \ldots, h_{3n}$ uniformly and independently chosen. For any such h, define the random variable C to be the number of colliding pairs,

$$C = \sum_{\{x,y\} \subseteq E, x \ne y} \chi_{[h(x) = h(y)]}.$$

The expectation of C is

$$\mathbf{E}[C] = \sum_{\{x,y\} \subseteq E, x \ne y} \Pr[h(x) = h(y)] = \binom{|E|}{2} \frac{1}{|R|} < \frac{|E|}{8r}.$$

Hence by Markov's inequality

$$\Pr[C \ge \varepsilon |E|/4] \le \frac{1}{2}. \tag{3}$$

We say a point $\alpha \in R$ is a *unique image* if there is a unique $x \in E$ such that $h(x) = \alpha$. Suppose $C \leq \varepsilon|E|/4$, then there can be at most $\varepsilon|E|/2$ many $x \in E$ involved in a collision, i.e., such that there exists some $y \neq x$, $y \in E$, $h(x) = h(y)$. At least $(1 - \varepsilon/2)|E| \geq |E|/2$ elements of E are mapped to a unique image. Also by assumption $|T| > \varepsilon|E|$, at least $\varepsilon|E|/2$ many elements from T are mapped to a unique image.

For each h_i, the sampling procedure will produce $O(n/\varepsilon)$ points in time $(n/\varepsilon)^{O(1)}$. The probability that the procedure fails to produce any point from T is bounded by the sum of probabilities of the following events: (E1) One did not get a good estimate U; or else, (E2) $\forall 1 \leq i \leq 3n$, the collision set for h_i is large: $|C_i| \geq \varepsilon|E|/4$; or else, (E3) the first i for which the C_i is small, yet less than $4rn$ points from $h_i(E)$ are picked; or else, (E4) for this i the first $4rn$ points from $h_i(E)$ all do not produce points from T.

We have seen $\Pr[E1] \leq 2^{-3n}$. Also, $\Pr[E2] \leq 2^{-3n}$ by (3).

For E3, we use the following version of Chernoff Bound:

Chernoff Bound: For any $0 < p < 1$ and $0 < \delta \leq p(1-p)$, if $X_i, i = 1, \ldots, \ell$ are i.i.d Bernoulli 0-1 variables with $\Pr[X_i = 1] = p$, then

$$\Pr\left[\left|\sum_{i=1}^{\ell} X_i - p\ell\right| \geq \delta\ell\right] \leq 2e^{-\frac{\delta^2\ell}{2p(1-p)}}. \tag{4}$$

If $|C_i| \leq \varepsilon|E|/4$, then $|h_i(E)| \geq |E|/2 \geq U/32n$, thus a target α belongs to $h_i(E)$ has probability at least $|h_i(E)|/|R| \geq \frac{1}{2^7 rn}$. Thus in our case, $p \geq \frac{1}{2^7 rn}$, $\ell = 2^{10}r^2n^2$, and let $\delta = p/2$. Then a simple calculation gives

$$\Pr[E3] \leq 2e^{-rn} \leq 2e^{-2n}.$$

Finally for E4, for this h_i, $\varepsilon|E|/2$ many elements from T are mapped to unique images, thus each time a random $\alpha \in h_i(E)$ is picked, it has probability at least $\varepsilon|E|/(2|h_i(E)|) \geq \varepsilon/2 \geq 1/2r$ to give a sample point from T. (If $\alpha \in h_i(T)$ is a unique image, then the self-reducibility procedure with **SAT** *will* produce a pre-image from T.) It follows that $\Pr[E4] \leq (1 - 1/2r)^{4rn} < e^{-2n}$.

Adding up all the error probabilities, we get

$$\Pr[S' \cap T = \emptyset] \leq \frac{1}{2^{2n}}.$$

The procedure as stated will produce $O(n^2/\varepsilon)$ points. (This is sufficient for our Theorem 1.) However, for each hash function h_i one can check whether the collision set C_i is approximately small probabilistically using **SAT**, and proceed to produce $4rn$ samples only for the first h_i for which the C_i is found small. The modified procedure produces only $O(n/\varepsilon)$ points in $(n/\varepsilon)^{O(1)}$ time. This completes the proof of Lemma 1.

