RESEARCH STATEMENT

Abigail Raz (ajr224@math.rutgers.edu)

My research interests are in discrete mathematics, which includes both combinatorics and graph theory. I am particularly interested in questions where probability may be introduced either via using probabilistic tools to tackle deterministic problems or taking classical deterministic results and examining their behavior in a random setting. I also hope to expand my work to combinatorial problems that can be tackled with other non-combinatorial tools, for example, those with connections to model theory. In this statement I focus on the work I have done during my graduate career. The first problem takes a classical graph theoretic result and now asks the same question of the Erdős-Rényi random graph. The second is solely focused on a facet of the behavior of the Erdős-Rényi random graph. The final result diverges from the first two and provides a counter-example to an extension of the union-closed sets conjecture — an open problem in extremal combinatorics.

STRUCTURE OF THE LARGEST SUBGRAPHS OF $G_{n,p}$ WITH A GIVEN MATCHING NUMBER

This is joint work with my advisor Jeff Kahn; a preliminary manuscript should be available soon.

Recall that the matching number of a graph G is the size of a largest set of disjoint edges and is denoted $\nu(G)$. We say the *size* of a graph is the number of edges. In what follows "largest" will refer to the size of the graph. Let us say a graph G has the *EG Property* if for each k every largest subgraph with matching number k has one of two forms:

(a) All edges are within a set of vertices of size $2k+1$.
(b) All edges are incident to a set of vertices of size k.

In 1959, Erdős and Gallai proved the following theorem in [9].

Theorem 1. *K_n has the EG Property.*

Erdős conjectured that this result can be extended from K_n to $\mathcal{K} = \binom{[n]}{l}$ for all l.

Conjecture 1. *(Erdős' Matching Conjecture) The largest subhypergraphs of $\mathcal{K} = \binom{[n]}{l}$ with matching number k have* $\max\left\{\binom{l(k+1)-1}{l}, \binom{n}{l} - \binom{n-k}{l}\right\}$ *hyperedges.*

The case $l = 2$ is Theorem 1. The conjecture has also been proved for $l = 3$ [11, 12, 21] and when k is not too close to n/l [11, 15]. Note that as k changes the optimal configuration shifts between two forms.

Result: In [17], I showed two regimes of p where Theorem 1 can be extended to $G_{n,p}$ (the usual Erdős-Rényi random graph) and one where it cannot.

Theorem 2. *If $p \geq \frac{8\log n}{n}$ or $p \ll 1/n$, then with high probability* [1] *$G_{n,p}$ has the EG Property. Furthermore, if $\beta/n < p < \frac{\gamma \log n}{n}$, where $\gamma < 1/3$ and $\beta > 4\log 2$ then w.h.p. $G_{n,p}$ does not have the EG Property.*

[1] With high probability ("w.h.p.") means with probability tending to 1 as $n \to \infty$

Theorem 2 gives a good rough understanding of the ranges of p where we do or do not expect the EG-property. The most interesting part of the argument is for the upper range, where the proof uses the Tutte-Berge formula together with various probabilistic arguments and tools. In the middle range we show that the EG property fails at $k = \nu(G_{n,p})$. Further details on the proof can be found in my unabridged research statement.

Future Work: Going forward, I would like to close the two gaps for p in Theorem 2. I also wish to understand, in the range where $G_{n,p}$ does not have the EG-property, what happens for k other than (and maybe not too close to) $\nu(G_{n,p})$, since the negative part of the theorem considers only $k = \nu(G_{n,p})$. It would also be interesting to see what, if anything, can be said about the forms of the largest subgraphs with matching number k when $G_{n,p}$ fails to have the EG-property.

TIGHT UPPER TAIL BOUNDS FOR THE NUMBER OF CYCLES IN $G_{n,p}$

This is joint work with my advisor Jeff Kahn; a preliminary manuscript should be available soon.

Let $\mathbb{G} = G(n,p)$ be the usual (Erdős-Rényi) random graph. For a fixed graph H define $\xi_H = \xi_H^{n,p}$ to be the number of copies of H in $\mathbb{G}$. It is a much studied and surprisingly difficult problem to understand the *upper tail* of the distribution of ξ_H, for example, to estimate

$$\mathbb{P}(\xi_H > 2\mathbb{E}\xi_H). \tag{1}$$

(The naive first guess, that this probability behaves like $\exp[-\Omega(\mathbb{E}\xi_H)]$, turns out to be far from the truth.) The best result for general H is due to Janson, Oleszkiewicz, and Ruciński who proved, in [26], that for any H, p, and η we have

$$\exp[-O_{H,\eta}(M_H(n,p)\ln(1/p))] < \mathbb{P}(\xi_H > (1+\eta)\mathbb{E}\xi_H) < \exp[-\Omega_{H,\eta}(M_H(n,p))]. \tag{2}$$

