Deterministic Extractors - Lecture Notes

Speaker: Avi Wigderson
Scribe: Zeev Dvir

February 4, 2009

1 Motivation

Randomness is used in many places in our daily lives. Some examples are gambling, statistics, algorithms, cryptography etc. For such applications one typically assumes a supply of completely unbiased and independent randim bits. This raises the problem of where to get these assumed random bits from. We could try and use natural sources of randomness such as sun spots, the stock market or the weather, which seem unpredictable at least to some degree. The problem is that these natural sources are not completely random, only somewhat random, in the sense that samples from them are often correlated and biased. The goal of deterministic extractors is to *extract* completely random bits from such "weak random sources". One hopes to obtain extractors which would achieve this task for as large and general families of weak sources. We will concentrate on the problem of extracting only a single random bit, which turns out to demonstrate the variety of problems, results and techniques in this area (in the general case we would like to get many uncorrelated bits). We will not discuss here seeded extractors (e.g. see survey [Sha02] and the recent [DW08]).

2 The formal model

Let $\mathcal{D}$ be a finite domain and let $\mathcal{S}$ be a family of distributions over $\mathcal{D}$. An ϵ-extractor for $\mathcal{S}$ is a function

$$f : \mathcal{D} \mapsto \{1, -1\}$$

such that for every $X \in \mathcal{S}$ we have

$$|\mathbb{E}[f(X)]| \leq \epsilon.$$

In words, even if an adversary chooses a distribution $X \in \mathcal{S}$, the output bit $f(X)$ will be nearly unbiased (up to ϵ). Of course, for most applications we want f to be efficiently computable. E.g. if the elements of $\mathcal{D}$ are representable by n-bit sequences, we'd like f to be computable in time polynomial in n. We proceed to give some examples.

3 Von Neumann's extractors

Let $\mathcal{D} = \{0,1\}^n$ and for $\delta < 1$ denote by

$$B_\delta = \{X = (X_1, \ldots, X_n) | X_i \textit{ are i.i.d and } \delta\textit{-biased}\}$$

the family of distributions sampled by n biased and independent coin tosses. Von Neumann [vN51] observed that one can extract random bits from these distributions (without knowing the value of δ) by the following simple procedure. Consider the first two coin tosses (X_1, X_2) the probability of getting $(0,1)$ and $(1,0)$ are the same. Therefore, we could output 1 if we saw the first and -1 if we got the second. If the output (X_1, X_2) was either $(1,1)$ or $(0,0)$ we will try again (that is, look at the next pair (X_3, X_4)). The only way this could fail is if all the pairs (X_{2i-1}, X_{2i}) are bad (in which case we might output a default value, say 1). This procedure defines an efficiently computable function f which is an $\epsilon = \epsilon(\delta)$-extractor for every B_δ in which $\epsilon = \exp(n)$, corresponding to failing on all pairs.

The extractor above strongly uses the fact that the variables X_i are i.i.d., namely have the *same* bias. Consider the family of random sources $X = (X_1, \ldots, X_n)$ in which we just guarantee independence of the coins X_i, but the biases can differ, as long as they do not exceed δ. Even these sources can be handled very easily, simply by taking $f(X) = X_1 \oplus X_2 \oplus \cdots \oplus X_n$. It is easy to see that f is an $\epsilon = \epsilon(\delta)$-extractor for every B'_δ with $\epsilon = \exp(n)$.

In the two examples above it seems that the constant amount of entropy in each bit was a key to the possibility of extraction, and its exponentially small bias. A more general family of sources was studied by Santha and Vazirani [SV86]. Here, we remove independence as well, and leave only the entropy guarantee for every bit. Let SV_δ be the family of distributions in which for every i random variable X_i is δ-biased *conditioned* on *any* outcome values of the previous coin tosses $(X_1, \ldots, X_{i-1})$. Here, we get an impossibility of extraction result even for every specific δ.