5 In Search of Irrefutable Proofs

Let $L \in \mathrm{S}_2^p$ be defined as in (1)(2). If $x \in L$, then there exists y that beats all z. We call such a y an *irrefutable proof* w.r.t. P. Similarly if $x \notin L$ there are *irrefutable proofs* w.r.t. P, namely any z which beats all y. We have shown that membership $x \in L$ is decidable in $\mathrm{ZPP}^{\mathrm{NP}}$. However in neither case have we produced, in general, an irrefutable proof.

Say $x \in L$, then one simple case is already problematic when we have a polynomial number of y_i's and according to **SAT** there are no z that beat all these y_i's. While this is sufficient to conclude that $x \in L$ (and hence an irrefutable proof y exists), it does not help in locating one such. Moreover, suppose it happens to be that most $y \in \{0,1\}^n$ beats most but not all $z \in \{0,1\}^n$ w.r.t. P, then our proof of Theorem 1 in fact will not find an irrefutable proof with high probability.

However, for any $L \in \mathrm{S}_2^p$, we *can* find an irrefutable proof w.r.t. *some* predicate also defining L.

Theorem 3 *For every $L \in \mathrm{S}_2^p$, there is a P-time predicate Q defining L, such that irrefutable proof w.r.t. Q can be found in* $\mathrm{ZPP}^{\mathrm{NP}}$.[1]

Given L defined via P, define Q as follows:

$$Q(x; y_1, \ldots, y_m; z_1, \ldots, z_m) = 1 \Longleftrightarrow |\{(i,j) \mid 1 \le i, j \le m, P(x, y_i, z_j) = 1\}| > \frac{m^2}{2},$$

where x is the input to L, $y_i, z_j \in \{0,1\}^n$, the length $n = |x|^{O(1)}$ is determined by P, and $m = 7n$ or $7n+1$, whichever is odd.

It is clear that Q is defined symmetrically. Also Q defines L: if $x \in L$, one can take all y_i to be an irrefutable proof y w.r.t. P. The case $x \notin L$ is symmetric.

We claim that in $\mathrm{ZPP}^{\mathrm{NP}}$ we can find an irrefutable proof w.r.t. Q in the following strong sense: Suppose $x \in L$, it will find a sequence $y_1, \ldots, y_m$ such that $\forall z \in \{0,1\}^n$,

$$|\{i \mid P(x, y_i, z) = 1, 1 \le i \le m\}| > m/2, \tag{5}$$

and symmetrically if $x \notin L$.

By symmetry, we assume $x \in L$, and have found out this is so in $\mathrm{ZPP}^{\mathrm{NP}}$. The sequence $y_1, \ldots, y_m$ is defined inductively. $y_1, \ldots, y_k$ defines $\{\mathcal{Z}_k\}_{k \ge 0}$, a sequence of partitions of $Z = \{0,1\}^n$. $\mathcal{Z}_k = \{Z_{k0}, Z_{k1}, \ldots, Z_{kk}\}$ consists of $k+1$ disjoint subsets of Z, where $Z_{k,i}$ consists of those z for which exactly i of $y_1, \ldots, y_k$ beat it. Formally, for $\mathcal{Z}_0$, let $Z_{00} = Z$. For $k \ge 1$, $\mathcal{Z}_k$ is defined as: $\forall z \in Z$, let

$$c_k(z) = c_{y_1,\ldots,y_k}(z) = |\{j \mid P(x, y_j, z) = 1, 1 \le j \le k\}|;$$

then for $0 \le i \le k$,

$$Z_{k,i} = \{z \in Z \mid c_k(z) = i\}.$$

Suppose $\mathcal{Z}_k$ and $y_1, \ldots, y_k$ have been defined. For any y, it divides $Z_{k,i}$ into two parts, $Z_{k,i}^{\epsilon} = \{z \in Z_{k,i} \mid P(x, y, z) = \epsilon\}$, for $\epsilon = 0, 1$. We want to choose $y = y_{k+1}$, so that $|Z_{k,i}^1| \ge \frac{3}{4}|Z_{k,i}|$, for all $0 \le i \le k$. Our y_{k+1} will be chosen probabilistically, and we will

