The Erdős Matching Conjecture

and related questions

Andrey Kupavskii

University of Birmingham

https://www.birmingham.ac.uk/

Moscow Institute of Physics and Technology

https://mipt.ru/english/

Definitions

Denote $\binom{[n]}{k}$: the set of all k-element subsets of $[n]$.

A subset $\mathcal{F} \subset \binom{[n]}{k}$ is called a *family*.

A *matching of size s* in $\mathcal{F}$: s pairwise disjoint sets $F_1, \ldots, F_s \in \mathcal{F}$.

The *matching number $\nu(\mathcal{F})$* of $\mathcal{F}$: the size of the largest matching in $\mathcal{F}$.

If $\nu(\mathcal{F}) = 1$, then $\mathcal{F}$ is *intersecting*: any two sets in $\mathcal{F}$ intersect.

The extremal quantity

Define

$$e_k(n,s) := \max\left\{|\mathcal{F}| : \mathcal{F} \subset \binom{[n]}{k},\ \nu(\mathcal{F}) < s\right\}.$$

Theorem (Erdős-Ko-Rado, 1938-1961)

$$e_k(n,2) = \binom{n-1}{k-1} \quad \text{for} \quad n \geqslant 2k.$$

The theorem is tight. Consider the family $\left\{A \in \binom{[n]}{k} : 1 \in A\right\}$.

For $n = 2k$ the family $\binom{[2k-1]}{k}$ has the same cardinality.

The extremal quantity

Define

$$e_k(n,s) := \max\left\{|\mathcal{F}| : \mathcal{F} \subset \binom{[n]}{k},\ \nu(\mathcal{F}) < s\right\}.$$

Theorem (Erdős-Ko-Rado, 1938-1961)

$$e_k(n,2) = \binom{n-1}{k-1} \quad \text{for} \quad n \geqslant 2k.$$

The theorem is tight. Consider the family $\left\{A \in \binom{[n]}{k} : 1 \in A\right\}$.

For $n = 2k$ the family $\binom{[2k-1]}{k}$ has the same cardinality.

The extremal quantity

Define

$$e_k(n,s) := \max\left\{|\mathcal{F}| : \mathcal{F} \subset \binom{[n]}{k},\ \nu(\mathcal{F}) < s\right\}.$$

Theorem (Erdős-Ko-Rado, 1938-1961)

$$e_k(n,2) = \binom{n-1}{k-1} \quad \text{for} \quad n \geqslant 2k.$$

The theorem is tight. Consider the family $\left\{A \in \binom{[n]}{k} : 1 \in A\right\}$.

For $n = 2k$ the family $\binom{[2k-1]}{k}$ has the same cardinality.

The Erdős Matching Conjecture

How to construct a large family $\mathcal{F} \subset \binom{[n]}{k}$, satisfying $\nu(\mathcal{F}) < s$?

$$\mathcal{A}^{(k)}(n,s) := \left\{ A \in \binom{[n]}{k} : A \cap [s-1] \neq \emptyset \right\}, \quad \mathcal{B}^{(k)}(n,s) := \binom{[sk-1]}{k}.$$

We have $|\mathcal{A}^{(k)}(n,s) = \binom{n}{k} - \binom{n-s+1}{k}$, $|\mathcal{B}^{(k)}(n,s)| = \binom{sk-1}{k}$.

The Erdős Matching Conjecture, 1965

For $n \geqslant sk$ we have

$$e_k(n,s) = \max\left\{ |\mathcal{A}^{(k)}(n,s)|, |\mathcal{B}^{(k)}(n,s)| \right\}.$$

Put $x := s/n$. If k is fixed and $s \to \infty$: $|\mathcal{A}|/\binom{n}{k} \to 1 - (1-x)^k$,
$|\mathcal{B}|/\binom{n}{k} \to (kx)^k$.