Theorem 3.1 ([SV86]). *Fix any $\delta \in (0,1)$. The no function f is not an ϵ-extractor SV_δ with $\epsilon < \delta$.*

There are several proofs of this theorem, and the reader is encouraged to find one.

4 Affine sources – small fields

We will now impose algebraic structure on the domain $\mathcal{D}$.

Let $\mathcal{D} = \mathbb{F}_2^n$ and define the following family of distributions:

$$L_{2,k} = \{X \quad \textit{uniform on a k-dimensional affine subspace of } \mathcal{D}\}.$$

When k is large enough the following folklore theorem shows the existence of a simple extractor for this family

Theorem 4.1. *Suppose $k > (1/2 + \alpha)n$ and define*

$$f(x_1, \ldots, x_n) = x_1x_2 + x_3x_4 + \ldots + x_{n-1}x_n.$$

Then f is an ϵ-extractor for $L_{2,k}$ with $\epsilon = \exp(-\alpha n)$.

On the other hand, it is not hard to show that almost every f will be an extractor for $L_{2,k}$ with k being roughly logarithmic in n. This follows from a probabilistic argument, estimating the probability that a random f is not $\exp(-k)$-extractor for a fixed such X, and then taking a union bound over all subspaces. But finding an explicit f which is an extractor for such small dimensional subspaces is considered a very hard task. The current state of the art is the following result of Bourgain [Bou07].

Theorem 4.2 ([Bou07]). *There exists an efficient f which is an ϵ-extractor for $L_{2,k}$ for every $k = \Omega(n)$ and $\epsilon = \exp(-\Omega(n))$. In fact, this f is a polynomial of constant degree (a function of n/k) in the variables X_i.*

The proof is quite intricate, and uses among other things tools from arithmetic combinatorics. A particularly useful one, which is key to many other developments, is the following result of Bourgain.

Theorem 4.3 ([Bou08]). *Let $\chi : \mathbb{F}_q \mapsto \mathbb{C}$ be a non trivial additive character. Let $A_1, \ldots, A_s \subset \mathbb{F}_p$ with $|A_i| > p^\delta$. Suppose $s > C/\delta$ for sufficiently large constant C, then*

$$\left| \sum_{a_i \in A_i} \chi(a_1 \cdot \ldots \cdot a_s) \right| \leq p^{-\delta'},$$

with $\delta' > C^{-s}$.

5 Affine sources – large fields

Here we'll see that the extraction problem from subspaces becomes significantly easier, and the results are much stronger, when the field size grows with n. Moreover, the techniques use exponential sums estimates, and we'll see how such classical results themselves may be viewed as extractors.

Let $\mathcal{D} = \mathbb{F}_p^n$ with p a prime larger than n^4. We define the family of distribution $L_{p,k}$ in the same way as before, only that now the subspaces are over $\mathbb{F}_p$. The following theorem of Gabizon and Raz [GR05] gives an extractor for this family, even for lines ($k = 1$).

Theorem 5.1 ([GR05]). *There exists an efficient f which is an ϵ-extractor for $L_{p,k}$ for every $k \geq 1$ and with $\epsilon = 1/n$.*

Note that the higher the dimension k, the larger the entropy in these random variables, so we should expect ϵ to decrease with k. It is an interesting open problem to improve the bias to ϵ to $p^{\Omega(k)}$, which is achievable by a random function.

The theorem above uses the following deep result from Algebraic Geometry due to Weil.

Theorem 5.2 ([Wei48]). *Let χ be the quadratic character of $\mathbb{F}_p$. Let $g \in \mathbb{F}_p[z]$ be a polynomial which is not a square and has degree d. Then, for Z uniform in $\mathbb{F}_p$*

$$|\mathbb{E}_Z[\chi(g(Z)]| \leq d/\sqrt{p}.$$

We can define the following family of sources in $\mathbb{F}_p$

$$P_d = \{X = g(Z) \,:\, Z \textit{ uniform in } \mathbb{F}_p\}$$

and interpret Weil's theorem as saying that χ is an ϵ extractor for P_d with $\epsilon = d/\sqrt{p}$. Note that χ is efficiently computable.