Thus they determined the upper tail up to a factor of $\ln(1/p)$ in the exponent. The definition of $M_H(n,p)$ can be found in [26], but we omit it here. The first progress towards closing the $\ln(1/p)$ gap was made by DeMarco and Kahn in [7] and Chatterjee in [3] who independently closed it for triangles, showing that the lower bound is the truth (up to a constant in the exponent). Later, in [6], DeMarco and Kahn closed the $\ln(1/p)$ gap for cliques. Additionally, the gap has been closed for all graphs and large p (i.e. $p > n^{-\alpha_H}$) [4, 20] and so-called strictly balanced graphs H and low p (i.e $p \leq n^{-v/e}\log^{C_h} n$) [30, 31, 27]. Recently, in [5], Cook and Dembo closed the gap — including determining the correct constant in the exponent — for cycles when $p \gg n^{-1/2}$ (among other results).

Result: In [18], I closed the gap for all fixed cycles (up to a constant in the exponent) and any p. Formally, for any fixed l let $\xi_l = \xi_l(\mathbb{G})$ be the number of C_l's (cycles of length l) in $\mathbb{G}$.

Theorem 3. *For any fixed l, $\eta > 0$, and $p \in [0,1]$,*

$$\mathbb{P}(\xi_l > (1+\eta)\mathbb{E}\xi_l) < \exp[-\Omega_{\eta,l}(\min\{n^2p^2\ln(1/p), n^l p^l\})].$$

(This matches the lower bound in (2) in the case of cycles — where $M_{C_l}(n,p) = n^2p^2$.) The proof relies on various applications of two standard large deviation bounds. The constant goal

is to balance the fact that a vertex of large degree may be in many cycles with the fact that the probability we have many vertices of large degree is small.
Future Work: In the future, a first question is to see if the methods used in our paper can lead to a simpler proof of the result for cliques found in [6] and potentially extend the methods to tackle the question of bounding the upper tail of ξ_H for all fixed regular graphs. In [6], DeMarco and Kahn conjectured that the lower bound given in [26] is always the truth. However, in [28], this was recently disproved by Šileikis and Warnke for an infinite family of graphs and p close to their appearance threshold. In [28] they still conjecture that the lower bound given in [26] is the truth for any strictly balanced graph or with the additional assumption that p is sufficiently above the appearance threshold. Additionally, they leave open the question of formulating a new upper tail conjecture for graphs which are not strictly balanced.

A COUNTER-EXAMPLE TO AN EXTENSION OF THE UNION-CLOSED SETS CONJECTURE

We say a family of sets, $\mathcal{A}$, is union-closed if for all $A, B \in \mathcal{A}$ we have $A \cup B \in \mathcal{A}$. The union-closed sets conjecture states that if a finite family of sets $\mathcal{A} \neq \{\emptyset\}$ is union-closed, then there is an element which belongs to at least half the sets in $\mathcal{A}$. In 2001, D. Reimer showed in [24] that the average set size of a union-closed family, $\mathcal{A}$, is at least $\frac{1}{2}\log_2|\mathcal{A}|$. In order to do so, he showed that all union-closed families satisfy a particular condition (that we will refer to as "Reimer's condition"), which in turn implies the preceding bound. The question of whether Reimer's condition alone is enough to imply that there is an element in at least half of the sets was raised in the context of T. Gowers' polymath project on the union-closed sets conjecture [13].
Result: I exhibited a counter-example with ground set $\{1, \ldots, 8\}$ that satisfies Reimer's condition, but fails to have an element in at least half the sets. The counter-example is minimal (both in the size of the ground set and the number of sets in the family) and can be found in [23]. I generated the counter-example by constructing an auxiliary directed graph and examining how changes to the potential counter-example affected the degrees of the auxiliary digraph. Using this one can show that a counter-example exists and easily recover it.
Future Work: It is important to note that this counter-example is very far from union-closed. In the future, it would be interesting to investigate whether additional counter-examples that are closer to being union-closed can be obtained by similar methods. I believe this line of inquiry would be well suited to an undergraduate student, particularly as this could be attacked with combinatorial machinery combined with clever computer searches.

CONTINUING AND FUTURE WORK

In the previous section I noted a few directions for further research connected to my results. Here I discuss a couple of additional problems that are also of interest to me.

1. A question in percolation.
   In the standard model of percolation theory, we consider the d-dimensional integer lattice (the graph consisting of the vertex set $\mathbb{Z}^d$ together with edges between any two points

with Euclidean distance 1). Percolation theory, generally, examines the behavior of the random subgraph of $\mathbb{Z}^d$ where each edge is, independently, "open" with probability p and "closed" with probability $1-p$. A standard first question is "What is the probability that the origin can reach infinitely many vertices in our random subgraph?". The following question concerns only a finite $n \times m$ subset of $\mathbb{Z}^2$ and instead of choosing edges to be open or closed, we now assign them directions (thus producing a random directed graph).