The Erdős Matching Conjecture

How to construct a large family $\mathcal{F} \subset \binom{[n]}{k}$, satisfying $\nu(\mathcal{F}) < s$?

$$\mathcal{A}^{(k)}(n,s) := \left\{ A \in \binom{[n]}{k} : A \cap [s-1] \neq \emptyset \right\}, \quad \mathcal{B}^{(k)}(n,s) := \binom{[sk-1]}{k}.$$

We have $\quad |\mathcal{A}^{(k)}(n,s) = \binom{n}{k} - \binom{n-s+1}{k}, \qquad |\mathcal{B}^{(k)}(n,s)| = \binom{sk-1}{k}.$

The Erdős Matching Conjecture, 1965

For $n \geqslant sk$ we have

$$e_k(n,s) = \max\left\{ |\mathcal{A}^{(k)}(n,s)|, |\mathcal{B}^{(k)}(n,s)| \right\}.$$

Put $x := s/n$. If k is fixed and $s \to \infty$: $\quad |\mathcal{A}|/\binom{n}{k} \to 1-(1-x)^k,$

$$|\mathcal{B}|/\binom{n}{k} \to (kx)^k.$$

The Erdős Matching Conjecture

How to construct a large family $\mathcal{F} \subset \binom{[n]}{k}$, satisfying $\nu(\mathcal{F}) < s$?

$$\mathcal{A}^{(k)}(n,s) := \left\{A \in \binom{[n]}{k} : A \cap [s-1] \neq \emptyset\right\}, \quad \mathcal{B}^{(k)}(n,s) := \binom{[sk-1]}{k}.$$

We have $|\mathcal{A}^{(k)}(n,s) = \binom{n}{k} - \binom{n-s+1}{k}$, $\quad |\mathcal{B}^{(k)}(n,s)| = \binom{sk-1}{k}$.

The Erdős Matching Conjecture, 1965

For $n \geqslant sk$ we have

$$e_k(n,s) = \max\left\{|\mathcal{A}^{(k)}(n,s)|, |\mathcal{B}^{(k)}(n,s)|\right\}.$$

Put $x := s/n$. If k is fixed and $s \to \infty$: $|\mathcal{A}|/\binom{n}{k} \to 1-(1-x)^k$,
$|\mathcal{B}|/\binom{n}{k} \to (kx)^k$.

Asymptotic fractional version of the EMC.

A fractional matching for $\mathcal{F} \subset 2^{[n]}$: a function $w : \mathcal{F} \to [0,1]$, such that

$$\sum_{F \in \mathcal{F} : i \in F} w(F) \leqslant 1 \qquad \text{for every element } i \in [n].$$

Fractional matching number $\nu^*(\mathcal{F})$:
the size of the largest fractional matching in $\mathcal{F}$.

Conjecture A (Alon et. al., 2012)
Let $x \in [0, 1/k]$ be fixed and let $\mathcal{F}_n \subset \binom{[n]}{k}$ be a sequence of families such that $\nu^*(\mathcal{F}) \leqslant xn$. Then

$$\limsup_{n \to \infty} \frac{|\mathcal{F}|}{\binom{n}{k}} \leqslant \max\left\{1 - (1-x)^k, (kx)^k\right\}.$$

Asymptotic fractional version of the EMC.

A fractional matching for $\mathcal{F} \subset 2^{[n]}$: a function $w : \mathcal{F} \to [0,1]$, such that

$$\sum_{F \in \mathcal{F} : i \in F} w(F) \leqslant 1 \qquad \text{for every element } i \in [n].$$

Fractional matching number $\nu^*(\mathcal{F})$:

the size of the largest fractional matching in $\mathcal{F}$.

Conjecture A (Alon et. al., 2012)

Let $x \in [0, 1/k]$ be fixed and let $\mathcal{F}_n \subset \binom{[n]}{k}$ be a sequence of families such that $\nu^*(\mathcal{F}) \leqslant xn$. Then

$$\limsup_{n \to \infty} \frac{|\mathcal{F}|}{\binom{n}{k}} \leqslant \max\left\{1 - (1-x)^k, (kx)^k\right\}.$$

Deviation of sums of random variables

$\mathbf{X} := (X_1, \ldots, X_k)$: $X_i \geqslant 0$ are i.i.d. random variables, $\mathrm{E}[X_i] = x$.

$$m_k(x) := \sup_{\mathbf{X}} \Pr[X_1 + \ldots + X_k \geqslant 1].$$

Note: $m_k(x) = 1$ for $x \geqslant 1/k$.

