Another Proof that $\mathcal{BPP} \subseteq \mathcal{PH}$ (and more)

Oded Goldreich David Zuckerman

March 11, 2016

Abstract

We provide another proof of the Sipser–Lautemann Theorem by which $\mathcal{BPP} \subseteq \mathcal{MA}$ ($\subseteq \mathcal{PH}$). The current proof is based on strong results regarding the amplification of $\mathcal{BPP}$, due to Zuckerman (1996). Given these results, the current proof is even simpler than previous ones. Furthermore, extending the proof leads to two results regarding $\mathcal{MA}$: $\mathcal{MA} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$ (which seems to be new), and that two-sided error $\mathcal{MA}$ equals $\mathcal{MA}$. Finally, we survey the known facts regarding the fragment of the polynomial-time hierarchy that contains $\mathcal{MA}$.

An early version of this work appeared as TR97-045 of *ECCC*. The current revision is quite minimal.

Keywords: BPP, The Polynomial-Time Hierarchy, Interactive Proof Systems (AM and MA), Randomness–Efficient Error Reduction (Amplification).

Note. This paper appeared in the volume *Studies in Complexity and Cryptography*, Springer LNCS, Vol. 6650, 2011. The current version is a post-publication revision, correcting a missing credit.

1 Introduction

Non-trivial results, showing containment of fundamental complexity classes in one another, are quite rare. One of the first such results is Sipser's Theorem [14] by which $\mathcal{BPP}$ is contained in the Polynomial-Time Hierarchy. A simpler proof, placing $\mathcal{BPP}$ even lower in this hierarchy, was presented by Lautemann [11]. Although not stated in these (subsequently introduced) terms, Lautemann's proof actually establishes the following:

Theorem 1 (The Sipser–Lautemann Theorem): $\mathcal{BPP} \subseteq \mathcal{MA}$.

See definitions in next section.

The contents of this note. In this note, we present an alternative proof of the Sipser–Lautemann Theorem. Our proof relies on powerful results regarding randomness–efficient error reduction (a.k.a amplification) for $\mathcal{BPP}$. Given these powerful results, our proof is almost a triviality.

Using similiar arguments, we show that $\mathcal{MA} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$ (re-establishing a theorem of Zachos and Heller [16] by which $\mathcal{BPP} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$). It follows that $\mathcal{NP}^{\mathcal{BPP}} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$ (re-establishing a

theorem of Zachos and Furer [15]). To the best of our knowledge, the first result was not known before.

In summary, the purpose of this note is three-fold: Firstly to demonstrate the power of the currently known results regarding randomness–efficient error reduction. We believe that these results have not been fully assimilated into complexity theory and are yet to be exploited by it. Secondly we wish to focus attention on the fragment of the polynomial-time hierarchy that contains $\mathcal{MA}$. It seems that this fragment gives rise to some challenges which may be within our current reach. Finally, we take the oppertunity to prove the aforementioned new result.

Organization. The core of this work (i.e., the alternative proof of the Sipser–Lautemann Theorem) is presented in Sections 2 and 3.1. This alternative proof is further discussed in Section 3.2, and applied in the context of two-sided MA in Section 3.3. The same proof strategy is then applied to show that $\mathcal{MA}$ is contained in $\mathcal{ZPP}^{\mathcal{NP}}$ (see Section 4). Finally, we conclude with a brief survey of the complexity classes around $\mathcal{MA}$ (see Section 5).

2 Background

(For further background, see Section 5.)

2.1 BPP and randomness-efficient error reduction

Definition 1 (the class BPP): *For any set S, we denote by χ_S the characteristic function of the set; that is, $\chi_S(x) = 1$ if $x \in S$ and $\chi_S(x) = 0$ otherwise. A set S is in $\mathcal{BPP}$ if there exists a probabilistic polynomial-time machine M such that for every $x \in \{0,1\}^*$*

$$\mathrm{Prob}[M(x) \neq \chi_S(x)] \leq \frac{1}{3}$$

where the probability is taken uniformly over the internal coin tosses of M.

The error probability in the foregoing procedure can be reduced by repetitions (a process hereafter referred to as *amplification*). The obvious way of doing so transforms a machine (as above) that, on input x, uses $p(|x|)$ coins into a machine having error probability at most $2^{-t(|x|)}$ that uses $O(t(|x|) \cdot p(|x|))$ coins (for any polynomial t). More efficient amplification procedures, utilizing Expander Random Walks, yield the same error bound while using only $p(|x|) + (4 + o(1)) \cdot t(|x|)$ coins (see survey [6]). In particular, *for any constant $c > 4$, using a sufficiently large polynomial t, we get a procedure that uses $c \cdot t(|x|)$ coins and has error probability at most $2^{-t(|x|)}$*. An alternative construction due to Zuckerman [17] provides, *for any constant $c > 1$* and sufficiently large polynomial t, a procedure that *uses $c \cdot t(|x|)$ coins and has error probability at most $2^{-t(|x|)}$*. What is remarkable in the last procedure is that the number of coins used is essentially the logarithm of the error bound. Put in other words, the number of "bad" coin sequences can be made any (constant) root of the total number of coin sequences. In particular,

Theorem 2 (Zuckerman's randomness-efficient amplification of BPP [17]): *For any set S in $\mathcal{BPP}$, there exists a polynomial-time recognizable binary relation R and a polynomial p such that*

$$|\{r \in \{0,1\}^{p(|x|)} : R(x,r) \neq \chi_S(x)\}| < 2^{p(|x|)/3}.$$

2.2 The complexity class MA

Definition 2 (the class MA): *A set S is in $\mathcal{MA}$ if there exists a polynomial-time recognizable 3-ary relation V and polynomials p, q such that*

- *If $x \in S$, then there exists $w \in \{0,1\}^{q(|x|)}$ such that for every $r \in \{0,1\}^{p(|x|)}$ it holds that $V(x,w,r) = 1$.*
- *If $x \notin S$, then for every $w \in \{0,1\}^{q(|x|)}$ it holds that*

$$\mathrm{Prob}_r[V(x,w,r) = 1] \;\leq\; \frac{1}{2}$$

where the probability is taken uniformly over all $r \in \{0,1\}^{p(|x|)}$.