Question 1. *Take a $n \times m$ sublattice of $\mathbb{Z}^2$ (e.g. the subgraph induced by all vertices with coordinates (i,j) where $1 \leq i \leq n$ and $1 \leq j \leq m$). We randomly assign each edge a direction. The vertical edges will be assigned "up" with probability $\frac{1}{2}$ (and "down" also with probability $\frac{1}{2}$). The horizontal edges will be assigned "right" with probability p (and "left" with probability $1-p$). Let E be the event that there is a directed path from the left side to the right side. Is $\mathbb{P}(E)$ monotone with p?*

This question was told to me by Bhargav Narayanan and seems obviously true. Yet, to my knowledge, it remains open. If this statement is true a next step would be to more precisely describe the behavior of $\mathbb{P}(E)$. Percolation theory contains many similar questions, and, while some have simple answers, many more remain open.

2. Piercing axis-parallel boxes.
Let $\mathcal{D}$ be a family of boxes in $\mathbb{R}^d$ with axis-parallel edges. One can ask the following two basic questions:

   (a) What is the maximum number of disjoint boxes (call this $\nu_d(\mathcal{D})$)?

   (b) What is the minimum number of points needed to pierce every box (call this $\tau_d(\mathcal{D})$)?

   The following is a brief summary of what is known in the deterministic case. A result of Gallai (see [16]) says that every family of intervals, $\mathcal{D}$, (i.e. boxes in $\mathbb{R}$) have $\nu_1(\mathcal{D}) = \tau_1(\mathcal{D})$. In [32] Wegner conjectured that τ_2/ν_2 is bounded by 2 for families of rectangles (i.e boxes in $\mathbb{R}^2$), while, in [14], Gýarfás and Lehel conjectured that τ_2/ν_2 is bounded by a constant. The best known lower bound, $\tau_2 \geq \lfloor 5\nu_2/3 \rfloor$, is attained by a construction due to Fon-Der-Flaass and Kostochka in [10], while a simple result of Karoyli's, in [19], shows that for families of axis-parallel boxes in $\mathbb{R}^d$ we have $\tau_d \leq \nu_d(1+\log \nu_d)^{d+1}$. A few recent papers, such as [22], improved this when every two intersecting boxes intersect at a corner.

   While it would be interesting to tackle the deterministic question, one can also investigate the case of random boxes. Here a random sub-interval of [0,1] is determined by two random endpoints each chosen independently with respect to the uniform measure on [0,1]. Thus a d-dimensional random box is the product of d independent random sub-intervals of [0,1]. For a family of n independently chosen d-dimensional random boxes we let $\nu_d(n)$ be the largest number of pairwise disjoint boxes and $\tau_d(n)$ be the smallest number of points needed to pierce every box. In 2001, Coffman, Lueker, Spencer and Winker, in [8], showed that, w.h.p., for $d = 2$ we have $\nu_2 = \Theta(\sqrt{n})$ and for $d \geq 3$ we have $\Omega(\sqrt{n}) \leq \nu_d(n) \leq O(\sqrt{n \log^{d-1} n})$. In 2011, in [29], Tran showed that for any fixed $d \geq 2$ w.h.p. we have $\Omega_d(\sqrt{n}\log^{d/2-1} n) = \tau_d(b) = O_d(\sqrt{n}\log^{d/2-1}\log\log n)$. Furthermore, he conjectured that the lower bound is the truth. It is unlikely that the methods used in Tran's result can be easily extended to a proof of this conjecture; however, it would be interesting to see a different auxiliary hypergraph than the one used by Tran yields a better bound.

References

[1] N. Alon and J.H Spencer. *The Probabilistic Method*. New York: Wiley, 2015.

[2] J. Beck and W. Chen. *Irregularities of Distribution*. Cambridge: Cambridge Univ. Pr., 1987.

[3] S. Chatterjee. "The missing log in large deviations for triangle counts". In: *Random Structures & Algorithms* 40 (4 2011), pp. 437–451.

[4] S. Chatterjee and A. Dembo. "Nonlinear large deviations". In: *Adv. Math.* 299 (2016), pp. 396–450.

[5] N. A. Cook and A. Dembo. "Large deviations of subgraph counts for sparse Erdős-Rényi graphs". In: *ArXiv e-prints* (Sept. 2018). arXiv: 1809.11148 [math.PR].

[6] R. DeMarco and J. Kahn. "Tight upper tails bounds for cliques". In: *Random Structures & Algorithms* 41.4 (), pp. 469–487. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20440.