[1] Technically $\mathrm{ZPP}^{\mathrm{NP}}$ is a language class, and thus not for search problems. However the slight abuse of notation is harmless here. The theorem says that a probabilistic P-time algorithm using **SAT** can find *some* irrefutable proof w.h.p. and it never produces a non-irrefutable proof.

argue that it satisfies this condition w.h.p. In other words, let $p_{k,i} = \frac{|Z^1_{k,i}|}{|Z_{k,i}|}$ (if $|Z_{k,i}| = 0$, we let $p_{k,i} = 1$), then we require that

$$p_{k,i} \geq 3/4 \tag{6}$$

for all $k \geq 0$ and $0 \leq i \leq k$. Note that $Z_{k+1,i+1} = Z^1_{k,i} \cup Z^0_{k,i+1}$, if $y = y_{k+1}$.

Lemma 3 *Let $\{\mathcal{Z}_k\}_{k\geq 0}$ be any sequence of partitions of Z, where each $Z_{k,i}$ is divided into a disjoint union $Z_{k,i} = Z^0_{k,i} \cup Z^1_{k,i}$ and $Z_{k+1,i+1} = Z^1_{k,i} \cup Z^0_{k,i+1}$. Suppose $p_{k,i}$ as defined above satisfy (6), then*

$$Z_{m,0} = Z_{m,1} = \ldots = Z_{m,\lfloor \frac{m}{2} \rfloor} = \emptyset,$$

where $m = 7n$ or $7n+1$, whichever is odd.

We will prove Lemma 3 after we complete the proof of Theorem 3 assuming the lemma.

With $\mathcal{Z}_k$ defined and $y_1, \ldots, y_k \in \{0,1\}^n$ in hand, we can apply Lemma 1 (with $\varepsilon = 3/4$) to each $Z_{k,i}$, $0 \leq i \leq k$, and probabilistically produce samples $Z'_{k,i} \subseteq Z_{k,i}$, where each $|Z'_{k,i}|$ is polynomially bounded, and such that

$$\begin{aligned} & \Pr[(\exists y \in \{0,1\}^n)\ y \text{ beats all } Z'_{k,i},\ 0 \leq i \leq k, \text{ yet } \exists i,\ y \text{ beats at most } \tfrac{3}{4} \text{ of } Z_{k,i}] \\ \leq\ & 2^n \cdot (k+1) \cdot \frac{1}{2^{2n}}. \end{aligned}$$

For polynomially bounded k, this is exponentially small.

Assume such y does not exist, then we can ask our **SAT** oracle to find a y_{k+1}, via self-reducibility, that beats all $Z'_{k,i}$, $0 \leq i \leq k$. Such y_{k+1} certainly exists since $x \in L$, and, since all such y beat at least $3/4$ of $Z_{k,i}$, (6) is satisfied with this y_{k+1} for all $0 \leq i \leq k$. Now it follows from Lemma 3 that the sequence $y_1, \ldots, y_m$ *is* an irrefutable proof w.r.t. Q in the strong sense of (5). Thus except with exponentially small probability $O(n^2/2^n)$ we find an irrefutable proof w.r.t. Q. One more query to **SAT** confirms this.

This completes the proof of Theorem 3 modulo Lemma 3, to which we turn next. Our proof of Lemma 3 will be probabilistic in nature. It should be pointed out that this use of probability has nothing to do with the probabilistic construction of $\mathcal{Z}_k$ in the proof of Theorem 3. The statement of Lemma 3 is completely deterministic.

We define an ensemble of r.v. $\{\tilde{c}_k(z) : z \in Z\}_{k\geq 0}$ where for each $k \geq 0$, the family $\{\tilde{c}_k(z) : z \in Z\}$ is i.i.d. and defined as follows: $\forall z \in Z$, $\tilde{c}_0(z) = 0$, and if $\tilde{c}_k(z) = i$ then $\tilde{c}_{k+1}(z) = i+1$ or i with probability $p_{k,i}$ and $1 - p_{k,i}$ respectively. Let $\tilde{\mathcal{Z}}_k = \{\tilde{Z}_{k0}, \tilde{Z}_{k1}, \ldots, \tilde{Z}_{kk}\}$ be defined as follows: For $0 \leq i \leq k$,

$$\tilde{Z}_{k,i} = \{z \in Z \mid \tilde{c}_k(z) = i\}.$$

We can show that

Claim: The expectation $\mathbf{E}|\tilde{Z}_{k,i}| = |Z_{k,i}|$, for all $k \geq 0$ and $0 \leq i \leq k$.

