Pseudorandomness Program

Release Spring 2017

Simons Institute

Feb 06, 2017

CONTENTS

CHAPTER

ONE

ADDITIVE COMBINATORICS

Problem 1 ((submitted by Pseudonym))**.** Given a graph G.

See [vadhan2012pseudorandomness].

CHAPTER TWO

CRYPTOGRAPHY

Problem 2 ((submitted by Pseudonym))**.** Given a graph G.

See [vadhan2012pseudorandomness].

CHAPTER

THREE

EXPANDERS AND EXTRACTORS

3.1 Open Problems presented at the Simons Workshop on Expanders and Extractors, January 30-Feb. 3, 2017

Compiled by Noga Alon.

Partial Steiner systems of large girth, Nati Linial

Let us first recall that the *girth* of a graph G is the least integer g such that there is a set of g vertices in G that spans at least g edges. We seek an analogous notion for 3-uniform hypergraphs. In fact we only deal with *linear* hypergraphs H where every two hyperedges share at most one vertex. As defined by Erdős, the girth of H is the smallest integer $g \geq 4$ such that there is a set of g vertices in H that spans at least $g - 2$ hyperedges. He conjectured that there exist Steiner Triple Systems with arbitrarily high girth, but despite considerable attempts over the years, the state of our knowledge concerning this problem is quite bad. I therefore formulate a variant that may be more accessible:

Question: Does there exist $c > 0$ and n-vertex 3-uniform hypergraphs with at least cn^2 hyperedges and arbitrarily high girth?

Cliques in near Ramanujan Graphs, Noga Alon

An (n, d, λ)-graph is a d-regular graph on n vertices so that all eigenvalues but the top one are in absolute value at most λ. Let G be an $(n, d, 100\sqrt{d})$-graph. It is known that:

1. There is a constant $c > 0$ so that if $d \geq cn^{2/3}$ then G contains a triangle.
2. The statement in (1) is tight up to the constant c.
3. There is a constant $c > 0$ so that if $d \geq cn^{4/5}$ then G contains a copy of K_4.

Open: Is (3) tight up to the value of c? That is, is there an $(n, d, 100\sqrt{d})$-graph containing no K_4, where $d = \Omega(n^{4/5})$? It is not even known whether or not there is such a graph with $n^{2/3} = o(d)$.

On the extractable entropy from zero-fixing sources, Gil Cohen

Let $n \geq k \geq 0$ be integers. An n-bit random variable X is called a *k-zero-fixing source* if there exists a subset of indices $R \subseteq [n]$ such that the marginal of X when projected to R is uniformly distributed, and the remaining bits of X are fixed to zero. The parameter k is called the entropy of X. Let $m = m(n, k)$ be the largest integer for which there exists a function $\mathsf{Ext}_k : \{0, 1\}^n \to \{0, 1\}^m$ with the following property: for any k-zero-fixing source X, $\mathsf{Ext}_k(X)$ is within statistical distance $1/5$ from uniform (here $1/5$ is an arbitrary choice of a small constant.) We stress that Ext_k does not get the description of the source X (namely, the set R) but rather a single sample from X. Such a function Ext_k is called an *extractor for k-zero-fixing sources*. We would like to have a better understanding of the function $m(n, k)$ that, informally speaking, captures the amount of extractable entropy from zero-fixing sources. Clearly, $m(n, k) \leq k$. By a straightforward counting argument, for any $k > \log_2 \log n + \log_2 \log \log n + \Omega(1)$, $m(n, k) = k - O(1)$. That is, when the entropy k is high enough, one can extract essentially all the entropy from the source. What about smaller entropies? For any k, $m(n, k) \geq 0.5 \log_2 k - O(1)$ as can be seen by considering the function that takes the Hamming weight of the input X modulo $\Theta(\sqrt{k})$ [kamp2006deterministic][reshef2013extractors]. That is to say, a logarithmic amount of entropy is always extractable. Interestingly, this simple function is optimal for small enough k. Put formally, $m(n, (\log^* n)^{2/3}) \le 0.5 \log_2 k + O(1)$ [cohen2015zero]. The natural open problem is to understand the behavior of $m(n, k)$ and close the gap between $k = \Omega(\log \log n)$ and $k = O((\log^* n)^{2/3})$. In particular, it is not clear if $m(n, k)$ has a threshold behavior, namely, if there exists a function $\tau(n)$ such that for $k = \omega(\tau(n))$, $m(n, k) = k - O(1)$ whereas for $k = o(\tau(n))$, $m(n, k) \le 0.5 \log_2 k + O(1)$.

Beating the expander mixing lemma for small sets, David Zuckerman

The expander mixing lemma asserts that in a d-regular graph G on n nodes, for sets S and T of size k, we have

$$|e(S,T) - dk^2/n| < \lambda k,$$

where $\lambda = \max(\lambda_2, -\lambda_n)$ is the largest nontrivial eigenvalue in absolute value. Even for optimal $\lambda = O(\sqrt{d})$, this is useless when $k < c\sqrt{n}$. On the other hand, when $d > (n/k)\log(n/k)$, most d-regular graphs achieve an upper bound of $O(k\sqrt{(dk/n)\log(n/k)})$. This is better than the expander mixing lemma for $k = o(n)$.

The problem is to give an upper bound $f(G, k)$ that is better than λk for some interesting graphs or for most or even many graphs. For dense graphs, this corresponds to two-source extractors. Bounds on line-point incidence graphs are also known. It would be extremely interesting to have a general method.

Spectral radius problem for free groups, Emmanuel Breuillard

If μ is a probability measure on a non-abelian free group F, let $\sigma(\mu) := \|T_\mu\|$ be the norm of the convolution operator on $\ell^2(F)$

$$T_\mu : f \mapsto \mu * f,$$

where $\mu * f(x) = \sum_{g \in G} f(g^{-1}x)\mu(x)$.

Is it true that for every $\epsilon > 0$ there is $\delta > 0$ such that for all probability measures μ on F, the condition $\sigma(\mu) > \epsilon$ implies that there is a coset xH of a cyclic subgroup H of F such that $\mu(xH) > \delta$?

Explicit Coding Power Series, Anup Rao

We are interested in giving an explicit description of a formal power-series over a finite field F with some nice properties. The motivation comes from several applications related to coding.

We say that a power series $P(X) = p_0 + p_1 X + \ldots$ is ϵ-*sparse* if there is a finite k such that of the first k coefficients of $P(X)$, at most ϵk of them are non-zero.

Definition: $P(X)$ is an ϵ-*coding power series* if for every polynomial $g(X)$ with $0/1$ coefficients, the power series $(1 + Xg(X))P(X)$ is not ϵ-sparse.

Fact: For every $\epsilon > 0$, there is a finite field F for which a random power series $P(X)$ will be a coding power series with positive probability.

Open Problem: Give an explicit example of a coding power-series.

BIBLIOGRAPHY

[CS15] Gil Cohen and Igor Shinkar. Zero-fixing extractors for sub-logarithmic entropy. In *International Colloquium on Automata, Languages, and Programming*, 343–354. Springer, 2015.

[KZ06] Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2006.

[RV13] Yakir Reshef and Salil Vadhan. On extractors and exposure-resilient functions for sublogarithmic entropy. *Random Structures & Algorithms*, 42(3):386–401, 2013.

[Vad12] Salil P Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.