The class $\mathcal{MA}$, introduced by Babai [1], consists of sets having a Merlin–Arthur proof system: The prover (Merlin) sends a certificate (denoted w above) to the verifier (Arthur) who assesses it probabilistically (by tossing coins r and applying the predicate V). Merlin–Arthur proof systems are a degenerate type of interactive proof systems (introduced by Goldwasser, Micali and Rackoff [8] and Babai [1]). Actually, in a Merlin–Arthur proof system there is no real interaction. Instead, it is instructive to view $\mathcal{MA}$ as *the* randomized version of $\mathcal{NP}$: Here the "certificates" (for membership) can be verified via a randomized procedure and errors may occur (yet with bounded probability).

3 A Proof of the Sipser–Lautemann Theorem

3.1 The proof itself

Using Zuckerman's efficient amplification of BPP, we present the following MA proof system. Specifically, we will refer to the relation R and the polynomial p guaranteed in Theorem 2.

The protocol. On input x, both parties compute $m = p(|x|)$, and proceed as follows.

1. Merlin tries to select $r' \in \{0,1\}^{m/2}$ such that $R(x, r'r'') = 1$ for all $r'' \in \{0,1\}^{m/2}$. Merlin sends r' to Arthur.
2. Upon receiving r', Arthur selects $r'' \in \{0,1\}^{m/2}$ uniformly and accepts if and only if $R(x, r'r'') = 1$.

Analysis of the foregoing protocol. If $x \in S$, then there are at most $2^{m/3}$ possible r's for which $R(x,r) = 0$. Thus there are at most $2^{m/3}$ prefixes $r' \in \{0,1\}^{m/2}$ for which some r'' exists so that $R(x, r'r'') = 0$. Merlin may just select any of the other $2^{m/2} - 2^{m/3}$ prefixes and make Arthur always accept. On the other hand, if $x \notin S$, then there are at most $2^{m/3}$ possible r's for which $R(x,r) = 1$. Thus, for each $r' \in \{0,1\}^{m/2}$, it holds that

$$\mathrm{Prob}_{r'' \in \{0,1\}^{m/2}}[R(x, r'r'') = 1] \;\leq\; \frac{2^{m/3}}{2^{m/2}} \;\ll\; \frac{1}{2}.$$

3.2 Discussion

Let us review our proof strategy. Starting with Theorem 2, we partitioned the space of all (2^m) possible coin-tosses outcomes into $(2^{m/2})$ subsets of equal size. We then used the following two facts:

1. The number of bad outcomes is smaller than the number of subsets (and so there exists a subset with no bad outcomes). This was used to analyze the case $x \in S$.
2. The number of bad outcomes is much smaller than the size of each subset (and so each subset contains a majority of good outcomes). This was used to analyze the case $x \notin S$.

Thus, what we have used is the fact that number of bad outcomes is much smaller than the square root of the total number of outcomes. We stress that the fact that any BPP-machine can be transformed into a machine for which the foregoing holds (i.e., Theorem 2) is highly non-trivial. We believe that this fact (or known generalizations of it) may find further applications in complexity theory.

Comparison to Lautemann's proof. Recall that Lautemann's proof has the prover send the verifier $t = m/\log_2 m$ strings, $s_1, ..., s_t$, and the verifier tosses coins $r \in \{0,1\}^m$ and accepts iff $R(x, r \oplus s_i) = 1$ holds for some i. The existence of an appropriate sequence of strings is proven by an elementary probabilistic argument. Actually, s_1 may be any fixed string (e.g., 0^m) and so needs not be sent (by the prover). We observe that IF we start with R as guaranteed by Theorem 2, then $t = 2$ suffices. This gets us very close to the proof above. In fact, the probabilistic argument of Lautemann reduces to the trivial counting argument above. Thus, using Theorem 2 allows also a simplification of Lautemann's argument, although the proof presented earlier is believed to be simpler: Technically speaking, we have the prover send only $m/2$ bits (rather than m required in the simplified Lautemann's argument), the verifier tosses only $m/2$ coins (again, rather than m), and the predicate R is evaluated only once (rather than twice).

3.3 Two-sided error equals one-sided error for MA

Both Lautemann's proof and our proof can be extended to show that a two-sided error version of $\mathcal{MA}$ equals the one-sided error defined above. (This provides an alternative proof to the one presented in [15].) We mention that interactive proof systems with zero error collapse to $\mathcal{NP}$, whereas for all (higher than MA) levels of the interactive proof hierarcy, the two-sided error version equals the one-sided one [5].

Definition 3 (two-sided version of MA): *A set S is in $\mathcal{MA}_2$ if there exists a polynomial-time recognizable 3-ary relation V and polynomials p, q such that*

- *If $x \in S$, then there exists $w \in \{0,1\}^{q(|x|)}$ such that*

$$\mathrm{Prob}_r[V(x, w, r) = 1] \geq \frac{2}{3}.$$

- *If $x \notin S$, then for every $w \in \{0,1\}^{q(|x|)}$ it holds that*

$$\mathrm{Prob}_r[V(x, w, r) = 0] \geq \frac{2}{3}.$$

In both cases, the probability is taken uniformly over all $r \in \{0,1\}^{p(|x|)}$.

