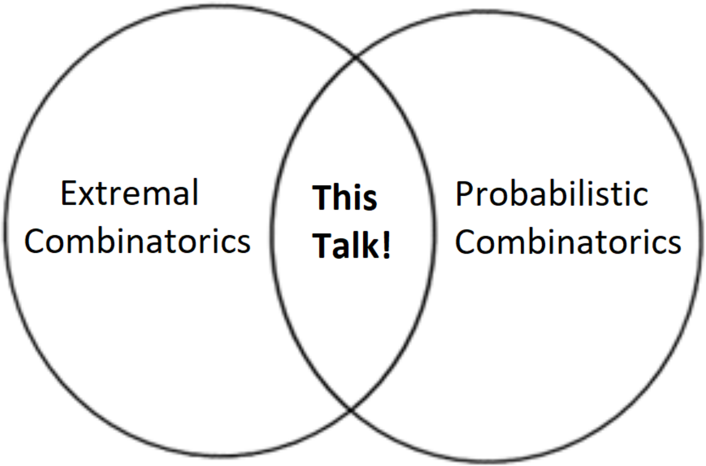
Extremal Problems for Random Objects

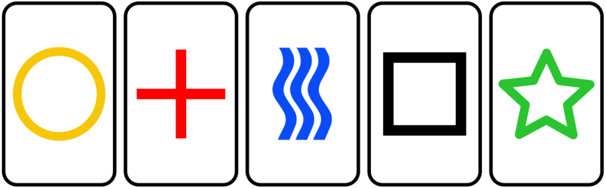
Sam Spiro, Rutgers University


Extremal
Combinatorics
This
Talk!
Probabilistic
Combinatorics

Part I: Card Guessing with Feedback

In the "Complete Feedback Model," we start with a deck of mn cards where there are n card types each appearing with multiplicity m.

In the "Complete Feedback Model," we start with a deck of mn cards where there are n card types each appearing with multiplicity m.



Let $\mathcal{C}^+_{m,n}$ and $\mathcal{C}^-_{m,n}$ be the maximum and minimum expected scores that the player can get in the complete feedback model.

Let $\mathcal{C}^+_{m,n}$ and $\mathcal{C}^-_{m,n}$ be the maximum and minimum expected scores that the player can get in the complete feedback model.

Theorem (Diaconis-Graham, 1981)

For n fixed,

$$\mathcal{C}^{\pm}_{m,n} = m \pm c_n\sqrt{m} + o_n(\sqrt{m}).$$

Another Game

Another Game

In the "partial feedback model", the Guesser guesses the next card and is only told whether their guess was correct or not.

Another Game

In the "partial feedback model", the Guesser guesses the next card and is only told whether their guess was correct or not. If $\mathcal{P}^{+}_{m,n}$ is the maximum expected score in this model,

$$m \leq \mathcal{P}^{+}_{m,n}$$

Another Game

In the "partial feedback model", the Guesser guesses the next card and is only told whether their guess was correct or not. If $\mathcal{P}^+_{m,n}$ is the maximum expected score in this model,

$$m \leq \mathcal{P}^+_{m,n} \leq \mathcal{C}^+_{m,n}$$

Another Game

In the "partial feedback model", the Guesser guesses the next card and is only told whether their guess was correct or not. If $\mathcal{P}^+_{m,n}$ is the maximum expected score in this model,

$$m \le \mathcal{P}^+_{m,n} \le \mathcal{C}^+_{m,n} = m + o_n(m)$$

Another Game

In the "partial feedback model", the Guesser guesses the next card and is only told whether their guess was correct or not. If $\mathcal{P}^+_{m,n}$ is the maximum expected score in this model,

$$m \leq \mathcal{P}^+_{m,n} \leq \mathcal{C}^+_{m,n} = m + o_n(m).$$

What happens when n is large?

Theorem (Diaconis-Graham-He-S., 2020)

For m fixed,

$$\mathcal{C}^{+}_{m,n} \sim H_m \log(n),$$
$$\mathcal{C}^{-}_{m,n} = \Theta(n^{-1/m}),$$

where H_m is the mth harmonic number.

