RESEARCH STATEMENT

Pat Devlin (prd41@math.rutgers.edu)

My research interests are in discrete mathematics, which includes combinatorics, graph theory, and theoretical computer science. I tend to gravitate to problems that are easy to state yet have proved resistant to established approaches. I especially like problems that seem potentially amenable to other, ideally surprising areas of mathematics such as algebra, Fourier analysis, topology, and probability.

I have written on diverse topics in discrete math including algorithms [11, 7], hypergraphs [14], sequences [12, 15], information theory [3], and extremal combinatorics [13, 10]. To provide a more cohesive narrative, I discuss here four papers connected not by content so much as by their association with a main theme of my research: probability in discrete mathematics.

I offer quick, high-level accounts of each of the four followed by some discussion of continuing research plans. A more in-depth research statement is available on my webpage.

Matrices with large permanent [7]

Leonid Gurvits [19] proved that the permanent of an $n \times n$ matrix A over $\mathbb{C}$ is bounded by its operator norm via $|\text{perm}(A)| \leq \|A\|^n$. Motivated by questions related to boson sampling and quantum computing, Scott Aaronson and Travis Hance [1] asked for a characterization of matrices for which this bound is nearly tight. Appealing to inverse Littlewood-Offord theory (developed by Terry Tao and Van Vu), several papers [2, 22] attempted to address this question with only limited successs. We settled it by proving (a quantified version of) the following, which was a conjecture of Aaronson.

Theorem 1. *If $|perm(A)| \geq \|A\|^n/n^{100}$, then A must have a readily identifiable form where virtually every row and column is dominated by a single entry of very large modulus.*

Although the result deterministically holds for all matrices over $\mathbb{C}$, our proof is entirely probabilistic, using Talagrand's inequality, hypercontractivity, and Khintchine's inequality. It would be interesting to extend our results to obtain deterministic algorithms approximating $|\text{perm}(A)|$ to within an additive error of $\pm\varepsilon\|A\|^n$, which was a driving motivation of [1].

Permutations and entropy [3]

We prove the following, which was conjectured by Tom Leighton and Ankur Moitra [21] in connection with the algorithmic problem of sorting under partial information.

Theorem 2. *If σ is a random (not necessarily uniform) permutation of $\{1, 2, \ldots, n\}$ satisfying*

$$|\mathbb{P}(\sigma(i) < \sigma(j)) - 1/2| > \varepsilon \qquad \forall i \neq j \tag{1}$$

for fixed $\varepsilon > 0$, then σ has entropy at most $(1 - \delta)\log(n!)$, where $\delta > 0$ depends only on ε.

That is, the assumption (1) implies a significant loss of information relative to the entropy of a uniform distribution (namely $\log(n!)$). Leighton and Moitra proved this in the special case where $\mathbb{P}(\sigma(i) < \sigma(j)) > 1/2 + \varepsilon$ for all $i < j$. Our proof uses a mix of probabilistic and graph theoretic techniques including a version of Szemerédi's regularity lemma, a coupling argument, and martingale concentration results. Continuing our work on this, we would like to better understand the dependence of δ on ε, for which we conjecture $\delta = C\varepsilon^4$ should suffice.

Stability in the Erdős–Ko–Rado theorem [13]

For $n \geq 2k+1$, the *Kneser graph*, $K(n,k)$, has as vertices the k-element subsets of $[n] := \{1,2,\ldots,n\}$ with vertices A and B adjacent iff $A \cap B = \emptyset$. Recall that the *independence number*, $\alpha(G)$, of a graph G is the maximum size of a set of vertices containing no edges. In the language of Kneser graphs, the classical Erdős–Ko–Rado theorem [17] says $\alpha(K(n,k)) = \binom{n-1}{k-1}$ (and that a largest independent set consists of all k-sets containing some fixed element of $[n]$).

Following a trend of considerable recent interest, Béla Bollobás and various co-authors [8] considered this classical result in a probabilistic setting and asked when the same behavior is likely to hold in the random subgraph $K_p(n,k) \subseteq K(n,k)$ gotten by retaining edges independently with probability p.

Combined with earlier work, our results completely determine the order of magnitude of the "threshold" for this property. A formal discussion is given in the unabridged research statement on my webpage. Here, I just highlight one case addressing what both [8] and [6] identified as the most interesting aspect of the problem:

Theorem 3. *There is a fixed $\varepsilon > 0$ such that for $n = 2k+1$ and $p > 1-\varepsilon$,*

$$\lim_{n\to\infty} \mathbb{P}[\alpha(K_p(n,k)) = \alpha(K(n,k))] = 1.$$

The key steps of the proof rely on spectral techniques and results from *Fourier analysis on the slice* to show that certain collections of vertices contain many edges in the Kneser graph. A natural conjecture is that the above result should hold for all $p > 3/4$; though proving this conjecture seems quite difficult, it is the most obvious direction for future research.

