

Current Research

Shlomo Hoory

November 16, 2005

I am interested in various problems in extremal graph theory, expander graphs, coding theory, and computational complexity. My favorite tools are algebraic methods, spectral analysis, Markov chains, and the probabilistic method.

I am currently studying the following problems:

1. Extremal problems on non-regular graphs

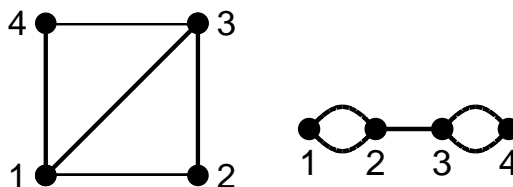
- (a) Perhaps the most important topic in extremal graph theory, is to determine the Turán number of a graph H . This is the maximal number of edges a graph of size n can have, without having a subgraph isomorphic to H . When H is non-bipartite, the asymptotics of this function is determined by the celebrated Erdős-Stone theorem, up to an $1 + o(1)$ factor. However, it is a longstanding problem to find good estimates on the Turán number of bipartite graphs, and in particular of even length cycles, C_{2k} .

In my past work in [1, 3, 4], I developed a technique for obtaining lower bounds on extremal properties of graphs, using entropy arguments. Recently I found one more application for this technique, in a joint work with Jacques A. Verstraete, and Felix Lazebnik, concerning a conjecture of Sidorenko 1991: For any $l < k$, the number of length $2k$ cycles is maximized in a C_{2l} free graph, when the graph is a Turán graph for C_{2l} . We have made significant progress in proving the conjecture for the cases $l = k - 1$, and $l = k - 2$. Our main tool is a generalization of my work [1]. That work gives a lower bound on the number of locally one to one embeddings of the length k path into a graph of a specified edge density. Our recent generalization, is to prove a similar result when embedding trees instead of paths.

- (b) In [1] we gave an upper bound on the girth of a graph of a specified size and degree distribution. A natural question is to prove a corresponding lower bound by constructing a family of graphs

with a specified degree distribution and large girth. However, this task is difficult even for the d -regular case. The naive approach to the problem, Erdős and Sachs 1963, still leaves a factor of two between the upper and lower bounds on the girth. When d is a prime power plus one, there are number theoretic constructions, such as Lubotzky, Phillips and Sarnak 1988, that reduce the gap to a factor of $3/2$. For irregular graphs, even a factor two gap is unknown. I propose to apply the naive approach also to irregular graphs. To this end one needs an initial graph with a specified degree distribution, but relatively few short cycles. The obvious solution is to pick the initial graph at random using the standard configuration model. However, I have realized that there are examples where one can do better. It seems to me that the inferiority of the standard configuration model is its lack of high level planning based on the given degree distribution. This problem naturally led me to the next problem.

- (c) Of all graphs with a specified degree distribution, which graphs minimize the largest eigenvalue λ_1 of their adjacency matrix? Currently, I can solve the problem for some degree distributions. For example, for the distribution where half the vertices have degree two and half have degree three, the optimal graphs are the lifts of one of the following four vertex graphs:



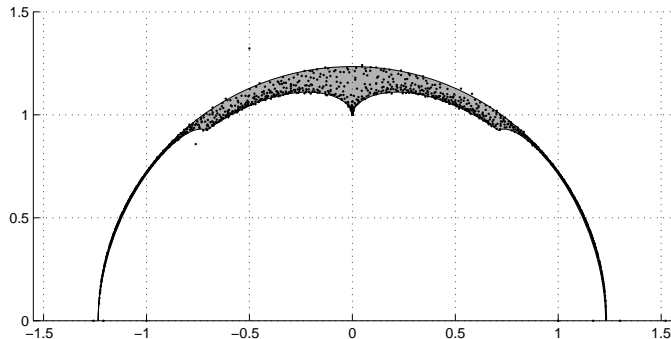
For general degree sequences the problem is more complex, and I suggest partitioning it into two independent subproblems. The first one is the fractional problem, which is to find non-negative *real* edge weights satisfying the degree constraints and minimizing the largest eigenvalue. The second subproblem is to convert a weighted graph into an actual graph. Here I suggest to generalize the well known model of random lifts of graphs, to accommodate lifts of weighted graphs.