To prove the theorem of Gabizon and Raz (even for the case $k = 1$, which implies the general case viewing a subspace as a union of lines) we define the following polynomial

$$g(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^{2i+1}$$

and observe that for every 1-dimensional subspace V, the restriction of g to V is a nonzero polynomial in one variable z, of degree at most $2n + 1$ and is not a square (since every monomial has odd degree). The extractor of [GR05] is therefore given by $f(x) = \chi(g(x))$.

6 Polynomial sources

Extensions of affine sources to sources defined using higher degree polynomial equations were studied in [DGW07, Dvi08]. In [DGW07] the model of affine source was extended to polynomial sources which are sources *sampled* by low degree polynomials. That is, sources which are sampled by choosing an element of $\mathbb{F}^k$ uniformly and then applying a polynomial mapping from $\mathbb{F}^k$ to $\mathbb{F}^n$ on it. The results of [DGW07] give an extractor for sources sampled by a polynomial mapping of degree d when the field size is at least $d^{O(n)}$. This result uses among other things the exponential sum estimate of Bombieri which extends Weil's exponential sum (Theorem 5.2) to sums over a curve. More formally, the variable Z in Theorem 5.2 now ranges over a curve in $\mathbb{F}^n$ and not over $\mathbb{F}$.

In [Dvi08] a different model for low degree sources was studied. This time the source is uniform over the set of zeros of a system of polynomial equations (a variety) of bounded degree d. One result in [Dvi08] gives an extractor for arbitrary sources of this kind when the field size is at least $d^{\Omega(n^2)}$. This result relies among other things on the exponential sum estimate of Bombieri mentioned above. The other result in [Dvi08] gives an extractor when the field size is at least $d^{O(1)}$ and under the additional constraint that the variety contains at least $|\mathbb{F}|^{n/2}$ points. This last result uses the following theorem of Deligne, which is a strengthening of Weil theorem to higher dimensions.

Theorem 6.1 ([Del74]). *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of degree d and let f_d denote its homogenous part of degree d. Suppose f_d is smooth. Then, for every non trivial additive character $\chi : \mathbb{F} \mapsto \mathbb{C}^*$ we have*

$$\left| \sum_{x \in \mathbb{F}^n} \chi(f(x)) \right| \leq (d-1)^n \cdot p^{n/2}.$$

7 Independent blocks

Now we leave the algebraic setting and move to a more combinatorial one, where the weak sources have several independent parts, each with sufficient entropy.

Let $\mathcal{D} = \{0,1\}^n \times \{0,1\}^n$. Define the family

$$I_{2,k} = \{X \quad uniform\ on \quad A_1 \times A_2, \ \ |A_i| \geq 2^k\}$$

of all distributions composed of two independent blocks, each distributed uniformly on a set of size at least 2^k. As before, for $k > (1/2 + \alpha)n$ we have a very simple extractor given by the inner product function $f(x,y) = (-1)^{<x,y>}$. Again, there are several proofs and the reader is encouraged to find one.

Again, a simple probabilistic argument that most functions f will be $\exp(-k)$-extractor for every $k >> \log n$. The current state of the art for explicit construction is also due to Bourgain.

Theorem 7.1 ([Bou05]). *There exists an efficient f which is an extractor for $I_{2,k}$ with $k = 0.4999 \cdot n$.*

This theorem again uses arithmetic combinatorics, in particular the famous sum-product growth in finite fields of Bourgain, Katz and Tao [BKT04], which has been useful for other extractors for independent blocks mentioned below. Before turning to them, we note that while no extractors for two independent blocks with lower entropy (e.g. $k = n/3$) exists, much better results can be proved for a related notion called a *disperser.*

A disperser relaxes the demand that the output of f will be close to uniform, and only demands that it will be non constant. In other words, $f(X) = \{-1, 1\}$ for every X in the given family. Note that in this case only the supports of the given distributions matter. We note that dispersers are a very natural objects in Ramsey theory, in which

f may be viewed as a 2-coloring of $\mathcal{D}$ with the property that no subset X in our family is monochromatic. In particular, the family $I_{2,k}$ corresponds to the edges of complete bipartite graphs (equivalently, the entries of a matrix), which should be 2-colored so as to avoid every large monochromatic complete subgraph (equivalently, submatrix). Explicit such colorings f correspond to explicit bipartite Ramsey graphs, and symmetric colorings f (in which $f(x,y) = f(y,x)$) correspond to explicit Ramsey graphs. We note that the bipartite case is significantly harder.