Fractional matchings in k-out hypergraphs [14]

Hypergraphs are extremely useful generalizations of graphs but are notoriously difficult to work with. An *r-uniform hypergraph $\mathcal{H}$* on vertex set V is a collection of r-subsets of V—thus 2-uniform hypergraphs coincide with graphs. A *perfect matching* of a hypergraph is a subset of the members of $\mathcal{H}$ ("edges") that partitions the vertex set. The linear programming relaxation of this is a *perfect fractional matching*, that is a nonnegative weighting of the edges of $\mathcal{H}$ for which the weights of the edges containing any $v \in V$ sum to 1.

For any $r > 2$, determining if an r-uniform hypergraph has a perfect matching is an NP-complete problem [20], making the question both important and computationally intractible (unless P = NP). Motivated by this and by a conjecture of Alan Frieze and Gregory Sorkin [18], we prove the following, which extends earlier results for graphs to hypergraphs. (The natural k-out model is defined in the unabridged version.)

Theorem 4. *For each r, there exists a constant $C = C(r)$ such that if $k > C$, then the r-uniform k-out hypergraph has a perfect fractional matching almost surely (i.e., with probability tending to 1 as $|V| \to \infty$).*

Moreover, for r-uniform r-partite hypergraphs (again, defined in online version), a stronger result precisely characterizes (almost surely) the optimal fractional covers. A key step in our proof is establishing that a certain expansion-type property deterministically implies the existence of perfect fractional matchings in r-uniform hypergraphs. The most natural (albeit difficult) open problem for future research would be to extend these results to perfect matchings; however, a more promising direction would be to determine the best possible value of $C(r)$.

Continuing and future work

I am interested in many different areas within discrete math. The above descriptions (and their corresponding parts in the unabridged version) already mention quite a few intriguing research questions that I am still thinking about. Here I discuss just two examples (among many) of other problems that interest me.

(a) *Relating the chromatic number of a graph to those of its subgraphs*

Suppose G is a graph on n vertices with chromatic number $\chi(G) = \chi$, and let G_p be the random subgraph of G where each edge is included independently with probability p. Alon, Krivelevich, and Sudakov [4] proved $\mathbb{E}[\chi(G_{1/2})] \geq C\chi/\log(n)$, and Boris Bukh suggested this could be improved to $\mathbb{E}[\chi(G_{1/2})] \geq C\chi/\log(\chi)$. To get a feel for this conjecture, let us consider $\alpha(G_{1/2})$, for which we can show $\mathbb{E}[\alpha(G_{1/2})] \leq C\alpha(G)\log(n/\alpha(G))$ using Túran's theorem together with Boole's inequality. This proves Bukh's conjecture when $\chi(G) = \Theta(n/\alpha(G))$, and we thus need only consider the rare graphs for which this is not the case. As intermediate steps, I propose the weaker conjectures

Conjecture 1: $\mathbb{E}[\chi(G_{1/2})] \geq C\chi_f/\log(\chi_f)$,

Conjecture 2: $\mathbb{E}[\chi(G_p)] \geq c_p\chi^p$,

where χ_f is the fractional chromatic number of G and $c_p > 0$ is a constant depending on p. As before, Conjecture 1 holds when $\chi_f = \Theta(n/\alpha(G))$, and the hope is that we might be able to take advantage of linear programming duality. As for Conjecture 2, if $1/p$ is an integer, it holds with $c_p = 1$; however, taking G to be a large odd cycle already shows we need $c_p \leq 2/3^p$. In fact, even proving $\mathbb{E}[\chi(G_{0.99})] \geq C\chi^{0.5+\varepsilon}$ would be interesting.

(b) *Intersecting families of graphs*

Let H be a fixed graph, and suppose $\mathcal{F}$ is a family of graphs on a common vertex set V such that $A \cap B$ contains a copy of H for all $A, B \in \mathcal{F}$. How large can $|\mathcal{F}|$ be?

Normalizing, let $\mu(\mathcal{F}) = |\mathcal{F}|/2^{|V|(|V|-1)/2}$—thus $\mu(\mathcal{F})$ is the fraction of graphs on V that are in $\mathcal{F}$. Since $\mathcal{F}$ is intersecting, $\mu(\mathcal{F}) \leq 1/2$, whereas we can acheive $\mu(\mathcal{F}) = 2^{-|E(H)|}$ by taking $\mathcal{F}$ to be the collection of graphs containing some fixed copy of H.