2. The non-backtracking spectrum of graphs

A non-backtracking path in a graph is a path v_0, v_1, \dots, v_l , where $v_i \neq v_{i+2}$ for all i . Counting non-backtracking paths or cycles is a crucial step in obtaining important results, such as lower bounding the number of vertices in a graph of specified girth and degree sequence [1], and Friedman's proof of Alon's Second Eigenvalue Conjecture, 2005. The way to formulate the problem algebraically is to replace the standard adjacency matrix A by a matrix B indexed by directed edges, whose uv, wz entry is one if $v = w$ and $u \neq z$, and is zero otherwise. Compared with the well studied spectrum of A , not much is known about the eigenvalues of the matrix B , or equivalently, the poles of the Ihara zeta function. We refer to B 's spectrum as the non-backtracking spectrum of the graph. In the irregular case, the non-backtracking spectrum does not have any simple correspondence with the standard spectrum, and merits study in its own right.

An important lesson learned from studying the standard spectrum of graphs, is that it is worthwhile to investigate the spectrum of random graphs and their limit object, the universal covering tree. McKay 1981, proved that the eigenvalue distribution of random regular graphs, obeys a limit law similar to Wigner's semicircle law for random symmetric matrices with i.i.d. normal entries, 1958. By analogy, the limit law for the non-backtracking spectrum, which is a distribution on complex numbers, should correspond to removing the symmetry requirement in Wigner's semicircle law, which is Girko's circular law, proved by Bai 1997.

In a joint work with Omer Angel and Joel Friedman, we explore the non-backtracking spectra of high lifts and the universal cover of fixed base graphs. We have obtained a recipe to determine the non-backtracking spectrum of the universal cover of a graph. In the figure, the base graph G is taken as K_4 minus an edge. The (upper half of the) non-backtracking spectrum of G 's universal cover is depicted in solid gray, and the non-backtracking spectrum of a random 300-lift of G is shown as black points. We conjecture that the limit spectrum of high lifts is the spectrum of the universal cover.



I have a variety of additional problems on the extremal properties of the spectrum and non-backtracking spectrum of graphs. An example is the conjecture that for any graph, the Perron (largest) eigenvalue is at least one plus the Perron non-backtracking eigenvalue. One technical difficulty in dealing with the non-backtracking spectrum is that a distribution on complex numbers cannot be determined by its moments. I would like to find a way to overcome this difficulty.

3. Monotone circuits for the majority and sorting networks

In [6], we investigated the size and depth complexity of monotone circuits for the majority function. There are two fundamental constructions of logarithmic depth monotone circuits for the majority. One is the AKS sorting network, and the other is a construction of Valiant. Our main result is a randomized construction of a simple circuit of depth $5.3 \log n + O(1)$, and size $O(n^3)$. That is, a circuit of virtually the same depth as Valiant's formula, but with a significantly smaller size. Our construction can be considered as a partial derandomization of Valiant's formula, where achieving full derandomization is a major open question. There are various interesting problems that I would like to pursue further:

- For all we know, our construction may work as is with size $O(n^2)$. This depends on the tightness of a certain union bound, which deserves a more detailed study.
- As the layer size of our circuit goes down exponentially with the height, it is natural to consider the construction of the first few layers as a separate problem. It turns out that this is an interesting discrepancy problem in hypergraphs.

- One other approach to the problem is to consider possible connections to sorting, that may imply an $O(n^2)$ circuit for the majority without increasing the depth too much. One can regard a sorting network as a function $f : \{0,1\}^n \mapsto \{0,1\}^n$ that preserves the number of ones, and maps the 2^n possible inputs to $n+1$ outputs. I define an N -sorter as a circuit that preserves the number of ones while reducing the number of outputs to N . If one can construct an N -sorter for $N = 2^{n^{o(1)}}$, with size $O(n^2)$ and small depth, it will imply the desired $O(n^2)$ circuit for the majority.
- As part of our construction, we build a linear circuit of optimal depth for the promise version of the problem, where the number of 1's in the input is promised to be either less than one third, or greater than two thirds. In an attempt to derandomize this circuit, we use a belief-propagation message-passing algorithm on a good bipartite expander graph. We base our analysis on results of Burshtein and Miller 2001, concerning belief-propagation decoding. It would be of significant interest if one can expand upon this connection between seemingly unrelated areas.

4. Mixing times and cryptographic applications

Motivated by cryptographic problems concerning block ciphers such as DES and AES, I proved the following in [5], [2], and subsequent work: The composition of approximately nk random permutations drawn from a certain simple and small family of permutations on the n -dimensional binary cube results in a permutation that is nearly k -wise independent. This upper bound is tight, with the obvious lower bound, up to a polylogarithmic factor. Stated differently, I proved that the rate of progress up the scale of k -wise independence, is almost as fast as possible. This is a consequence of the fact that underlying Markov chain has a large log-Sobolev constant, implying that entropy increases quickly. I hope to gain insight as to why such families of simple permutations function so well. Furthermore, I wish to find alternative measures of progress other than closeness to k -wise independence, that are appealing from the mathematical view point, and have a closer connection with the security of such cryptographic primitives.