To prove this claim, we induct on k, the case $k = 0$ being trivial. Suppose the claim holds for k and for all $0 \leq i \leq k$. Consider $k+1$ and $1 \leq i \leq k+1$. The case $\mathbf{E}|\tilde{Z}_{k+1,0}| = |Z_{k+1,0}|$ follows from the rest, and the fact that the total cardinality is 2^n.

Denote by $\mathbf{E}_{\leq k}$ the expectation taken w.r.t. stages up to k. Since $|\tilde{Z}_{k+1,i}| = \sum_{z \in Z} \chi_{[z \in \tilde{Z}_{k+1,i}]}$, it follows that, for $1 \leq i \leq k+1$,

$$\begin{aligned}
\mathbf{E}|\tilde{Z}_{k+1,i}| &= \sum_{z \in Z} \mathbf{E}[\chi_{[z \in \tilde{Z}_{k+1,i}]}] \\
&= \sum_{z \in Z} \mathbf{E}\left[\chi_{[z \in \tilde{Z}_{k,i}]} \cdot \chi_{[z \in \tilde{Z}_{k+1,i}]} + \chi_{[z \in \tilde{Z}_{k,i-1}]} \cdot \chi_{[z \in \tilde{Z}_{k+1,i}]}\right] \\
&= \sum_{z \in Z} \left\{ \mathbf{E}_{\leq k}\left[\chi_{[z \in \tilde{Z}_{k,i}]}\right] \cdot (1 - p_{k,i}) + \mathbf{E}_{\leq k}\left[\chi_{[z \in \tilde{Z}_{k,i-1}]}\right] \cdot (p_{k,i-1}) \right\} \\
&= (1 - p_{k,i})\mathbf{E}_{\leq k}|\tilde{Z}_{k,i}| + p_{k,i-1}\mathbf{E}_{\leq k}|\tilde{Z}_{k,i-1}| \\
&= (1 - p_{k,i})|Z_{k,i}| + p_{k,i-1}|Z_{k,i-1}| \\
&= |Z_{k+1,i}|.
\end{aligned}$$

We next define a second ensemble of r.v. $\{\underline{c}_k(z) : z \in Z\}_{k \geq 0}$, where again, for fixed $k \geq 0$, the family $\{\underline{c}_k(z) : z \in Z\}$ is i.i.d. and defined simply as the sum of k Bernoulli independent 0-1 variables with $p = 3/4$. More formally, $\underline{c}_k(z) = \sum_{j=1}^{k} I_j(z)$, where $I_j(z)$ are i.i.d. 0-1 variables with $\Pr[I_j(z) = 1] = 3/4$. Then $\underline{\mathcal{Z}}_k = \{\underline{Z}_{k0}, \ldots, \underline{Z}_{kk}\}$ is defined:

$$\underline{Z}_{k,i} = \{z \in Z \mid \underline{c}_k(z) = i\}.$$

We can "realize" $\tilde{\mathcal{Z}}_k$ via $\underline{\mathcal{Z}}_k$ by a "nibbling" process. Note that $\underline{c}_0(z) = 0$, and $\underline{c}_{k+1}(z) = \underline{c}_k(z) + I_k(z)$. Define a third ensemble $\{c_k^*(z) : z \in Z\}_{k \geq 0}$ via $\underline{c}_k(z)$ as follows: $c_0^*(z) = 0$, and $c_{k+1}^*(z) = c_k^*(z) + I_k(z) + \Delta$, where the "nibble" Δ is a 0-1 r.v. dependent on $c_k^*(z)$ and $I_k(z)$: If $I_k(z) = 1$ then $\Delta = 0$, if $I_k(z) = 0$, and $i = c_k^*(z)$, then $\Delta = 1$ with probability $4p_{k,i} - 3$, and $\Delta = 0$ with probability $4(1 - p_{k,i})$. Note that $0 \leq 4p_{k,i} - 3 \leq 1$. Given $c_k^*(z)$, the combined effect of $I_k(z) + \Delta$ is a Bernoulli 0-1 variable taking value 1 with probability exactly $p_{k,i}$, independent for every z.