Conjecture B (Łuczak, Mieczkowska, Šileikis, 2017)

$$m_k(x) = \max\left\{1 - (1-x)^k, (kx)^k\right\}.$$

Case $k = 2$ was resolved by Hoeffding and Shrikhande (1955).

Related conjectures of Samuels (1966) and Feige (2006) speak about random variables that are not necessarily identically distributed.

Results of Samuels (1966, 1968) imply Conjecture B for $k = 3, 4$.

Deviation of sums of random variables

$\mathbf{X} := (X_1, \ldots, X_k)$: $X_i \geqslant 0$ are i.i.d. random variables, $\mathrm{E}[X_i] = x$.

$$m_k(x) := \sup_{\mathbf{X}} \Pr[X_1 + \ldots + X_k \geqslant 1].$$

Note: $m_k(x) = 1$ for $x \geqslant 1/k$.

Conjecture B (Łuczak, Mieczkowska, Šileikis, 2017)

$$m_k(x) = \max\left\{1 - (1-x)^k, (kx)^k\right\}.$$

Case $k = 2$ was resolved by Hoeffding and Shrikhande (1955).

Related conjectures of Samuels (1966) and Feige (2006) speak about random variables that are not necessarily identically distributed.

Results of Samuels (1966, 1968) imply Conjecture B for $k = 3, 4$.

Conjectures A and B are equivalent

(Alon, Frankl, Huang, Rödl, Ruciński, Sudakov, 2012)

Conjecture B $\Rightarrow$ Conjecture A. Take the largest $\mathcal{F}$ with $\nu^*(\mathcal{F}) \leqslant xn$. By LP-duality: $(\nu^* = \tau^*)$ there exists $w : [n] \to [0,1]$, such that

$$\sum_{i \in [n]} w(i) = xn, \quad \text{and} \quad \sum_{i \in F} w(i) \geqslant 1 \text{ for every } F \in \mathcal{F}.$$

Define a random variable: $w(t)$ for a randomly chosen $t \in [n]$. Form $\mathbf{v} := (w(t_1), \ldots, w(t_k))$. Then

$$m_k(x) \geqslant \Pr\left[\sum_{i=1}^{k} w(t_i) \geqslant 1\right] \gtrsim \Pr[\mathbf{v} \in \mathcal{F}] = \frac{|\mathcal{F}|}{\binom{n}{k}}.$$

What do we know about EMC?

True for $k = 2$ (Erdős and Gallai, 1959)

True for $k = 3$ (Łuczak and Mieczkowska, 2014, for large s; Frankl, 2017, for all s).

True for $n > n_0(s, k)$ (Erdős, 1965)

True for $n > 2k^3 s$ (Bollobás, Daykin, Erdős, 1976)

True for $n > 100ks^2$ (Frankl, Füredi, 1987)

True for $n > 3k^2 s$ (Huang, Loh, Sudakov, 2012)

True for $n \geqslant (2s-1)k - s$ (Frankl, 2013)

Connections to large deviation bounds, frac versions (Alon et. al. 2012)

Equivalence of Conjectures A, B (Łuczak, Mieczkowska, Šileikis, 2017)

$e_k(n, s) \leqslant (s-1)\binom{n-1}{k-1}$ (Frankl, 1987)

What do we know about EMC?

True for $k = 2$ (Erdős and Gallai, 1959)

True for $k = 3$ (Łuczak and Mieczkowska, 2014, for large s; Frankl, 2017, for all s).

True for $n > n_0(s, k)$ (Erdős, 1965)

True for $n > 2k^3 s$ (Bollobás, Daykin, Erdős, 1976)

True for $n > 100ks^2$ (Frankl, Füredi, 1987)

True for $n > 3k^2 s$ (Huang, Loh, Sudakov, 2012)

True for $n \geqslant (2s-1)k - s$ (Frankl, 2013)

Connections to large deviation bounds, frac versions (Alon et. al. 2012)

Equivalence of Conjectures A, B (Łuczak, Mieczkowska, Šileikis, 2017)

$e_k(n, s) \leqslant (s-1)\binom{n-1}{k-1}$ (Frankl, 1987)

New results

Theorem (AK, Frankl, 2018+)

There exists s_0 such that the EMC is true for $s \geqslant s_0$, any k and $n \geqslant \frac{5}{3}sk - \frac{2}{3}s$.