Theorem 3 [15, Thm 2(i)]: $\mathcal{MA} = \mathcal{MA}_2$.

Proof: Clearly, $\mathcal{MA} \subseteq \mathcal{MA}_2$, and so we focus on showing that $\mathcal{MA}_2 \subseteq \mathcal{MA}$. Let S be an arbitrary set in $\mathcal{MA}_2$. For every $x \in S$, we consider w as guaranteed by the first condition of Definition 3, whereas for $x \notin S$ we consider any $w \in \{0,1\}^{q(|x|)}$. Both Lautemann's proof and our proof extend to promise problems in BPP, and in particular to the following BPP promise problem, $\Pi = (\Pi_{\mathrm{YES}}, \Pi_{\mathrm{NO}})$, where

$$\begin{aligned}
\Pi_{\mathrm{YES}} &\stackrel{\mathrm{def}}{=} \left\{(x,w) : \mathrm{Prob}_r[V(x,w,r)=1] \ \geq \ \frac{2}{3}\right\} \\
\Pi_{\mathrm{NO}} &\stackrel{\mathrm{def}}{=} \{(x,w) : x \notin S\} \\
&\subseteq \left\{(x,w) : \mathrm{Prob}_r[V(x,w,r)=0] \ \geq \ \frac{2}{3}\right\}
\end{aligned}$$

In particular, the amplification technique of Zuckerman applies also to this case and so we obtain a predicate V' and a polynomail q' such that

$$\forall (x,w) \in \Pi_{\mathrm{YES}} \qquad |\{r \in \{0,1\}^{q'(|x|)} : V'(x,w,r)=0\}| < 2^{q'(|x|)/3} \tag{1}$$

$$\forall (x,w) \in \Pi_{\mathrm{NO}} \qquad |\{r \in \{0,1\}^{q'(|x|)} : V'(x,w,r)=1\}| < 2^{q'(|x|)/3} \tag{2}$$

Thus, we augment the MA-protocol of Section 3.1 as follows. On input x, with $m = q'(|x|)$, Merlin sends (w,r'), where $|r'| = m/2$, and Arthur uniformly selects $r'' \in \{0,1\}^{m/2}$ and accepts if and only if $V'(x,w,r'r'') = 1$. As before, in case $x \in S$, by sending an adequate (w,r'), Merlin can make Arthur accept for every choice of r''; whereas, in case $x \notin S$, for any choice of (w,r'), Arthur accepts with negligible probability. It follows that $S \in \mathcal{MA}$. ■

4 MA is Contained in ZPP with an NP-oracle

The machines in the following definition may halt with a non-Boolean output (which may be interpreted as abstaining from a decision regarding membership).

Definition 4 (the class ZPP): *A set* S *is in* $\mathcal{ZPP}$ *if there exists a probabilistic polynomial-time machine* M *such that for every* $x \in \{0,1\}^*$

$$\begin{aligned}
\mathrm{Prob}[M(x) = \chi_S(x)] &\geq \frac{1}{2} \\
\mathrm{Prob}[M(x) = 1 - \chi_S(x)] &= 0
\end{aligned}$$

where the probability is taken uniformly over the internal coin tosses of M.

Thus, the ZPP machine either gives the correct answer or gives no answer at all (i.e., a non-Boolean output is interpreted as no output). Clearly $\mathcal{ZPP} = \mathcal{RP} \cap \mathrm{co}\mathcal{RP}$ (actualy, $\mathcal{ZPP}$ is sometimes defined this way).

4.1 BPP is contained in ZPP with an NP-oracle

We start by providing an alternative proof to a result of Zachos and Heller.

Theorem 4 [16, P. 132, Cor. 3]: $\mathcal{BPP} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$.

Proof: Using the same amplification and notations as in Section 3.1, we construct a probabilistic polynomial-time oracle machine, M, that on input x operates as follows (where $m = p(|x|)$):

1. Selects $\sigma \in \{0,1\}$ uniformly (as guess for $\chi_S(x)$);
2. Selects $r' \in \{0,1\}^{m/2}$ uniformly;
3. Queries the oracle on whether (x, σ, r') is in the following co$\mathcal{NP}$ set
$$\{(y, \tau, u) : \forall v \in \{0,1\}^{|s|}\ ,\ R(y, uv) = \tau\}. \tag{3}$$
4. If the oracle answers YES, then the machine outputs σ. Otherwise it halts with no output.

Recall that by the foregoing amplification, for any x, the following holds:

- For each r', it holds that
$$|\{r'' \in \{0,1\}^{m/2} : R(x, r'r'') \neq \chi_S(x)\}| < 2^{m/2},$$
and so the oracle never answers YES on query $(x, 1 - \chi_S(x), r')$. Thus, the machine never outputs the wrong answer.
- On the other hand, it holds that
$$\mathrm{Prob}_{r'}[\forall r'' \in \{0,1\}^{m/2}\ ,\ R(x, r'r'') = \chi_S(x)] > \frac{1}{2}$$
and so with probability at least 1/4, over the choices of σ and r', the oracle answers YES (and the machine produces a (correct) 0-1 output).

Using straightforward amplification, the theorem follows. ■

4.2 Extension to MA

Combining ideas from the last two proofs, we obtain.