Theorem (Diaconis-Graham-He-S., 2020)

For m fixed,

$$\mathcal{C}^+_{m,n} \sim H_m \log(n),$$
$$\mathcal{C}^-_{m,n} = \Theta(n^{-1/m}),$$

where H_m is the mth harmonic number.

With this we have the trivial bounds

$$m \leq \mathcal{P}^+_{m,n} \leq \mathcal{C}^+_{m,n} = O_m(\log n).$$

$$m \leq \mathcal{P}^{+}_{m,n} \leq O_m(\log n).$$

$$m \le \mathcal{P}^{+}_{m,n} \le O_m(\log n).$$

Theorem (Diaconis-Graham-He-S., 2020)

There exist $c, C > 0$ such that if n is sufficiently large in terms of m, we have

$$m + c\sqrt{m} \le \mathcal{P}^{+}_{m,n}$$

$$m \leq \mathcal{P}^{+}_{m,n} \leq O_m(\log n).$$

Theorem (Diaconis-Graham-He-S., 2020)

There exist $c, C > 0$ such that if n is sufficiently large in terms of m, we have

$$m + c\sqrt{m} \leq \mathcal{P}^{+}_{m,n} \leq m + Cm^{3/4} \log m.$$

$$m \le \mathcal{P}^{+}_{m,n} \le O_m(\log n).$$

Theorem (Diaconis-Graham-He-S., 2020)

There exist $c, C > 0$ such that if n is sufficiently large in terms of m, we have

$$m + c\sqrt{m} \le \mathcal{P}^{+}_{m,n} \le m + Cm^{3/4}\log m.$$

Theorem (Z. Nie, 2022)

If $n \gg m$, then

$$\mathcal{P}^{+}_{m,n} = m + \Theta(\sqrt{m}).$$

A Partial Proof

A Partial Proof

Lemma

Assume that we have played in the partial feedback model for $t-1$ rounds such that we have guessed card type i a total of g_i times

A Partial Proof

Lemma

Assume that we have played in the partial feedback model for $t-1$ rounds such that we have guessed card type i a total of g_i times, and let S be the total number of points scored.

A Partial Proof

Lemma

Assume that we have played in the partial feedback model for $t-1$ rounds such that we have guessed card type i a total of g_i times, and let S be the total number of points scored. Given this, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S}.$$

A Partial Proof

Corollary

$$\mathcal{P}^{+}_{m,n} \leq 3m + o(m).$$

A Partial Proof

Corollary

$$\mathcal{P}^{+}_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S}$$

A Partial Proof

Corollary

$$\mathcal{P}^+_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S} \approx \frac{m}{mn - g_i},$$

A Partial Proof

Corollary

$$\mathcal{P}^{+}_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S} \approx \frac{m}{mn - g_i},$$

At most one i is guessed more than $mn/2$ times.

A Partial Proof

Corollary

$$\mathcal{P}^+_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S} \approx \frac{m}{mn - g_i},$$

At most one i is guessed more than $mn/2$ times. Every other j has $\Pr[\pi_t = j] \leq \frac{2}{n}$ for all t.

A Partial Proof

Corollary

$$\mathcal{P}^+_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S} \approx \frac{m}{mn - g_i},$$

At most one i is guessed more than $mn/2$ times. Every other j has $\Pr[\pi_t = j] \leq \frac{2}{n}$ for all t. Thus in expectation at most $mn \cdot (2/n) = 2m$ cards are guessed correctly from this part

A Partial Proof

Corollary

$$\mathcal{P}^+_{m,n} \leq 3m + o(m).$$

For all i and t, we have

$$\Pr[\pi_t = i] \leq \frac{m}{mn - g_i - S} \approx \frac{m}{mn - g_i},$$

At most one i is guessed more than $mn/2$ times. Every other j has $\Pr[\pi_t = j] \leq \frac{2}{n}$ for all t. Thus in expectation at most $mn \cdot (2/n) = 2m$ cards are guessed correctly from this part, and in total at most $3m$ are guessed correctly in expectation. $\square$

Card Guessing

Card Guessing

The probability of drawing four aces in a row with a deck shuffled uniformly at random is $1/270725$.

Adversarial Card Guessing

The probability of drawing four aces in a row with a deck shuffled uniformly at random is $1/270725$.

Adversarial Card Guessing

The probability of drawing four aces in a row with a deck shuffled uniformly at random is $1/270725$.