Consequently, Conjectures A and B hold for $x < \frac{3}{5k-2}$. Previous best due to the equivalence and the result of Frankl: $x < \frac{1}{2k-1}$.

We also get a bound $e_k(n,s) \leqslant c(s-1)\binom{n-1}{k-1}$, where $c < 1$ and depends on sk/n.

Dirac-type thresholds

$m_k^d(n)$ ($f_k^d(n)$): minimum d-degree in $\mathcal{F} \subset \binom{[n]}{k}$ that guarantees the existence of a perfect (fractional) matching.

Theorem (Alon et. al., 2012; Treglown and Zhao, 2016)

If $\limsup_{n\to\infty} f_k^d(n)/\binom{n-d}{k-d} = c^*$, then

$$\limsup_{n\to\infty} m_k^d(n)/\binom{n-d}{k-d} = \max\{c^*, 1/2\}.$$

If $c^* < 1/2$, then we know $m_k^d(n)$ exactly for large n. Also,

$$f_k^d(n) \leqslant e_{k-d}(n, n/k) + 1.$$

Corollary (Kupavskii, Frankl, 2018+)

Determination of c^* for $d \geqslant 2k/5$; exact values of $m_k^d(n)$ for $d \geqslant 3k/8$.

Previous best: for $d \geqslant k/2$: Pikhurko (2008) for $d \leqslant k-2$;
Rödl, Ruciński, Szemerédi (2006) for $d = k-1$.

Dirac-type thresholds

$m_k^d(n)$ ($f_k^d(n)$): minimum d-degree in $\mathcal{F} \subset \binom{[n]}{k}$ that guarantees the existence of a perfect (fractional) matching.

Theorem (Alon et. al., 2012; Treglown and Zhao, 2016)

If $\limsup_{n\to\infty} f_k^d(n)/\binom{n-d}{k-d} = c^*$, then

$$\limsup_{n\to\infty} m_k^d(n)/\binom{n-d}{k-d} = \max\{c^*, 1/2\}.$$

If $c^* < 1/2$, then we know $m_k^d(n)$ exactly for large n. Also,

$$f_k^d(n) \leqslant e_{k-d}(n, n/k) + 1.$$

Corollary (Kupavskii, Frankl, 2018+)

Determination of c^* for $d \geqslant 2k/5$; exact values of $m_k^d(n)$ for $d \geqslant 3k/8$.

Previous best: for $d \geqslant k/2$: Pikhurko (2008) for $d \leqslant k-2$;
Rödl, Ruciński, Szemerédi (2006) for $d = k-1$.

Proof ingredients

Take the approach of Frankl as a base. The original approach uses:

1. Shifting.
2. Shadows of families with small matching numbers.
3. Inequality on the sum of sizes of cross-dependent families.

We add the following ingredients:

4. Better bounds on shadows.
5. Concentration inequalities for intersections of families and matchings.
6. Induction

Proof ingredients

Take the approach of Frankl as a base. The original approach uses:

1. Shifting.
2. Shadows of families with small matching numbers.
3. Inequality on the sum of sizes of cross-dependent families.

We add the following ingredients:

4. Better bounds on shadows.
5. Concentration inequalities for intersections of families and matchings.
6. Induction

Proof ingredients

Take the approach of Frankl as a base. The original approach uses:

1. Shifting.
2. Shadows of families with small matching numbers.
3. Inequality on the sum of sizes of cross-dependent families.

We add the following ingredients:

4. Better bounds on shadows.
5. Concentration inequalities for intersections of families and matchings.
6. Induction.

A concentration inequality

Consider a family $\mathcal{F} \subset \binom{[n]}{k}$ for $n = kt$ of *density* $\alpha := |\mathcal{F}|/\binom{n}{k}$.

Take a t-matching $\mathcal{M} \subset \binom{[n]}{k}$ uniformly at random.

Define a random variable $\eta := |\mathcal{M} \cap \mathcal{F}|$. Then $\mathrm{E}[\eta] = \alpha t$.