Theorem 5 (seemingly new): $\mathcal{MA} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$.

Observing that $\mathcal{NP}^{\mathcal{BPP}} \subseteq \mathcal{MA}_2$ (see Fact 6), and using Theorems 3 and 5, we conclude that $\mathcal{NP}^{\mathcal{BPP}} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$.

Fact 6 (folklore) $\mathcal{NP}^{\mathcal{BPP}} \subseteq \mathcal{MA}_2$.

Proof: Let $S \in \mathcal{NP}^{\mathcal{BPP}}$. Then, for every $x \in S$, we instruct Merlin to send a transcript of an accepting computation of the non-deterministic polynomial-time oracle-machine, and instruct Arthur to verify the validity of transcript as well as the correctness of the the oracle answers (by running a probabilistic decision procedure of negligible two-sided error). ■

Proof of Theorem 5: Let $S \in \mathcal{MA}$ and consider the same promise problem $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ as in the proof of Theorem 3. Furthermore, consider the set $\Pi'_{\text{YES}} \subseteq \Pi_{\text{YES}}$ that consists of all pairs (x, w) such that for all $r \in \{0,1\}^{p(|x|)}$ it holds that $V(x, w, r) = 1$, and recall that for every $x \in S$ there exists $w \in \{0,1\}^{q(|x|)}$ such that $(x, w) \in \Pi'_{\text{YES}}$.

We construct a probabilistic polynomial-time oracle machine, M, that on input x and access to an NP-oracle, first attempts to find w such that $(x, w) \in \Pi_{\text{YES}}$, and next verifies that $(x, w) \in \Pi_{\text{YES}}$ indeed holds. Following is a detailed description of the operation of M (as well as key observations towards its analysis). On input x, where $n = q(|x|)$ and $k = p(|x|)$, machine M proceeds as follows.

Step 1: Attempting to find a good w. The machine uniformly selects $r_1, ..., r_{2n} \in \{0,1\}^k$, and queries the NP-oracle on whether there exists a $w \in \{0,1\}^n$ such that $\wedge_{i=1}^{2n} V(x, w, r_i) = 1$. If the answer is NO, then M halts with output 0, otherwise M iteratively recovers the bits of such a string w (by $|w|$ additional queries) and proceeds to the next step. Specifically, all queries have the form $(x, w', r_1, ..., r_{2n})$, and each such query is answered by a YES if and only if there exists a $w'' \in \{0,1\}^{n-|w'|}$ such that $\wedge_{i=1}^{2n} V(x, w'w'', r_i) = 1$.

Note that if $x \in S$, then a string w such that $\wedge_{i=1}^{2n} V(x, w, r_i) = 1$ exists (e.g., consider w such that $(x, w) \in \Pi'_{\text{YES}}$), and so Step 1 must be completed while finding such a string w. On the other hand, for each $(x, w) \notin \Pi_{\text{YES}}$, the probability that $\wedge_{i=1}^{2n} V(x, w, r_i) = 1$ holds, where $r_1, ..., r_{2n}$ are selected uniformly in $\{0,1\}^k$, is at most $(2/3)^{2n}$, and it follows that

$$\text{Prob}_{r_1,...,r_{2n}}[\exists w \text{ s.t. } (x, w) \notin \Pi_{\text{YES}} \text{ and } \wedge_{i=1}^{2n} V(x, w, r_i) = 1]] \leq 2^n \cdot (2/3)^{2n},$$

which is exponentially vanishing (in n).

Step 2: Verifying that w is good (i.e., $(x, w) \in \Pi_{\text{YES}}$). The machine treats (x, w) as an input to the promise problem Π and proceeds as in the proof of Theorem 4. Specifically, by using the same amplification as in the proof of Theorem 3, we obtain a verification procedure V' that satisfies Eq. (1)-(2). Letting $m = q'(|x|)$, machine M selects an $m/2$-bit long random prefix r', and queries the NP-oracle on whether all $m/2$-bit long suffixes make the predicate V' evaluate to 1 (i.e., whether for every r'' it holds that $V'(x, w, r'r'') = 1$). If the oracle answers YES, then M halts with output 1; otherwise, M halts with no output. (We stress that we never output 0 in this step.)

If $x \in S$, then Step 1 never halts but rather always yields a string w (for Step 2). Furthermore, with overwhelmingly high probability, the string w satisfies $(x, w) \in \Pi_{\text{YES}}$. Thus, with overwhelmingly high probability, Step 2 accepts. On the other hand, if $x \notin S$, then with overwhrlmingly high probability Step 1 halts (with output 0). Furthermore, if the procedure continued to Step 2 with some string w, then $(x, w) \in \Pi_{\text{NO}}$ (since $x \notin S$). In this case, the oracle will always answer NO, and M will halt with no output. Thus, for any x, the machine never errs, and with overwhelmingly high probability it produces the correct output. ■

5 The Bigger Picture: Complexity Classes Around MA

(For a wider perspective on interactive proofs, see [7, Sec. 9.1],)

5.1 Definitions

All the (binary and 3-ary) relations that are mentioned in the following definitions are only satisfied by arguments of polynomially related length (i.e., all tuples in a relation have arguments that are of length that is polynomial in the length of the first argument). Likewise, all quantifiers range over arguments of such lengths.