More precisely, we are now considering a two player game played by Shuffler and Guesser.

Adversarial Card Guessing

The probability of drawing four aces in a row with a deck shuffled uniformly at random is $1/270725$.

More precisely, we are now considering a two player game played by Shuffler and Guesser. Let $\mathcal{C}_{m,n}(G, S)$ be the expected number of points Guesser scores when the two players follow strategies G, S.

$$\Theta_m(n^{-1/m}) \le \mathcal{C}_{m,n}(G, \mathrm{Uniform}) \le H_m \log n + o_m(\log n).$$

$$\Theta_m(n^{-1/m}) \leq \mathcal{C}_{m,n}(G, \mathrm{Uniform}) \leq H_m \log n + o_m(\log n).$$

Theorem (S., 2021)

There exists a strategy S' *for Shuffler so that*

$$\mathcal{C}_{m,n}(\mathsf{G}, \mathsf{S}') \leq \log n + o_m(\log n),$$

and this bound is best possible.

Theorem

There exists a strategy S' *for Shuffler so that* $\mathcal{C}_{m,n}(\mathsf{G}, \mathsf{S}') \leq \log n + o_m(\log n)$.

Theorem

There exists a strategy S' *for Shuffler so that* $\mathcal{C}_{m,n}(\mathsf{G},\mathsf{S}') \leq \log n + o_m(\log n)$.

A strategy that gives this is the "greedy strategy", which is such that if there are r types of cards remaining in the deck, then Shuffler draws each of these card types with probability r^{-1}.

Theorem

There exists a strategy S' *for Shuffler so that* $\mathcal{C}_{m,n}(\mathsf{G}, \mathsf{S}') \leq \log n + o_m(\log n)$.

A strategy that gives this is the "greedy strategy", which is such that if there are r types of cards remaining in the deck, then Shuffler draws each of these card types with probability r^{-1}. E.g. if the deck has a hundred 1's and one 2, we draw a 1 or 2 with probability $\frac{1}{2}$.

Theorem

There exists a strategy S' *for Shuffler so that* $\mathcal{C}_{m,n}(\mathsf{G}, \mathsf{S}') \leq \log n + o_m(\log n)$.

A strategy that gives this is the "greedy strategy", which is such that if there are r types of cards remaining in the deck, then Shuffler draws each of these card types with probability r^{-1}. E.g. if the deck has a hundred 1's and one 2, we draw a 1 or 2 with probability $\frac{1}{2}$. This gives the desired bound due to a variant of the coupon collector problem. $\square$

Theorem (S., 2021)

The greedy strategy is the unique strategy that minimizes the number of correct guesses if Guesser tries to maximize their score.

Theorem (S., 2021)

The greedy strategy is the unique strategy that minimizes the number of correct guesses if Guesser tries to maximize their score.

Interestingly, the greedy strategy is also the "unique" strategy which maximizes the number of correct guesses if Guesser tries to minimize their score.

Future Problems

Future Problems

Problem

Prove non-trivial bounds for the partial feedback model with adversarial shufflings.

Future Problems

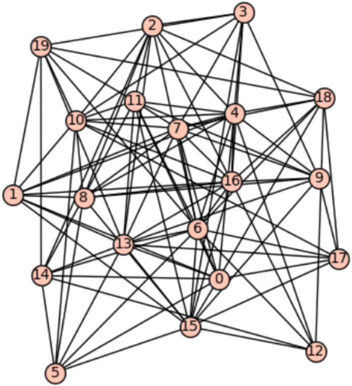
Problem

Prove non-trivial bounds for the partial feedback model with adversarial shufflings.

Conjecture

The minimum expected score one can get with partial feedback is asymptotic to m.

Part II: Turán's Problem in Random Graphs

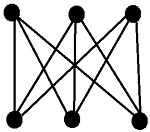


Define the Turán number $\mathrm{ex}(n, F)$ to be the maximum number of edges that an F-free graph on n vertices can have.

Define the Turán number $\mathrm{ex}(n, F)$ to be the maximum number of edges that an F-free graph on n vertices can have.