[7] R. DeMarco and J. Kahn. "Upper tails for triangles". In: *ArXiv e-prints* (May 2010). arXiv: 1005.4471 [math.PR].

[8] E.G. Coffman, G.S. Leuker, J. Spencer and P.M. Winkler. "Packing random rectangles". In: *Probab. Theory Relat. Fields* 120 (2001), pp. 585–599.

[9] P. Erdős and T. Gallai. "On maximal paths and circuits of graphs". In: *Acta Mathematica Academiae Scientiarum Hungarica* 10.3-4 (1959), pp. 337–356.

[10] D. Fon-Der Flaass and A. Kostochka. "Covering boxes by points". In: *Discrete Mathematics* 120 (1-3 1993), pp. 269–275.

[11] P. Frankl. "Improved bounds for Erdős' Matching Conjecture". In: *Journal of Combinatorial Theory, Series A* 120 (2013), pp. 1068–1072.

[12] P. Frankl. "On the maximum number of edges in a hypergraph with given matching number". In: *Discrete Applied Mathematics* 216.3 (2017), pp. 562–581.

[13] T. Gowers. *FUNC1 - strengthenings, variants, potential counterexamples*. Jan. 2016. URL: https://gowers.wordpress.com/2016/01/29/func1-strengthenings-variants-potential-counterexamples.

[14] A. Gýarfás and J. Lehel. "Covering and coloring problems for relatives of intervals". In: *Discrete Mathematics* 55 (2 1985), pp. 167–180.

[15] H. Huang, P.-S. Loh, and B. Sudakov. "The size of a hypergraph and its matching number." In: *Combinatorics Probability and Computing* 21 (2012), pp. 442–450.

[16] A. Hajnal and J. Surányi. "Über die Auösung von Graphen in vollstäandie Teilgraphen". In: *Ann. Univ. Sci. Budapest, Eötvös Sect. Math* (1958), pp. 113–121.

[17] J. Kahn and A. Raz. "Structure of the largest subgraphs of $G_{n,p}$ with a given matching number". In Preperation. 2018+.

[18] J. Kahn and A. Raz. "Tight upper tail bounds for cycles". In Preperation. 2018+.

[19] G. Károlyi. "On point covers of parallel rectangles". In: *Periodica Mathematica Hungarica* 22 (2 1991), pp. 105–107.

[20] E. Lubetzky and Y. Zhao. "On the variational problem for upper tails in sparse random graphs". In: *Random Struct. Alg.* 50 (2017), pp. 420–436.

[21] T. Łuczak and K. Mieczkowska. "On Erdős' extremal problem on matchings in hypergraphs". In: *Journal of Combinatorial Theory, Series A* 124 (2014), pp. 178–194.

[22] M. Chudnovsky, S. Spirkl, and S. Zerbib. "Piercing axis-parallel boxes". In: *Electronic Journal of Combinatorics* 25.1 (2018).

[23] A. Raz. "Note on the union-closed sets conjecture". In: *Electronic Journal of Combinatorics* 24.3 (2017).

[24] D. Reimer. "An Average Set Size Theorem". In: *Combinatorics, Probability, and Computing* 12.1 (2003), pp. 89–93.

[25] I. Rival, ed. *Graphs and Order.* 1st ed. Vol. 147. Nato Science Series C: Springer, 1985.

[26] S. Janson, K. Oleszkiewicz and A. Ruciński. "Upper tails for subgraph counts in random graphs". In: *Israel Journal of Mathematics* 142 (2004), pp. 61–92.

[27] M. Šileikis. "On the upper tail of counts of strictly balanced subgraphs". In: *Electronic Journal of Combinatorics* 19.1 (2012).

[28] M. Šileikis and L. Warnke. "A counterexample to the DeMarco-Kahn Upper Tail Conjecture". In: *ArXiv e-prints* (Sept. 2018). arXiv: `1809.09595 [math.PR]`.

[29] L. Tran. "Piercing random boxes". In: *Random Structures & Algorithms* 28 (3 2011), pp. 365–380.

[30] V.H. Vu. "On the concentration of multivariate polynomials with small expectation." In: *Random Struct. Alg.* 16 (2000), pp. 344–363.

[31] L. Warnke. "On the missing log in upper tail estimates". In: *ArXiv e-prints* (Dec. 2016). arXiv: `1612.08561 [math.PR]`.

[32] G. Wegner. "Über eine kombinatorisch-geometrische Frage von Hadwiger und Debrunner". In: *Israel Journal of Mathematics* 3 (4 1965), pp. 187–198.

[33] D.B. West. *Introduction to Graph Theory.* New Jersey: Prentice Hall, 2001.