Definition 5 (traditional classes – classes of the 1970's:)

- *A set S is in $\Sigma_2^P = \mathcal{NP}^{\mathcal{NP}}$ (resp., $\Pi_2^P = \mathrm{co}\mathcal{NP}^{\mathcal{NP}}$) if there exists a polynomial-time recognizable 3-ary relation R such that*

$$\begin{aligned} S &= \{x : \exists y \forall z\ R(x,y,z) = 1\} \\ (\textit{resp.},\ S &= \{x : \forall y \exists z\ R(x,y,z) = 1\}). \end{aligned}$$

- *A set S is in $\Delta_2^P = \mathcal{P}^{\mathcal{NP}}$ if there exists a deterministic polynomial-time oracle machine M and a set $S' \in \mathcal{NP}$ such that $x \in S$ iff $M^{S'}(x) = 1$ $(\forall x)$.*

- *A set S is in $\mathcal{RP}$ if there exists a probabilistic polynomial-time machine M such that*

$$\begin{aligned} x \in S &\implies \mathrm{Prob}[M(x) = 1] \geq \frac{1}{2} \\ x \notin S &\implies \mathrm{Prob}[M(x) = 1] = 0 \end{aligned}$$

For any class $\mathcal{C}$, we define $\mathrm{co}\mathcal{C} \stackrel{\mathrm{def}}{=} \{\{0,1\}^* \setminus S : S \in \mathcal{C}\}$.

Definition 6 ($\mathcal{AM}$ [1] – a class of the 1980's:) *A set S is in $\mathcal{AM}$ if there exists a polynomial-time recognizable 3-ary relation V and polynomials p, q such that*

- *If $x \in S$, then for every $r \in \{0,1\}^{p(|x|)}$ there exists $w \in \{0,1\}^{q(|x|)}$ such that $V(x,r,w) = 1$.*
- *If $x \notin S$, then it holds that*

$$\mathrm{Prob}_r[\exists w \text{ s.t. } V(x,r,w) = 1] \leq \frac{1}{2}$$

where the probability is taken uniformly over all $r \in \{0,1\}^{p(|x|)}$.

In other words, the class $\mathcal{AM}$, introduced by Babai [1], consists of sets having an Arthur–Merlin proof systems: The verifier (Arthur) challenges the prover (Merlin) with a random query, denoted r, and given the prover's answer (denoted w) makes a decision using the predicate V. Thus, in contrast to Merlin–Arthur systems (where Arthur just (probabilistically) evaluates the validity of a "written proof"), in Arthur–Merlin systems we have a real interaction between the prover and the verifier. The class $\mathcal{AM}$ coincides with the class of sets having constant-round interactive proof systems [1, 9]. Thus, it is the lowest level of the hierarchy of "real" interactive proofs [1, 8] (i.e., interactive proofs that, unlike $\mathcal{NP}$ and $\mathcal{MA}$, are really interactive).

Definition 7 ($\mathcal{S}_2^P$ [4, 13] – a class of the 1990's:) *S is in $\mathcal{S}_2^P$ if there exists a polynomial-time recognizable 3-ary relation R such that for every $x \in \{0,1\}^*$*

$$\exists y \forall z \qquad R(x,y,z) = \chi_S(x) \tag{4}$$

$$\exists z \forall y \qquad R(x,y,z) = \chi_S(x) \tag{5}$$

The class $\mathcal{S}_2^P$ was introduced independently by Canetti [4] and Russell and Sundaram [13] with the motivation of providing a low "symmetric alternation class" that contains $\mathcal{BPP}$. Indeed, Canetti [4] has extended Lautemann's proof to show that $\mathcal{BPP} \subseteq \mathcal{S}_2^P$, whereas Russell and Sundaram [13] showed that $\mathcal{MA} \subseteq \mathcal{S}_2^P$ (and thus $\mathcal{BPP} \subseteq \mathcal{S}_2^P$).

5.2 Known Inclusions

We recall some known inclusions between the aforementioned classes. For sake of self-containment, we present proofs as well. Recall that, $\mathcal{BPP} \subseteq \mathcal{MA}$, by Theorem 1. We start with some simple *syntactical facts*:

1. $\mathcal{P} \subseteq \mathcal{RP} \subseteq \mathcal{NP} \subseteq \mathcal{MA}$.
2. $\mathcal{RP} \subseteq \mathcal{BPP}$.
3. $\mathcal{RP} \subseteq \mathrm{co}\mathcal{MA}$ (equiv., $\mathrm{co}\mathcal{RP} \subseteq \mathcal{MA}$).[1]
4. $\mathcal{NP} \cup \mathrm{co}\mathcal{NP} \subseteq \mathcal{P}^{\mathcal{NP}}$.
5. $\mathcal{AM} \subseteq \Pi_2^P$.
6. $\mathcal{S}_2^P \subseteq \Sigma_2^P \cap \Pi_2^P$.

   (Actually, the transparent syntactical facts are the inclusion $\mathcal{S}_2^P \subseteq \Sigma_2^P$ and the closure of $\mathcal{S}_2^P$ under complement.)
7. $\mathcal{ZPP}^{\mathcal{NP}} \subseteq \Sigma_2^P \cap \Pi_2^P$.

   (Here the transparent facts are $\mathcal{ZPP}^{\mathcal{NP}} \subseteq \mathcal{RP}^{\mathcal{NP}} \subseteq \mathcal{NP}^{\mathcal{NP}} = \Sigma_2^P$.)

We now turn to three non-trivial results.

Proposition 7 [1]: $\mathcal{MA} \subseteq \mathcal{AM}$.

Proof: We use a naive amplification to reduce the error probability in the Merlin–Arthur game so to obtain error that is substantially smaller than the reciprocal of the number of possible Merlin messages. Specifically, we obtain a polynomial-time recognizable 3-ary relation V and polynomials p, q such that

1. If $x \in S$, then there exists $w_0 \in \{0,1\}^{q(|x|)}$ such that for every $r \in \{0,1\}^{p(|x|)}$ it holds that $V(x, w_0, r) = 1$.