Theorem (Mantel 1907)

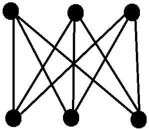
$$\mathrm{ex}(n, K_3) = \left\lfloor n^2/4 \right\rfloor.$$



Define the Turán number $\mathrm{ex}(n, F)$ to be the maximum number of edges that an F-free graph on n vertices can have.

Theorem (Mantel 1907)

$$\mathrm{ex}(n, K_3) = \left\lfloor n^2/4 \right\rfloor.$$



Theorem (Erdős-Stone 1946)

$$\mathrm{ex}(n, F) = \left(1 - \frac{1}{\chi(F) - 1} + o(1)\right)\binom{n}{2}.$$

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p.

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p. Let $\mathrm{ex}(G_{n,p}, F)$ be the maximum number of edges that an F-free subgraph of $G_{n,p}$ can have.

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p. Let $\mathrm{ex}(G_{n,p}, F)$ be the maximum number of edges that an F-free subgraph of $G_{n,p}$ can have. For example,

$$\mathrm{ex}(G_{n,1}, F) = \mathrm{ex}(n, F)$$

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p. Let $\mathrm{ex}(G_{n,p}, F)$ be the maximum number of edges that an F-free subgraph of $G_{n,p}$ can have. For example,

$$\mathrm{ex}(G_{n,1}, F) = \mathrm{ex}(n, F),$$

and with high probability

$$p \cdot \mathrm{ex}(n, F) \lesssim \mathrm{ex}(G_{n,p}, F) \lesssim p\binom{n}{2}.$$

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p. Let $\mathrm{ex}(G_{n,p}, F)$ be the maximum number of edges that an F-free subgraph of $G_{n,p}$ can have. For example,

$$\mathrm{ex}(G_{n,1}, F) = \mathrm{ex}(n, F),$$

and with high probability

$$p \cdot \mathrm{ex}(n, F) \lesssim \mathrm{ex}(G_{n,p}, F) \lesssim p\binom{n}{2}.$$

The lower bound is tight when $p = 1$.

Let $G_{n,p}$ be the random graph on n vertices where each edge is included independently and with probability p. Let $\mathrm{ex}(G_{n,p}, F)$ be the maximum number of edges that an F-free subgraph of $G_{n,p}$ can have. For example,

$$\mathrm{ex}(G_{n,1}, F) = \mathrm{ex}(n, F),$$

and with high probability

$$p \cdot \mathrm{ex}(n, F) \lesssim \mathrm{ex}(G_{n,p}, F) \lesssim p\binom{n}{2}.$$

The lower bound is tight when $p = 1$. The upper bound is tight if p is "small."

$$\frac{1}{2}p\binom{n}{2} \lesssim \mathrm{ex}(G_{n,p}, K_3) \lesssim p\binom{n}{2},$$

with the lower bound tight for $p = 1$ and the upper bound tight for $p \ll n^{-1/2}$.

$$\frac{1}{2}p\binom{n}{2} \lesssim \mathrm{ex}(G_{n,p}, K_3) \lesssim p\binom{n}{2},$$

with the lower bound tight for $p = 1$ and the upper bound tight for $p \ll n^{-1/2}$.

Theorem (Frankl-Rödl 1986)

Whp,

$$\mathrm{ex}(G_{n,p}, K_3) \sim \frac{1}{2}p\binom{n}{2} \qquad p \gg n^{-1/2}.$$

$$\frac{1}{2}p\binom{n}{2} \lesssim \mathrm{ex}(G_{n,p}, K_3) \lesssim p\binom{n}{2},$$

with the lower bound tight for $p = 1$ and the upper bound tight for $p \ll n^{-1/2}$.

Theorem (Frankl-Rödl 1986)

Whp,

$$\mathrm{ex}(G_{n,p}, K_3) \sim \frac{1}{2}p\binom{n}{2} \qquad p \gg n^{-1/2}.$$

Theorem (Conlon-Gowers, Schacht 2010)

Whp,

$$\mathrm{ex}(G_{n,p}, F) = p \cdot \left(1 - \frac{1}{\chi(F) - 1} + o(1)\right)\binom{n}{2} \qquad p \gg n^{-1/m_2(F)},$$

where $m_2(F) = \max\{\frac{e(F')-1}{v(F')-2} : F' \subseteq F\}$.

What happens for bipartite graphs?

What happens for bipartite graphs?