Theorem (AK, Frankl, 2018+)
For any $\beta > 0$, we have $\Pr\left[|\eta - \alpha t| \geqslant 2\beta\sqrt{t}\,\right] \leqslant 2e^{-\beta^2/2}$.

Proof outline

Assume $\mathcal{M} = \{M_1, \ldots, M_t\}$. We have $\eta = \eta_1 + \ldots + \eta_t$, where η_i indicates if $M_i \in \mathcal{F}$.

Define a martingale $X_0, \ldots, X_t$, where $X_i := \mathrm{E}[\eta \mid \eta_i, \ldots, \eta_1]$.

Note that $X_0 = \mathrm{E}[X_0]$ and $X_t = \eta$.

Assume $|X_i - X_{i-1}| \leqslant 2$ for any i.

Azuma-Hoeffding inequality (1963, 1967)

If $X_0, \ldots, X_t$ is a martingale and $|X_i - X_{i-1}| \leqslant 2$ for any $i \in [t]$, then $\Pr\left[|X_t - X_0| \geqslant 2\beta\sqrt{t}\right] \leqslant 2e^{-\beta^2/2}$.

Proof outline

Assume $\mathcal{M} = \{M_1, \ldots, M_t\}$. We have $\eta = \eta_1 + \ldots + \eta_t$, where η_i indicates if $M_i \in \mathcal{F}$.

Define a martingale $X_0, \ldots, X_t$, where $X_i := \mathrm{E}[\eta \mid \eta_i, \ldots, \eta_1]$.

Note that $X_0 = \mathrm{E}[X_0]$ and $X_t = \eta$.

Assume $|X_i - X_{i-1}| \leqslant 2$ for any i.

Azuma-Hoeffding inequality (1963, 1967)

If $X_0, \ldots, X_t$ is a martingale and $|X_i - X_{i-1}| \leqslant 2$ for any $i \in [t]$, then $\Pr\left[|X_t - X_0| \geqslant 2\beta\sqrt{t}\right] \leqslant 2e^{-\beta^2/2}$.

Proof outline

Assume $\mathcal{M} = \{M_1, \ldots, M_t\}$. We have $\eta = \eta_1 + \ldots + \eta_t$, where η_i indicates if $M_i \in \mathcal{F}$.

Define a martingale $X_0, \ldots, X_t$, where $X_i := \mathrm{E}[\eta \mid \eta_i, \ldots, \eta_1]$.

Note that $X_0 = \mathrm{E}[X_0]$ and $X_t = \eta$.

Assume $|X_i - X_{i-1}| \leqslant 2$ for any i. Why?

Azuma-Hoeffding inequality (1963, 1967)

If $X_0, \ldots, X_t$ is a martingale and $|X_i - X_{i-1}| \leqslant 2$ for any $i \in [t]$, then $\Pr\left[|X_t - X_0| \geqslant 2\beta\sqrt{t}\right] \leqslant 2e^{-\beta^2/2}$.

Proof of $|X_i - X_{i-1}| \leqslant 2$ for any i

$Y_{i-1} := \mathrm{E}[\eta \mid M_1, \dots, M_{i-1}]$ and $Y_i := \mathrm{E}[\eta \mid \eta_i, M_1, \dots, M_{i-1}]$.

It is sufficient to show $|Y_i - Y_{i-1}| \leqslant 2$.

Fix $M_1, \dots, M_{i-1}$, put $S := [n] \setminus (\cup_{j=1}^{i-1} M_j)$ and consider $\mathcal{F}' := \mathcal{F} \cap \binom{S}{k}$.

Kneser graph $KG_{S,k}$: vertices — $\binom{S}{k}$,
edges — pairwise disjoint sets.

$\mathcal{F}' \subset \binom{S}{k}$ gives an induced subgraph of $KG_{S,k}$.

Denote $\alpha' := |\mathcal{F}'|/\binom{|S|}{k}$, $e(\mathcal{F}')$: proportion of edges of $KG_{S,k}$ contained inside the subgraph induced on $\mathcal{F}'$.