[1] This syntactical fact can also be derived from $\mathcal{RP} \subseteq \mathcal{BPP}$, by using $\mathcal{BPP} \subseteq \mathcal{MA}$.

2. If $x \notin S$, then for every $w \in \{0,1\}^{q(|x|)}$ it holds that

$$\mathrm{Prob}_r[V(x,w,r)=1] \;<\; \frac{1}{2}\cdot 2^{-q(|x|)}.$$

Thus,

$$\begin{aligned}\mathrm{Prob}_r[\exists w \in \{0,1\}^{q(|x|)} : V(x,w,r)=1] \;&\leq\; \sum_{w\in\{0,1\}^{q(|x|)}} \mathrm{Prob}_r[V(x,w,r)=1] \\ &<\; \frac{1}{2}.\end{aligned}$$

We construct an Arthur–Merlin proof system (defined by a new predicate V') by merely reversing the order of moves in the foregoing proof system, and using essentially the same decision predicate as above: That is, we let $V'(x,r,w) \stackrel{\mathrm{def}}{=} V(x,w,r)$. This potentially makes the task of Merlin easier, and so we need only worry about the case $x \notin S$ (which we handle easily using the above bound). Specifically, for the case $x \in S$, we may use the string w_0 (guaranteed in Item 1) as Merlin's response to any challenge r (and so $V'(x,r,w_0) = V(x,w_0,r) = 1$ for all r's). For the case $x \notin S$ we use the bound in Item 2 and so $\mathrm{Prob}_r[\exists w \in \{0,1\}^{q(|x|)} : V'(x,r,w)=1] < 0.5$. The proposition follows. ■

Proposition 8 [13]: $\mathcal{MA} \subseteq \mathcal{S}_2^P$.

Proof: We use the same amplification as in the previous proof. Here we write the case of $x \notin S$ as

$$\forall w \in \{0,1\}^{q(|x|)} \qquad |\{r \in \{0,1\}^{p(|x|)} : V(x,w,r)=1\}| \;<\; 2^{p(|x|)-q(|x|)} - 1$$

We define a relation R (for the class $\mathcal{S}_2^P$) such that $R(x,y,z)=1$ if $|y|=|z|=q(|x|)$ and at least one of the following two conditions holds:

1. $y = w0^{p(|x|)-q(|x|)}$ and $V(x,w,z)=1$.
2. $z = w0^{p(|x|)-q(|x|)}$ and $V(x,w,y)=1$.

Clearly, this predicate is symmetric with respect to y and z; that is, condition (1) holds iff condition (2) holds. Thus, we only show, for any x, the existence of a string y such that, for all z's, $R(x,y,z) = \chi_S(x)$. Let us shorthand $m = p(|x|)$ and $n = q(|x|)$. For $x \in S$ there exists $w \in \{0,1\}^n$ such that for all $r \in \{0,1\}^m$ it holds that $V(x,w,r)=1$. Thus, there exists $y = w0^{m-n} \in \{0,1\}^m$ such that for all $z \in \{0,1\}^m$ it holds that $R(x,y,z)=1$. We now turn to the case where $x \notin S$: In this case,

$$\begin{aligned}|\{r : \exists w \text{ s.t. } V(x,w,r)=1\}| \;&\leq\; \sum_{w\in\{0,1\}^n} |\{r : V(x,w,r)=1\}| \\ &<\; 2^n \cdot (2^{n-m}-1) \\ &=\; 2^m - 2^n.\end{aligned}$$

Thus, there exists $r \in \{0,1\}^m \setminus \{0,1\}^n 0^{m-n}$ such that for every $w \in \{0,1\}^n$ it holds that $V(x,w,r) = 0$. Given such an r, we prove that for all z's $R(x,r,z)=0$. This holds since $R(x,r,z)=1$ requires either r ending with 0^{m-n} (which does not hold by our choice) or $z = w0^{n-m}$ with $V(x,w,r)=1$ (which again cannot hold). ■

Proposition 9 [13]: $\mathcal{P}^{\mathcal{NP}} \subseteq \mathcal{S}_2^P$.

Proof: Let S be an arbitrary set in $\mathcal{P}^{\mathcal{NP}}$, and let M be a (deterministic) polynomial-time oracle machine recognizing S when given access to the NP-complete set S'. We say that a string τ is a valid transcript of $M(x)$ if there exists *some* oracle such that τ describes the computation of M on input x and access to this oracle. Note that the oracle's answers in a valid transcript of $M(x)$ do *not* necessarily agree with the set S'. A valid transcript is said to be supported by a sequence of pairs $\overline{s}$ if for each oracle query q in the transcript τ that was answered by 1 there is a pair (q, w) in $\overline{s}$, where w is an NP-witness for membership of q in S'. A valid transcript is said to be consistent with a sequence of pairs $\overline{s}$ if for each oracle query q in the transcript τ that was answered by 0 there is no pair (q, w) in $\overline{s}$, where w is an NP-witness for membership of q in S'. We consider a fixed parsing of strings into pairs $(\tau, \overline{s})$, where $\overline{s}$ is a sequence of pairs.