Thus $c_k^*(z)$ has exactly the same distribution as $\tilde{c}_k(z)$. While $\tilde{c}_k(z)$ is independent from $\underline{c}_k(z)$, $c_k^*(z)$ is highly correlated with $\underline{c}_k(z)$: $\forall z, \forall k$,

$$\underline{c}_k(z) \leq c_k^*(z).$$

Thus, $\forall z, k, \ell$,

$$\Pr[\tilde{c}_k(z) \leq \ell] = \Pr[c_k^*(z) \leq \ell] \leq \Pr[\underline{c}_k(z) \leq \ell].$$

For $\underline{c}_k(z)$, the Chernoff bound (4) applies directly. Thus if $m \geq 7n$ and odd, we take $p = 3/4$ and $\delta = 1/4$ then a short calculation gives,

$$(\forall z)\ \Pr[\underline{c}_m(z) \leq \lfloor \frac{m}{2} \rfloor] \leq 2e^{-\frac{7}{6}n}.$$

Thus,

$$\sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} |Z_{m,i}| = \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \mathbf{E}|\tilde{Z}_{m,i}| = \sum_{z \in Z} \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \Pr[z \in \tilde{Z}_{m,i}] = \sum_{z \in Z} \Pr[\tilde{c}_m(z) \leq \lfloor \frac{m}{2} \rfloor] \leq 2^{n+1} e^{-\frac{7}{6}n} < 1.$$

But the cardinalities of the sets $Z_{m,i}$ are all non-negative integers, we must conclude that

$$Z_{m,0} = Z_{m,1} = \ldots = Z_{m, \lfloor \frac{m}{2} \rfloor} = \emptyset.$$

6 An Implication for Karp-Lipton

There has been a lot of work on the general theme inspired by the Karp-Lipton Theorem. For example, Mahaney [M80] showed that if the sparse oracle is itself in NP (i.e., NP has $\leq_T^p$-complete, not just $\leq_T^p$-hard sparse set) then PH collapses to Δ_2^p. Long [Lo82] extended this to co-sparse oracles. Arvind et. al. [AKSS95] showed that under the same assumption as in Karp-Lipton that **SAT** has small circuits then MA = AM. (See [HMO92] for a survey.)

Suppose NP has polynomial size circuits. The Karp-Lipton Theorem says that the Polynomial-time Hierarchy collapses to $\Sigma_2^p \cap \Pi_2^p$. Sengupta [Se00] pointed out that the same proof collapses the Polynomial-time Hierarchy to S_2^p. To see this we recount the "book" proof, but this time phrase it in terms of provers Y and Z. We only need to show that $\Pi_2^p \subseteq \mathrm{S}_2^p$, then it follows that $\Pi_2^p \subseteq \mathrm{S}_2^p \subseteq \Sigma_2^p$ and hence they are all equal.

Let L be any language in Π_2^p. There is a normal form $L = \{x \mid (\forall^p y)(\exists^p z)[P(x,y,z)]\}$, where P is a P-time predicate. By Cook's Theorem, without loss of generality we can assume that it takes the form

$$L = \{x \mid (\forall^p s)[\phi_{x,s} \in \mathbf{SAT}]\},$$

where $\phi_{x,s}$ is a boolean formula computable in P-time from x and s. Let the size of $\phi_{x,s}$ be bounded by $p(|x|)$ for some polynomial $p(\cdot)$.

Now to show membership in S_2^p we receive two strings y and z, from provers Y and Z respectively. We expect the string y to be a poly-size circuit for formulae of size up to $p(|x|)$. For a pair (y,z) we accept if and only if the circuit y says the boolean formula $\phi_{x,z}$ is satisfiable and by self-reducibility produced a satisfying assignment which satisfied it.

We note that there exists a relativized world where the Karp-Lipton Theorem cannot be improved to P^{NP} [H86, W85].