Conjecture

If F is a bipartite graph which is not a forest, then whp

$$\mathrm{ex}(G_{n,p}, F) = \begin{cases} \Theta(p \cdot \mathrm{ex}(n, F)) & p \gg n^{-1/m_2(F)}, \\ (1 + o(1))p\binom{n}{2} & p \ll n^{-1/m_2(F)}. \end{cases}$$

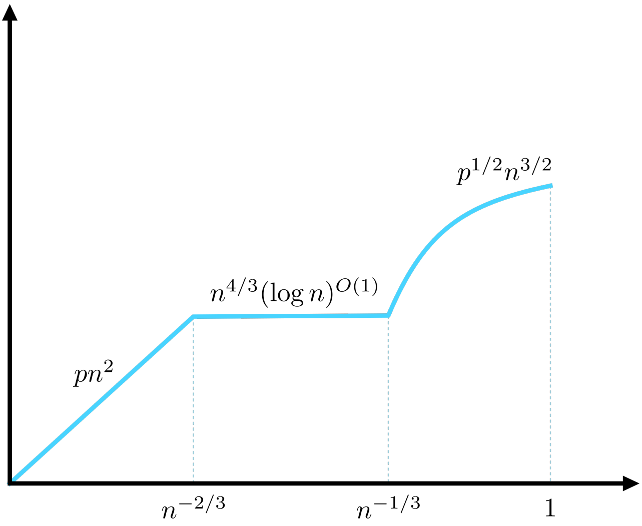
What happens for bipartite graphs?

Conjecture

If F is a bipartite graph which is not a forest, then whp

$$\mathrm{ex}(G_{n,p}, F) = \begin{cases} \Theta(p \cdot \mathrm{ex}(n, F)) & p \gg n^{-1/m_2(F)}, \\ (1 + o(1))p\binom{n}{2} & p \ll n^{-1/m_2(F)}. \end{cases}$$

This conjecture turns out to be completely false!



Plot of $\mathsf{ex}(G_{n,p}, C_4)$ (Füredi 1991)

Conjecture (McKinley-S.)

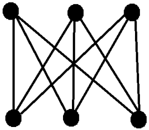
If F is a graph with $\mathrm{ex}(n, F) = \Theta(n^{\alpha})$ for some $\alpha \in (1, 2]$, then whp

$$\mathrm{ex}(G_{n,p}, F) = \max\{\Theta(p^{\alpha-1} n^{\alpha}), n^{2-1/m_2(F)} (\log n)^{O(1)}\},$$

provided $p \gg n^{-1/m_2(F)}$.

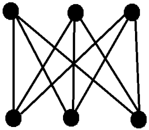
Theorem (Kővari-Sós-Turán 1954)

$$\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s}).$$



Theorem (Kővari-Sós-Turán 1954)

$$\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s}).$$

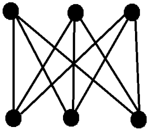


Theorem (Morris-Saxton 2013)

$$\mathrm{ex}(G_{n,p}, K_{s,t}) = O(p^{1-1/s} n^{2-1/s}) \text{ for } p \text{ large.}$$

Theorem (Kővari-Sós-Turán 1954)

$$\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s}).$$



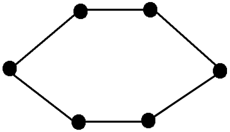
Theorem (Morris-Saxton 2013)

$$\mathrm{ex}(G_{n,p}, K_{s,t}) = O(p^{1-1/s} n^{2-1/s}) \textit{ for } p \textit{ large.}$$

Moreover, this bound is tight whenever $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$.

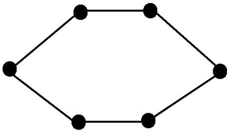
Theorem (Bondy-Simonovits 1974)

$$\mathrm{ex}(n, C_{2b}) = O(n^{1+1/b}).$$



Theorem (Bondy-Simonovits 1974)

$$\mathrm{ex}(n, C_{2b}) = O(n^{1+1/b}).$$

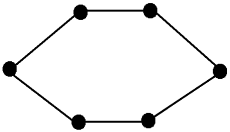


Theorem (Morris-Saxton 2013)

$$\mathrm{ex}(G_{n,p}, C_{2b}) = O(p^{1/b} n^{1+1/b}) \textit{ for } p \textit{ large.}$$