We are now ready to define a relation R (for the class $\mathcal{S}_2^P$): For $y = (\tau, \overline{s})$ and $z = (\tau', \overline{s}')$, we let $R(x, y, z) \stackrel{\text{def}}{=} \sigma$ if at least one of the following two conditions holds:

1. τ is a valid transcript of $M(x)$ with output σ, supported by $\overline{s}$ and consistent with $\overline{s}'$.
2. τ' is a valid transcript of $M(x)$ with output σ, supported by $\overline{s}'$ and consistent with $\overline{s}$.

In case none of the conditions hold, $R(x, y, z)$ may be defined arbitrarily. Intuitively, the quantification $\exists y \forall z$ guarantees that the transcript contained in y records correct oracle answers (since positive answers must be supported by NP-witnesses, whereas negative answers must be unrefutable by NP-witnesses to the opposite). Formally, we have to prove that R is well-defined, and that the actual execution transcript is both supportable and unrefutable (i.e., consistent with all valid sequences).

We first show that R is well-defined (i.e., it can not be the case that τ and τ' are both valid, supported and consistent but with different outputs). Here we use the fact that M is deterministic and so given the same oracle answers it must yield the same output. Also, if two valid transcripts differ on some oracle answer, then it cannot be that both transcripts are supported and consistent with respect to the same two sequences of pairs.[2] Finally, observe that for every x, there exists a pair $(\tau, \overline{s})$ with output $\chi_S(x)$ such that τ is a valid transcript of $M(x)$, supported by $\overline{s}$ and consistent with any possible sequence of pairs. ■

5.3 Conjectured Separations

Below we list some well-known conjectures.

Conjecture 1 (the leading conjecture of TOC): $\mathcal{P} \neq \mathcal{NP}$.

Conjecture 2 (most widely believed): $\mathcal{NP} \not\subseteq \mathcal{BPP}$.

Conjecture 3 (most widely believed): $\mathcal{NP} \neq \text{co}\mathcal{NP}$.

Conjecture 4 (widely believed): *The Polynomial-Time Hierarchy does not collapse.*

Conjecture 4 implies the following (see [3]):

[2] Consider the first conflicting answer and suppose, without loss of generality, that in transcript τ the answer is 1. Since τ is supported by a sequence of pairs $\overline{s}$, it cannot be the case that τ' (in which the answer to the same query is 0) is consistent with $\overline{s}$.

Conjecture 5 (widely believed): co$\mathcal{NP} \not\subseteq \mathcal{AM}$.

We believe that Conjecture 5 is interesting on its own; indeed, it is a natural extension of Conjecture 3.

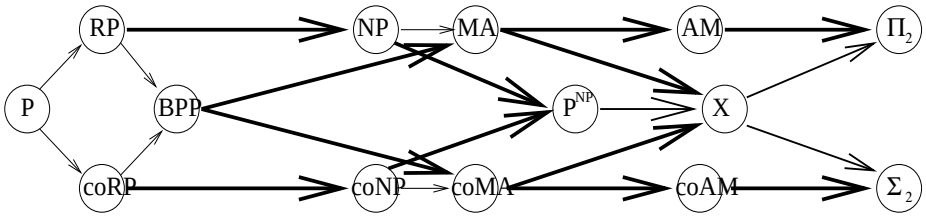

Figure 1: Arrows indicate containment between classes, with $\mathcal{C}_1 \to \mathcal{C}_2$ indicating that $\mathcal{C}_1 \subseteq \mathcal{C}_2$. Bolder (and bigger) arrows indicate conjectured gaps between the classes. The symbol X is a placeholder for either $\mathcal{S}_2^P$ or $\mathcal{ZPP}^{\mathcal{NP}}$ (and we do not know how these two classes are related).

5.4 Conjectured Inclusions

What we know combined with what is widely believed is depicted in Figure 1. We note that some of the inclusions that were not conjectured to be separations are believed to be equalities or "close to it". In particular, it is widely believed that $\mathcal{BPP}$ is very close to $\mathcal{P}$. This belief is supported, among other things, by the conjecture that (uniform) exponential-time cannot be computed by subexponential-size (non-uniform) circuits [2, 10]. We note that the latter conjecture holds provided there exist strong one-way functions (i.e., polynomial-time computable functions that cannot be inverted on typical images by subexponential-sized circuits).

The derandomization of BPP versus the derandomization of MA. We note that results about derandomization of $\mathcal{BPP}$ are likely to imply results on the derandomization of $\mathcal{MA}$. This holds provided that the former results extend also to the generalization of $\mathcal{BPP}$ to promise problems. We note that all known derandomization results have this feature. In the next proposition co$\mathcal{RP}$ denotes the class of *promise problems* of the form $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$, where there exists a probabilistic polynomial time machine M such that

$$\begin{aligned} x \in \Pi_{\text{YES}} &\implies \text{Prob}[M(x) = 1] = 1 \\ x \in \Pi_{\text{NO}} &\implies \text{Prob}[M(x) = 1] \leq \frac{1}{2} \end{aligned}$$

Proposition 10 (folklore): *Suppose that* co$\mathcal{RP} \subseteq$ DTIME$(t(n))$*, for a time constructible function* $t : \mathbb{N} \to \mathbb{N}$*. Then,* $\mathcal{MA} \subseteq \cup_{i \in \mathbb{N}}$NTIME$(t(n^i))$.