An old conjecture of Simonovits and Sós says that if H is a triangle, then $\mu(\mathcal{F}) \leq 1/8$, so the above construction is best possible. The first progress was by Chung, Graham, Frankl, and Shearer [9], who used an entropy argument to show that if H is not bipartite, then $\mu(\mathcal{F}) \leq 1/4$. Decades later Ellis, Filmus, and Friedgut [16] made a groundbreaking improvement of $\mu(\mathcal{F}) \leq 1/8$ for non-bipartite H (completely settling the triangle case).

For bipartite H, much less is known. If H is a disjoint union of stars, there are families with $\mu(\mathcal{F}) = 1/2 - o(1)$, and if H is a 3-edge path, there are families with $\mu(\mathcal{F}) > 2^{-|E(H)|}$. And that's all we know. In particular, *we don't know a single example of a bipartite H for which $\mu(\mathcal{F})$ is bounded away from $1/2$,* though Alon [5] conjectures that if H is not a disjoint union of stars, then this is always the case. This would follow from the special case that H is a 3-edge path, but even proving it for, say, $H = K_{100,100}$ would be very interesting. Thus far, my approaches have suggested intriguing interactions between random graphs, Fourier analysis, and information theory, which in turn have led to many questions of independent interest. For example, we get an isoperimetric-type problem:

Question: If $\mathcal{F} \subseteq 2^{[n]}$ is increasing (closed upwards) and balanced ($|\mathcal{F}| = 2^{n-1}$), with small vertex boundary ($|\{A \subseteq [n] \ : \ \exists i \ s.t. \ A \setminus \{i\} \notin \mathcal{F}, \ A \cup \{i\} \in \mathcal{F}\}| \ll 2^n$), must $\mathcal{F}$ be noticeably correlated with some weighted majority function?

References

[1] Scott Aaronson and Travis Hance. Generalizing and derandomizing Gurvits's approximation algorithm for the permanent. *Quantum Inf. Comput.*, 14(7-8):541–559, 2014.

[2] Scott Aaronson and Hoi Nguyen. Near invariance of the hypercube. *Israel Journal of Mathematics*, pages 1–33, 2016.

[3] Huseyin Acan, Pat Devlin, and Jeff Kahn. An entropy conjecture of Leighton and Moitra. 2016+. In preparation.

[4] Noga Alon, Michael Krivelevich, and Benny Sudakov. Subgraphs with a large cochromatic number. *Journal of Graph Theory*, 25(4):295–297, 1997.

[5] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2004.

[6] József Balogh, Béla Bollobás, and Bhargav P Narayanan. Transference for the Erdős–Ko–Rado theorem. In *Forum of Mathematics, Sigma*, volume 3, page e23. Cambridge Univ Press, 2015.

[7] Ross Berkowitz and Pat Devlin. A stability result using the matrix norm to bound the permanent. 2016+. Submitted. Preprint, arXiv:1606.07474.

[8] Béla Bollobás, Bhargav P Narayanan, and Andrei M Raigorodskii. On the stability of the Erdős–Ko–Rado theorem. *Journal of Combinatorial Theory, Series A*, 137:64–78, 2016.

[9] Fan RK Chung, Ronald L Graham, Peter Frankl, and James B Shearer. Some intersection theorems for ordered sets and graphs. *Journal of Combinatorial Theory, Series A*, 43(1):23–37, 1986.

[10] Pat Devlin. Integer subsets with high volume and low perimeter. *Integers*, 12, 2012.

[11] Pat Devlin and Edinah K Gnang. Some integer formula encodings and related algorithms. *Advances in Applied Mathematics*, 51(4):536–541, 2013.

[12] Pat Devlin and Edinah K Gnang. Primes appearing in prime tower factorization. 2014. arXiv:1204.5251.

[13] Pat Devlin and Jeff Kahn. On stability in the Erdős–Ko–Rado Theorem. *SIAM J. Discrete Math.*, 30(2):1283–1289, 2015.

[14] Pat Devlin and Jeff Kahn. Perfect fractional matchings for hypergraphs in the k-out model. 2016+. In preparation.

[15] Pat Devlin and Howard J Nuer. A strange family of Calabi-Yau 3-folds. *String-Math 2014*, 93:245, 2014.

[16] David Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Journal of the European Mathematical Society*, 14(3):841–885, 2012.

[17] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *The Quarterly Journal of Mathematics*, 12(1):313–320, 1961.

[18] Alan Frieze and Gregory B Sorkin. Efficient algorithms for three-dimensional axial and planar random assignment problems. *Random Structures & Algorithms*, 46(1):160–196, 2015.

[19] Leonid Gurvits. On the complexity of mixed discriminants and related problems. In *Mathematical Foundations of Computer Science 2005*, pages 447–458. Springer, 2005.

[20] Richard M Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[21] Tom Leighton and Ankur Moitra. On entropy and extensions of posets, 2011.

[22] Hoi Nguyen. On matrices of large permanent. Private communication, 2016.