5. Error correcting codes

One of the most intriguing developments in recent years in the area of coding theory is the construction of *Low Density Parity Check* (LDPC) codes where the performance of the *Belief Propagation* (BP) algorithm approaches the channel capacity, Richardson, Shokrollahi, and Urbanke 2001. Our current understanding of the BP algorithm is limited; One can analyze BP only when the number of cycles is very small or when the portion of erroneous messages is small, Burshtein and Miller 2001. I find it interesting that the graphs associated with capacity approaching LDPC codes are *irregular*, a fact which connects to my interest in irregular graphs. I propose to study the following questions:

- (a) Lately, it became evident that the configuration model is not the best possible for constructing LDPC codes achieving high capacity under the BP algorithm, Thorpe 2003. His work suggests using random lifts to yield better performance, which coincide with my conjectures about the optimal graphs for problem 1 (c). Namely, the problem of minimizing the largest eigenvalue of a graph with a prescribed degree sequence. I believe that these two optimization problems have a similar nature. Since the latter problem seems more amenable to systematic study, I suggest to use the results of such a study to point out the right random model for constructing good LDPC codes. Also I suggest that as graphs minimizing the largest eigenvalue have few short cycles, their corresponding LDPC codes suffer less from the degradation due to cycles introduced when decreasing the block length.
- (b) The state of the BP algorithm is a function on directed edges, and the computation is performed along non-backtracking paths. Therefore, it stands to reason that its analysis should be done via the matrix B mentioned earlier, which is the non-backtracking counterpart of the standard adjacency matrix A . I believe that the missing part in realizing this plan, is a non-backtracking theory of expansion, which I would like to investigate further.

6. Random Cayley graphs and high girth

As mentioned before, the construction of graphs of high girth is a notoriously difficult problem even for regular graphs. One approach

for constructing such graphs is to use an explicit construction such as Lubotzky, Phillips and Sarnak 1988. A different approach would be to try and use probabilistic tools. An obvious obstacle immediately encountered, is the lack of natural distributions on graphs that almost surely exhibit high girth. In a joint work with Alex Gamburd, Mehrdad Shahshahani and Bálint Virág, we have shown that a random d -regular Cayley graph almost surely has high girth, for a suitably chosen group. In our work we have considered the symmetric group S_n , and the projective linear group $\mathrm{PGL}(2, q)$. For the latter group we have proved that the normalized girth of a random Cayley graph a.a.s. is at least $1/3$, where the normalized girth is the girth divided by the base $d - 1$ logarithm of the graph size. This result follows from elementary arguments on the number of projective zeros of certain polynomials, and might be possible to improve using more advanced algebraic geometry. Experimental results indicate that the true answer might be $2/3$, the same as the normalized girth of the non-bipartite LPS graphs mentioned above, which are specific Cayley graphs on $\mathrm{PGL}(2, q)$. Therefore, it might be that the girth of non-bipartite LPS graphs is the typical value in this group.

One more natural question in this context is the following. Assume some finite group G . What is the length of the shortest non-trivial reduced word over the variables $g_1^{\pm 1}, \dots, g_k^{\pm 1}$, that is identically one for every assignment of g_1, \dots, g_k with values from G . We have solved the problem for $\mathrm{PGL}(2, q)$, and proved the answer is q .

Finding the answers for the above two questions in groups different from $\mathrm{PGL}(2, q)$ is an interesting question. I have some preliminary results for the symmetric group S_n , and would like to explore other groups as well.

References

- [1] Noga Alon, Shlomo Hoory, and Nathan Linial. The Moore bound for irregular graphs. *Graphs and Combinatorics*, 18(1):53–57, 2002.

- [2] Alex Brodsky and Shlomo Hoory. Simple permutation mix even better. Arxiv math.CO/0411098, submitted to Random Structures and Algorithms, 2005.
- [3] Shlomo Hoory. The size of bipartite graphs with a given girth. *Journal of Combinatorial Theory. Series B*, 86(2):215–220, 2002.
- [4] Shlomo Hoory. A lower bound on the spectral radius of the universal cover of a graph. *J. Combin. Theory Ser. B*, 93(1):33–43, 2005.
- [5] Shlomo Hoory, Avner Magen, Steve Myers, and Charles Rackoff. Simple permutations mix well. In *The 31st International Colloquium on Automata, Languages and Programming (ICALP)*, 2004.
- [6] Shlomo Hoory, Avner Magen, and Toniann Pitassi. Monotone circuits for the majority function. submitted to CCC 2006, 2005.