A Note on Extracting Randomness from Santha-Vazirani Sources

Omer Reingold*
Incumbent of the Walter and Elise Haas
Career Development Chair
Department of Computer Science
Weizmann Institute of Science
Rehovot, Israel
omer.reingold@weizmann.ac.il

Salil Vadhan†
Division of Engineering & Applied Sciences
Harvard University
Cambridge, MA, USA
salil@eecs.harvard.edu

Avi Wigderson‡
School of Mathematics
Institute for Advanced Study
Princeton, NJ, USA
avi@ias.edu

October 11, 2004

Abstract

Santha and Vazirani [SV86] proposed the notion of a *semi-random* source of bits (also known as *Santha-Vazirani* sources), and proved that it is impossible to (deterministically) extract an almost-uniform random bit from such a source (while it is possible given several independent Santha-Vazirani sources). We provide a simpler and more transparent proof of their impossibility result. Our proof has the advantage of applying to a natural strengthening of Santha-Vazirani sources. Moreover, our proof technique has been used in [DOPS04] to obtain impossibility results on doing cryptography with semi-random sources.

A *source* of length n is a random variable X taking values in $\{0,1\}^n$. We will denote the individual bits of X by $X = X_1 \cdots X_n$. If Y is a random variable taking values in $\{0,1\}$, the *bias* of Y is $|\Pr[Y=0] - \Pr[Y=1]|$, i.e. the smallest δ such that $(1-\delta)/2 \leq \Pr[Y=0] \leq (1+\delta)/2$.

Definition 1 *For $\delta \in [0,1]$, a source X of length n is a* Santha-Vazirani (SV) source *(or* semi-random source*) with* bias δ *if for every $i \in [n]$ and every $x_1, \ldots, x_i \in \{0,1\}$, the bias of X_i conditioned on $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$ is at most δ. That is,*

$$\frac{1-\delta}{2} \leq \Pr[X_i = x_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1}] \leq \frac{1+\delta}{2}.$$

*Most of this research was performed while at AT&T Labs - Research. Florham Park, NJ, and while visiting the Institute for Advanced Study, Princeton, NJ. Research was supported in part by US-Israel Binational Science Foundation Grant 2002246.

†Supported by US-Israel BSF Grant 2002246, NSF grant CCR-0133096, and ONR Grant N00014-04-1-0478. URL: http://eecs.harvard.edu/~salil.

‡Add acks.

Informally, in an SV source each bit is guaranteed to be slightly unpredictable given the previous bits.

We will give a simpler proof of the following theorem. The original paper [SV86] mentions that other simple proofs were previously found by Johan Håstad and Vijay Vazirani, but these were never published. The proof presented in [SV86] is attributed to Mihaly Gereb; the earlier conference version contained an even more involved proof.

Theorem 1 ([SV86]) *For every function* $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$ *and every* $\delta \in [0,1]$*, there is an SV source* X *of bias* δ *such that* $\mathrm{Ext}(X)$ *has bias at least* δ*.*

That is, there is no single function Ext that can extract a random bit of bias smaller than δ from every SV source of bias δ. Thus, the trivial extractor that outputs the first bit of its input is optimal.[1]

Our proof will actually apply to two stronger (i.e. more constrained) models of sources.

Definition 2 *For* $\delta \in [0,1]$*, source* X *of length* n *is a* strong SV source *with* bias δ *if for every* $i \in [n]$ *and every* $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \{0,1\}$*, the bias of* X_i *conditioned on* $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \ldots, X_n = x_n$ *is at most* δ*.*

That is, here we require that each bit is slightly unpredictable given *all* the others (not just the previous ones).

Definition 3 *For* $\delta \in [0,1]$*, a source* X *of length* n *is* δ-imbalanced *if for every* $x, y \in \{0,1\}^n$*,* $\Pr[X = x]/\Pr[Y = y] \leq (1+\delta)/(1-\delta)$*.*

Theorem 1 follows from the following easy lemmas.

Lemma 1 *For every function* $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$ *and every* $\delta \in [0,1]$*, there is a* δ*-imbalanced source* X *such that* $\mathrm{Ext}(X)$ *has bias at least* δ*.*

Proof: There exists a set $S \subset \{0,1\}^n$ of size exactly $2^n/2$ such that Ext is constant on S, say taking value $\sigma \in \{0,1\}$. Consider the source X that with probability $(1+\delta)/2$ outputs a uniformly selected element of S and with probability $(1-\delta)/2$ outputs a uniformly selected element of $\{0,1\}^n \setminus S$. By construction, $\mathrm{Ext}(X) = \sigma$ with probability at least $(1+\delta)/2$, so it has bias at least δ. Moreover, X assigns every string in $\{0,1\}^n$ probability mass either $((1+\delta)/2) \cdot (1/|S|) = (1+\delta)/2^n$ or $((1-\delta)/2)) \cdot (1/|\{0,1\}^n \setminus S|) = (1-\delta)/2^n$. Thus X is δ-imbalanced. ■

Lemma 2 *Every* δ*-imbalanced source is a strong SV source with bias* δ*.*

Proof: Let X be a δ-imbalanced source. For every $i \in [n]$, and every $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, the δ-imbalanced property implies that

$$\frac{1-\delta}{1+\delta} \leq \frac{\Pr[X = x_1 \ldots x_{i-1} 0 x_{i+1} \cdots x_n]}{\Pr[X = x_1 \ldots x_{i-1} 1 x_{i+1} \cdots x_n]} \leq \frac{1+\delta}{1-\delta}.$$

This is equivalent to saying that X_i has bias at most δ conditioned on $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \ldots, X_n = x_n$. ■

[1] However, if we relax the problem to allow the extractor access to several independent SV sources (as in [SV86, CG88]) or a small number of truly random bits (as in [NZ96]), then much better extraction is possible and these insights have led to a long and fruitful line of work.

Lemma 3 *Every strong SV source with bias δ is an SV source with bias δ.*

Proof: The bias of X_i conditioned on $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$ is at most the maximum, taken over all $x_{i+1}, \ldots, x_n$, of the bias of X_i conditioned on $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \ldots, X_n = x_n$ (because the former distribution is a convex combination of the latter set of distributions). ■

Thus, we conclude that it is impossible to extract a bit of bias less than δ from δ-imbalanced sources, strong SV sources of bias δ, and SV sources of bias δ.

References

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, April 1988.

[DOPS04] Yevgeniy Dodis, Shien Jin Ong, Manoj Prabhakaran, and Amit Sahai. On the (im)possibility of cryptography with imperfect randomness. To appear in *Proceedings of the Forty-Fifth Annual IEEE Symposium on the Foundations of Computer Science*, Rome, Italy, October 2004.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, February 1996.

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33(1):75–87, August 1986.