Theorem (Bondy-Simonovits 1974)

$$\mathrm{ex}(n, C_{2b}) = O(n^{1+1/b}).$$



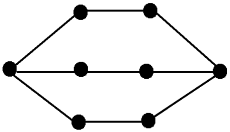
Theorem (Morris-Saxton 2013)

$$\mathrm{ex}(G_{n,p}, C_{2b}) = O(p^{1/b} n^{1+1/b}) \textit{ for } p \textit{ large.}$$

Moreover, this is tight whenever $\mathrm{ex}(n, \{C_3, C_4, \ldots, C_{2b}\}) = \Theta(n^{1+1/b})$.

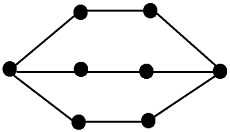
Theorem (Faudree-Simonovits 1974)

$$\mathrm{ex}(n, \theta_{a,b}) = O(n^{1+1/b}).$$



Theorem (Faudree-Simonovits 1974)

$$\mathrm{ex}(n, \theta_{a,b}) = O(n^{1+1/b}).$$



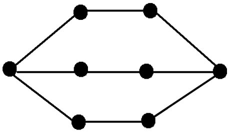
Theorem (McKinley-S. 2023)

For $a \geq 100$,

$$\mathrm{ex}(G_{n,p}, \theta_{a,b}) = O(p^{1/b} n^{1+1/b}) \textit{ for } p \textit{ large.}$$

Theorem (Faudree-Simonovits 1974)

$$\mathrm{ex}(n, \theta_{a,b}) = O(n^{1+1/b}).$$



Theorem (McKinley-S. 2023)

For $a \geq 100$,

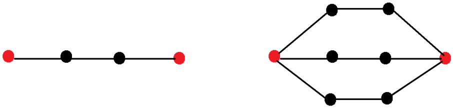
$$\mathrm{ex}(G_{n,p}, \theta_{a,b}) = O(p^{1/b} n^{1+1/b}) \textit{ for } p \textit{ large.}$$

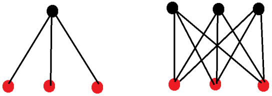
Moreover, this bound is tight whenever a is sufficiently large in terms of b.

Theorem (Bukh-Conlon 2015)

If T^ℓ is the "ℓth power of a balanced tree" and ℓ is sufficiently large, then

$$\mathrm{ex}(n, T^\ell) = \Omega(n^{2-\rho(T)}).$$

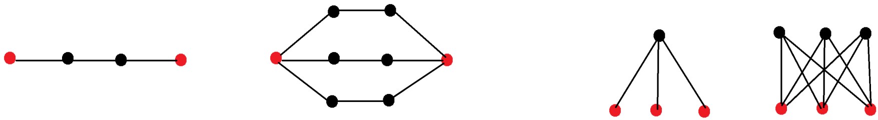




Theorem (Bukh-Conlon 2015)

If T^ℓ is the "ℓth power of a balanced tree" and ℓ is sufficiently large, then

$$\mathrm{ex}(n, T^\ell) = \Omega(n^{2-\rho(T)}).$$

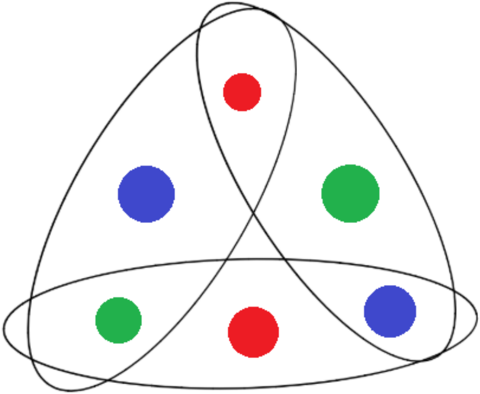


Theorem (S. 2022)

$$\mathrm{ex}(G_{n,p}, T^\ell) = \Omega(p^{1-\rho(T)} n^{2-\rho(T)}),$$

provided ℓ is sufficiently large.