Proof: Each set $L \in \mathcal{MA}$ gives rise to a promise problem $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$, where

$$\begin{aligned} \Pi_{\text{YES}} &\stackrel{\text{def}}{=} \{(x, w) : \forall r \in \{0,1\}^{p(|x|)} \quad V(x, w, r) = 1\} \\ \Pi_{\text{NO}} &\stackrel{\text{def}}{=} \{(x, w) : x \notin L\} \end{aligned}$$

with V and p as in Definition 2. Note that, for every $x \in L$ there exists $w \in \{0,1\}^{q(|x|)}$ such that $(x,w) \in \Pi_{\mathrm{YES}}$, whereas for every $x \notin L$ and every $w \in \{0,1\}^{q(|x|)}$ it holds that $(x,w) \in \Pi_{\mathrm{NO}}$. Also, for every $(x,w) \in \Pi_{\mathrm{NO}}$ it holds that

$$\mathrm{Prob}_{r \in \{0,1\}^{p(|x|)}}[V(x,w,r)=1] \;\leq\; \frac{1}{2}.$$

We conclude that $\Pi \in \mathrm{co}\mathcal{RP}$. Now, using the hypothesis, we have $\Pi \in \mathrm{DTIME}(t(n+q(n)))$, and so $L \in \mathrm{NTIME}(t(n+q(n)))$. The proposition follows. ∎

On the derandomization of MA (a comment added in revision). In light of recent derandomization results regarding $\mathcal{AM}$ (cf. [12]), one may question the conjecture $\mathcal{MA} \neq \mathcal{AM}$ (which is suggested by Figure 1). We note, however, that the aforementioned derandomization of $\mathcal{AM}$ seem to require stronger intractability assumptions than the ones used in the derandomization of $\mathcal{BPP}$ (and $\mathcal{MA}$).

Challenges. Indeed, all our challenges call for establishing some appealing inclusions (rather than separations).

1. Try to put $\mathcal{BPP}$ in $\mathcal{P}^{\mathcal{NP}}$. (Recall that $\mathcal{BPP}$ in in $\mathcal{ZPP}^{\mathcal{NP}}$.)
2. Try to put $\mathcal{MA}$ in $\mathcal{P}^{\mathcal{NP}}$. (This certainly implies (1).)
3. Try to put $\mathcal{RP}$ in co$\mathcal{NP}$. (Recall that $\mathcal{RP}$ is in co$\mathcal{MA}$.)
4. Try to put $\mathcal{AM}$ in $\Sigma_2^P \cap \Pi_2^P$.

Acknowledgments. We thank Lane Hemaspaandra for pointing out the fact that $\mathcal{NP}^{\mathcal{BPP}} \subseteq \mathcal{ZPP}^{\mathcal{NP}}$ was proved by Zachos and Furer [15]. Unfortunately, we were not aware of this fact at the time that the current article was published.

References

[1] L. Babai. Trading Group Theory for Randomness. In *17th STOC*, pages 421–429, 1985.

[2] L. Babai, L. Fortnow, N. Nisan and A. Wigderson. BPP has Subexponential Time Simulations unless EXPTIME has Publishable Proofs. *Complexity Theory*, Vol. 3, pages 307–318, 1993.

[3] R. Boppana, J. Håstad, and S. Zachos. Does Co-NP Have Short Interactive Proofs? *IPL*, 25, pages 127–132, 1987.

[4] R. Canetti. On BPP and the Polynomial-time Hierarchy. *IPL*, 57, pages 237–241, 1996.

[5] M. Fürer, O. Goldreich, Y. Mansour, M. Sipser and S. Zachos. On Completeness and Soundness in Interactive Proof Systems. *Advances in Computing Research*, Vol. 5 (Randomness and Computation, S. Micali, ed.), pages 429–442, 1989.

[6] O. Goldreich. A Sample of Samplers – A Computational Perspective on Sampling. *ECCC*, TR97-020, May 1997. See also *Studies in Complexity and Cryptography*, Springer LNCS, Vol. 6650, 2011.

[7] O. Goldreich. *Computational Complexity: A Conceptual Perspective*. Cambridge University Press, 2008.

[8] S. Goldwasser, S. Micali and C. Rackoff. The knowledge Complexity of Interactive Proofs. *SIAM J. on Computing,* Vol. 18 (1), pages 186–208, 1989.

[9] S. Goldwasser and M. Sipser. Private Coins versus Public Coins in Interactive Proof Systems. *Advances in Computing Research*, Vol. 5 (Randomness and Computation, S. Micali, ed.), pages 73–90, 1989.

[10] R. Impagliazzo and A. Wigderson. P=BPP if E requires exponential circuits: Derandomizing the XOR Lemma. In *29th STOC*, pages 220–229, 1997.

[11] C. Lautemann. BPP and the Polynomial Hierarchy. *IPL*, 17, pages 215–217, 1983.

[12] P.B. Miltersen and N.V. Vinodchandran. Derandomizing Arthur-Merlin Games using Hitting Sets. *Computational Complexity*, Vol. 14 (3), pages 256–279, 2005. Preliminary version in *40th FOCS*, 1999.

[13] A. Russell and R. Sundaram. Symmetric Alternation Captures BPP. *Journal of Computational Complexity*, to appear. Preliminary version in Technical Report MIT-LCS-TM-541, 1995.

[14] M. Sipser. A Complexity Theoretic Approach to Randomness. *15th STOC*, pages 330–335, 1983.

[15] S. Zachos and M. Fürer. Probabilistic Quantifiers vs. Distrustful Adversaries. In *Proc. FST-TCS*, Springer-Verlag, Lecture Notes in Computer Science (Vol. 287), pages 443–455, 1987.

[16] S. Zachos and H. Heller. A decisive characterization of BPP. *Information and Control*, Vol. 69 (1-3), pages 125–135, 1986.

[17] D. Zuckerman. Simulating BPP Using a General Weak Random Source. *Algorithmica*, Vol. 16, pages 367–391, 1996.