Proof of $|X_i - X_{i-1}| \leqslant 2$ for any i

$Y_{i-1} := \mathrm{E}[\eta \mid M_1, \dots, M_{i-1}]$ and $Y_i := \mathrm{E}[\eta \mid \eta_i, M_1, \dots, M_{i-1}]$.

It is sufficient to show $|Y_i - Y_{i-1}| \leqslant 2$.

Fix $M_1, \dots, M_{i-1}$, put $S := [n] \setminus (\cup_{j=1}^{i-1} M_j)$ and consider $\mathcal{F}' := \mathcal{F} \cap \binom{S}{k}$.

Kneser graph $KG_{S,k}$: vertices — $\binom{S}{k}$,
edges — pairwise disjoint sets.

$\mathcal{F}' \subset \binom{S}{k}$ gives an induced subgraph of $KG_{S,k}$.

Denote $\alpha' := |\mathcal{F}'| / \binom{|S|}{k}$, $e(\mathcal{F}')$: proportion of edges of $KG_{S,k}$ contained inside the subgraph induced on $\mathcal{F}'$.

$Y_{i-1} := \mathrm{E}[\eta \mid M_1, \ldots, M_{i-1}]$ and $Y_i := \mathrm{E}[\eta \mid \eta_i, M_1, \ldots, M_{i-1}]$.

$\frac{Y_{i-1}}{t-i+1} = \alpha'$ ("the density of $\mathcal{F}'$")

Y_i: random variable with two values.

If $\eta_i = 1$, then $\frac{Y_i - 1}{t-i} = \frac{2e(\mathcal{F}')}{\alpha'}$ ("the average degree of $\mathcal{F}'$")

λ': the second largest absolute value of an eigenvalue of $KG_{S,k}$.

We use the Alon-Chung bound:

$$\left|\frac{2e(\mathcal{F}')}{\alpha'} - \alpha'\right| \leqslant \frac{\lambda(1-\alpha')}{d}.$$

In Kneser graphs: $\frac{\lambda}{d} = \frac{1}{t-i}$. Therefore,

$$|Y_i - Y_{i-1}| \leqslant 2.$$

Similar for $\eta_i = 0$

$Y_{i-1} := \mathrm{E}[\eta \mid M_1, \ldots, M_{i-1}]$ and $Y_i := \mathrm{E}[\eta \mid \eta_i, M_1, \ldots, M_{i-1}]$.

$\frac{Y_{i-1}}{t-i+1} = \alpha'$ ("the density of $\mathcal{F}'$")

Y_i: random variable with two values.

If $\eta_i = 1$, then $\frac{Y_i - 1}{t-i} = \frac{2e(\mathcal{F}')}{\alpha'}$ ("the average degree of $\mathcal{F}'$")

λ': the second largest absolute value of an eigenvalue of $KG_{S,k}$.

We use the Alon-Chung bound:

$$\left| \frac{2e(\mathcal{F}')}{\alpha'} - \alpha' \right| \leqslant \frac{\lambda(1-\alpha')}{d}.$$

In Kneser graphs: $\frac{\lambda}{d} = \frac{1}{t-i}$. Therefore,

$$|Y_i - Y_{i-1}| \leqslant 2.$$

Similar for $\eta_i = 0$

Theorem (AK, Frankl, 2018+)

There exists s_0 such that the following holds. Fix $s \geqslant s_0$, k and $n \geqslant \frac{5}{3}sk - \frac{2}{3}s$. Then any family $\mathcal{F} \subset \binom{[n]}{k}$ with $\nu(\mathcal{F}) < s$ satisfies

$$|\mathcal{F}| \leqslant \binom{n}{k} - \binom{n-s+1}{k}.$$

Theorem (AK, Frankl, 2018+)

Take $k, t \in \mathbb{N}$ and $n = kt$. Fix $\mathcal{F} \subset \binom{[n]}{k}$ of density α. For any $\beta > 0$ the random variable $\eta := |\mathcal{M} \cap \mathcal{F}|$, where $\mathcal{M}$ is randomly chosen t-matching, satisfies

$$\Pr\left[|\eta - \alpha t| \geqslant 2\beta\sqrt{t}\,\right] \leqslant 2e^{-\beta^2/2}.$$