If one substitutes the predicate P in the definition (1)(2) of S_2^p by a predicate computable in NP $\cap$ co-NP, we get the class $\mathrm{S}_2[\mathrm{NP} \cap \text{co-NP}]$, and we can still prove the inclusion $\mathrm{S}_2[\mathrm{NP} \cap \text{co-NP}] \subseteq \mathrm{ZPP}^{\mathrm{NP}}$. Clearly $\mathrm{S}_2^p \subseteq \mathrm{S}_2[\mathrm{NP} \cap \text{co-NP}]$. It is open whether any of the following containments

$$\mathrm{S}_2^p \subseteq \mathrm{S}_2[\mathrm{NP} \cap \text{co-NP}] \subseteq \mathrm{ZPP}^{\mathrm{NP}}$$

is a proper containment. We note that under suitable hardness assumptions one can prove $\mathrm{P}^{\mathrm{NP}} = \mathrm{BPP}^{\mathrm{NP}}$ (see [KvM99]) and thus under these assumptions the above classes all collapse to P^{NP}.

Acknowledgement

I thank Venkat Chakaravarthy, Oded Goldreich, Lane Hemaspaandra, Alex Russell, Uwe Schöning, Alan Selman and Samik Sengupta for interesting discussions and comments. I also thank the anonymous referee for very helpful comments.

References

[AK97] V. Arvind and J. Köbler, *On Pseudorandomness and Resource-Bounded Measure*, Proceedings of Conference on the Foundations of Software Technology and Theoretical Computer Science, volume 1346 of Lecture Notes in Computer Science, pages 235-249. Springer-Verlag (1997).

[AK00] V. Arvind and J. Köbler, *Graph isomorphism is low for* $\mathrm{ZPP}^{\mathrm{NP}}$ *and other lowness results*, Proceedings of Annual Symposium on Theoretical Aspects of Computer Science (STACS) 431-442 (2000).

[AK02] V. Arvind and J. Köbler, *New Lowness Results for* $\mathrm{ZPP}^{\mathrm{NP}}$ *and Other Complexity Classes*. *The Journal of Computer and System Sciences* (JCSS) 65(2): 257-277 (2002).

[B85] L. Bábai, *Trading group theory for randomness*, Proceedings of ACM Symposium On Theory of Computing (STOC) 17:421–429(1985).

[BM88] L. Bábai and S. Moran, *Arthur-Merlin Games: A Randomized Proof System, and a Hierarchy of Complexity Classes*. *The Journal of Computer and System Sciences* (JCSS) 36(2): 254-276 (1988).

[AKSS95] V. Arvind, J. Köbler, U. Schöning and R. Schuler. *If NP has polynomial size circuits then MA=AM*, *Theoretical Computer Science* vol 137, 279-282 (1995).

[BM88] L. Bábai and S. Moran, *Arthur-Merlin Games : a randomized proof system, and a hierarchy of complexity classes*, *The Journal of Computer and System Sciences*, (JCSS) 36: 254-276 (1988).

[BCGKT94] N. Bshouty, R. Cleve, S. Kannan, R. Gavaldà and C. Tamon, *Oracles and Queries that are sufficient for Exact Learning*, Proceedings of the 17th Annual ACM conference on Computational Learning Theory, 130–19 (1994). *The Journal of Computer and System Sciences*, (JCSS) 52(3): 421–433 (1996).

[BGP00] M. Bellare, O. Goldreich and E. Petrank. *Uniform Generation of NP-witnesses using an NP-oracle*, *Inform. and Comp.*, Vol. 163, 510–526 (2000).

[C01] Jin-Yi Cai. $\mathrm{S}_2^p \subseteq \mathrm{ZPP}^{\mathrm{NP}}$.
Proc. 42nd IEEE Symposium on Foundations of Computer Science (FOCS), 620–628 (2001).

[C96] R. Canetti. *On BPP and the Polynomial-time Hierarchy. Information Processing Letters*, 57, pages 237–241 (1996).

[GMR85] S. Goldwasser, S. Micali and C. Rackoff, *The Knowledge Complexity of Interactive Proofs.* Proc. 17th ACM Symp. om Computing, Providence, RI, 1985, pp. 291-304. *SIAM J. Comput.* 18(1): 186-208 (1989).

[GS89] S. Goldwasser and M. Sipser, *Private coins versus public coins in interactive proof systems.* Proceedings of ACM Symposium On Theory of Computing (STOC) 18:59–68(1986). *Randomness and Computation*, S. Micali, editor, volume 5 of Advances in Computing Research, pages 73-90. JAI Press, Greenwich (1989).