Hypergraphs



Theorem (S.-Verstraëte 2021)

Let $K^r_{s_1,\dots,s_r}$ denote the complete r-partite r-graph with parts of sizes $s_1,\dots,s_r$. There exist constants $\beta_1,\beta_2,\beta_3,\gamma$ depending on $s_1,\dots,s_r$ such that, for s_r sufficiently large in terms of $s_1,\dots,s_{r-1}$, we have whp

$$\mathrm{ex}(G^r_{n,p},K^r_{s_1,\dots,s_r}) = \begin{cases} \Theta\left(pn^r\right) & n^{-r} \ll p \le n^{-\beta_1}, \\ n^{r-\beta_1+o(1)} & n^{-\beta_1} \le p \le n^{-\beta_2}(\log n)^{\gamma}, \\ \Theta(p^{1-\beta_3} n^{r-\beta_3}) & n^{-\beta_2}(\log n)^{\gamma} \le p \le 1. \end{cases}$$

Question

Does the McKinley-Spiro conjecture extend to hypergraphs?

Question

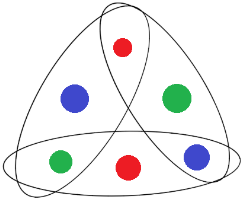
Does the McKinley-Spiro conjecture extend to hypergraphs?

Theorem (Nie-S. 2023 (Informal))

Any hypergraph which is not Sidorenko fails to have a flat middle range.

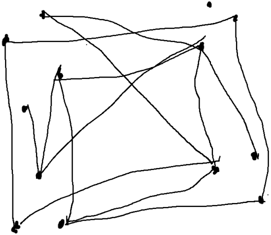
Other Hypergraph Results

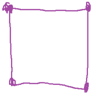
1. Solved for loose triangles (Nie-S.-Verstraëte 2020; Nie 2023)
2. Solved for loose even cycles of uniformity $r \geq 4$ (Mubayi-Yepremyan 2020; Nie 2023)
3. (Non-optimal) bounds for Berge cycles (S.-Verstraëte 2021; Nie 2023)
4. *Improved lower bound for non-Sidorenko hypergraphs (Nie-S. 2023)
5. *Lifting upper bounds from graphs to hypergraphs (Nie-S. 20XX++)



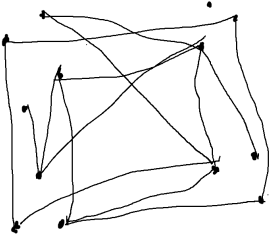
Lower Bound Techniques

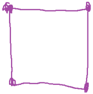
Lower Bound Techniques

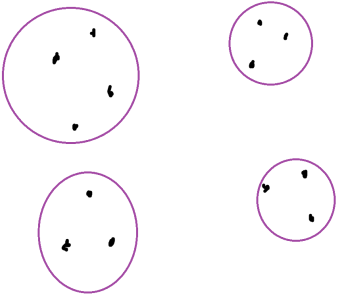
$G_{n,p} =$ 

$ex(P_n, F) =$ 

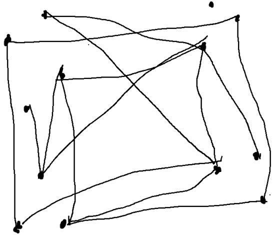
Lower Bound Techniques

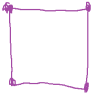
$G_{n,p} =$ 

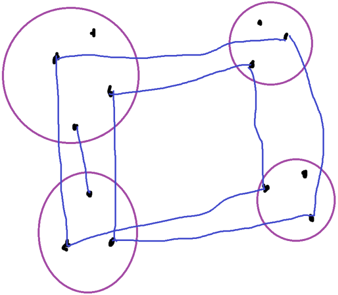
$ex(Pn, F) =$ 



Lower Bound Techniques

$G_{n,p} =$ 

$ex(pn, F) =$ 

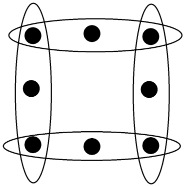


Future Problems

Future Problems

Problem

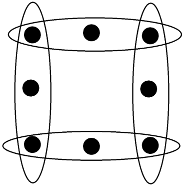
Prove tight bounds for the 3-uniform loose 4-cycles.



Future Problems

Problem

Prove tight bounds for the 3-uniform loose 4-cycles.



Problem

Prove tight bounds for subdivisions of complete bipartite graphs.

Thanks!