Explicit dispersers for $I_{2,k}$ were constructed in [BKS+05] for $k = \Omega(n)$ and the further improved in [BRSW06] for $k = \exp{(\log n)^{.9}}$ which is smaller than any polynomial in n. The proofs are very long and complex, and use quite a bit of extractor machinary for more than two blocks (mentioned below), as well as the result of Bourgain for 2 blocks above (we note that it is critical that his result works for entropy below $n/2$).

The family of r-block soruces $I_{r,k}$ is defined analogously for every $r \geq 2$. The domain $\mathcal{D} = (\{0,1\}^n)^r$ and every $X \in I_{r,k}$ is of the form $X_1 \times X_2 \times \cdots \times X_r$ with each X_i uniform on some subset A_i of size at least 2^k. Naturally, the extraction problem becomes easier the larger r is, and the problem of extracting from any constant number r of blocks was open for almost 20 years, since the aforementioned paper of Santha and Vazirani [SV86]. And indeed progress for larger r preceded the case $r = 2$ above. First, [BIW04] used the sum-product theorem [BKT04] to show that to efficiently extract from linearly small entropy $k = \alpha n$ it suffices to have $r = poly(1/\alpha)$ independent blocks. This was dramatically improved by Rao [Rao06] who showed (without arithmetic combinatroics) how to extract from polynomially small entropy $k = n^\beta$ using $r = poly(1/\beta)$ independent blocks. In both results $\epsilon = \exp(-k)$. It remains an open problem to reduce this entropy further, or obtain such extraction with r fixed independent of k, say $r = 3$ or even $r = 100$.

References

[BIW04] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, Washington, DC, USA, 2004. IEEE Computer Society.

[BKS+05] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: new constructions of condensers, ramsey graphs, dispersers, and extractors. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 1–10, New York, NY, USA, 2005. ACM Press.

[BKT04] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Func. Anal.*, 14:27–57, 2004.

[Bou05] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 2005.

[Bou07] J. Bourgain. On the construction of affine extractors. *Geometric And Functional Analysis*, 17(1):33–57, 2007.

[Bou08] J. Bourgain. Multilinear exponential sum bounds with optimal entropy assignement. GAFA (to apepar), 2008.

[BRSW06] B. Barak, A. Rao, R. Shaltiel, and A. Wigderson. 2-source dispersers for subpolynomial entropy and ramsey graphs beating the frankl-wilson construction. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 671–680, New York, NY, USA, 2006. ACM Press.

[Del74] P. Deligne. La conjecture de weil. *I , Inst. Hautes Etudes Sci. Publ. Math.*, 43:273–307, 1974.

[DGW07] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 52–62, Washington, DC, USA, 2007. IEEE Computer Society.

[Dvi08] Z. Dvir. Extractors for varieties. Manuscript, 2008.

[DW08] Z. Dvir and A. Wigderson. Kakeya sets, new mergers and old extractors. In *FOCS '08*, 2008.

[GR05] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 407–418, Washington, DC, USA, 2005. IEEE Computer Society.

[Rao06] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 497–506, New York, NY, USA, 2006. ACM Press.

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

[SV86] M. Santha and U. V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33:75–87, 1986.

[vN51] J. von Neumann. Various techniques used in connection with random digits. *Applied Math Series*, 12:36–38, 1951.

[Wei48] A. Weil. On some exponential sums. In *Proc. Nat. Acad. Sci. USA*, volume 34, pages 204–207, 1948.