[GZ97] O.Goldreich and D.Zuckerman, *Another Proof that BPP $\subseteq$ PH (and more).* Electronic Colloquium on Computational Complexity (ECCC), TR97-045, (1997).

[H86] H. Heller, *On relativized exponential and probabilistic complexity classes. Information and Control* 71 (3), 231-243 (1986).

[HMO92] L. A. Hemachandra, M. Ogiwara and O. Watanabe, *How Hard Are Sparse Sets?* Proceedings of Structure in Complexity Theory Conference: 222-238 (1992)

[H81] J. Hopcroft, *Recent Directions in Algorithmic Research.* Proceedings 5th GI Conference on Theoretical Computer Science, Springer-Verlag *Lecture Notes in Computer Science #104*, 123–134 (1981).

[JVV86] M. Jerrum, L. G. Valiant, V. V. Vazirani. *Random Generation of Combinatorial Structures from a Uniform Distribution. Theoretical Computer Science*, 43: 169-188 (1986).

[KL80] R. Karp and R. Lipton. *Some connections between nonuniform and uniform complexity classes.* Proceedings of the 12th ACM Symposium on Theory of Computing, 302–309. ACM Press, April 1980. An extended version has also appeared as: Turing machines that take advice, *L'Enseignement Mathématique*, 2nd series, 28, 191–209 (1982).

[KvM99] A. Klivans and D. van Melkebeek. *Graph Nonisomorphism has Subexponential Size Proofs Unless the Polynomial-Time Hierarchy Collapses.* Proceedings of ACM Symposium On Theory of Computing (STOC) 659-667 (1999).

[KW95] J. Köbler and O. Watanabe, *New collapse consequences of NP having small circuits.* ICALP *Lecture Notes in Computer Science* (LNCS) 944:196–207(1995). Journal version *SIAM J. Comput.* 28(1): 311-324 (1998).

[L83] C. Lautemann. *BPP and the Polynomial Hierarchy. Information Processing Letters* 17(4): 215-217 (1983).

[Lo82] T. Long. *A note on sparse oracles for NP. The Journal of Computer and System Sciences* (JCSS) vol 24, No. 2, pp 224–232 (1982).

[M80] S. Mahaney. *Sparse complete sets for NP: Solution of a conjecture of Berman and Hartmanis.* Proceedings of 21st IEEE Symposium of Foundations of Computer Science, (1980), pp 54–60. *J. Comput. System Sci.*, 25(2):130–143 (1982).

[OL93] M. Ogiwara and A. Lozano, *Sparse hard sets for counting classes. Theoretical Computer Science* 112, 255–276 (1993).

[OW91] M. Ogiwara and O. Watanabe, *On polynomial time bounded truth-table reducibility of NP sets to sparse sets. SIAM Journal on Computing* 20, 471–483 (1991).

[RS95] A. Russell and R. Sundaram. *Symmetric Alternation Captures BPP. Computational Complexity* 7(2): 152-162 (1998). A Preliminary version appeared in Technical Report MIT-LCS-TM-541, (1995).

[Se00] S. Sengupta, Personal communications, (2000).

[Si83] M. Sipser, *A Complexity Theoretic Approach to Randomness.* Proceedings of ACM Symposium On Theory of Computing (STOC) 330-335 (1983).

[St77] L. Stockmeyer. *The polynomial-time hierarchy. Theoretical Computer Science*, 3:1–22 (1977).

[St83] L. Stockmeyer. *The Complexity of Approximate Counting* (Preliminary Version). Proceedings of ACM Symposium On Theory of Computing (STOC) 1983: 118-126. Journal version: *On Approximation Algorithms for #P. SIAM J. Comput.* 14(4): 849-861 (1985).

[W85] C. B. Wilson, *Relativized circuit complexity. The Journal of Computer and System Sciences* (JCSS), 31, 169–181 (1985).

[ZF87] S. Zachos and M. Fürer, *Probabilistic quantifiers vs Distrustful adversaries.* Foundations of Software Technology and Theoretical Computer Science (FSTTCS) LNCS-287:449–455 (1987).

[Z96] D. Zuckerman. *Simulating BPP Using a General Weak Random Source. Algorithmica*, 16(4/5): 367-391 (1996)