Bipartite Decomposition of Graphs

Noga Alon, Tel Aviv U and IAS, Princeton.
I. Bipartite Decomposition

A bipartite decomposition of a graph $G=(V,E)$ is a family of pairwise edge disjoint complete bipartite subgraphs of $G$ so that every edge of $G$ belongs to exactly one of them.

$bc(G) =$ minimum number of bicliques in such a decomposition
Theorem [Graham-Pollak (72)]: $bc(K_n) = n - 1$

Proofs: Witsenhausen (72), Tverberg (82), Peck (84), Vishwanathan (08)

All apply linear algebra
Proofs give: If the adjacency matrix $A$ of $G$ has $p$ positive eigenvalues and $q$ negative ones then $bc(G) \geq \max \{p,q\}$

In particular, $bc(G) \geq (p+q)/2 = \text{rank}(A)/2$ (which is easy)

Extension [A (97)]: the minimum number of bicliques of $K_n$ so that each edge belongs to at least 1 of them and at most $t$ of them is $\Theta(t n^{1/t})$

(The precise minimum is not known even for $t=2$).
A geometric interpretation:

A family of d-dimensional convex polytopes is t-neighborly if for any two members A, B of the family \( d - t \leq \text{dim}(A \cap B) \leq d - 1 \).

Zaks (79), A(97): The maximum number of members in a t-neighborly family of d-dimensional aligned boxes is exactly the largest n so that \( K_n \) can be covered by d bicliques, with each edge covered at least once and at most t times.
Zaks(79): Thus, the maximum number of aligned d-dimensional boxes with disjoint interiors and with each pair having intersection of dimension d-1 is d+1.
II Related Invariants

By Graham-Pollak, for every $G$, $bc(G) \geq w(G) - 1$, where $w(G)$ is the maximum size of a clique in $G$.

Question [A-Seymour-Saks (90)]: Can we replace clique number by chromatic number, namely, is it true that for every $G$

$$bc(G) \geq \chi(G) - 1$$

Theorem [Huang-Sudakov (12)]: No!

There are graphs $G$ with $bc(G) \leq O(\chi(G)^{5/6})$.
Open:
how small can $bc(G)$ be as a function of $\chi(G)$?

Mubayi and Vishwanathan (09), A-Haviv (09):
For every $G$:

$$bc(G) \geq 2\sqrt{2 \log_2 \chi(G)}$$
Trivially, for a graph $G$ with adjacency matrix $A$
\[ \text{rank}(A) \leq 2 \text{bc}(G). \]

The rank-coloring conjecture [van Nuffelen (76), Fajtlowicz and Graffiti (87)]:

For every graph $G$, \[ \text{rank}(A) \geq \chi(G) - 1 \]

Counter-example: A-Seymour(89).

Better counter-examples: Razborov (92), Raz and Spieker (95), Nisan and Wigderson (95), Kushilevitz (95)
Open:
how small can the rank be as a function of $\chi(G)$?

Conjecture [Lovász and Saks (87)]: There is an $\epsilon > 0$ so that for every graph $G$ with chromatic number $\chi$ and adjacency matrix of rank $R$

$$R \geq 2^{(\log \chi)^\epsilon}$$

Equivalently: $\log R \geq (\log \chi)^\epsilon$
Best known bounds:

Lovett (14): For every graph with rank $R$ and chromatic number $\chi$: 

$$R \geq \Omega((\log \chi)^2)$$

Kushilevitz (95): There are graphs $G$ with 

$$R \cdot O(2^c(\log \chi)^{\log_6 3})$$

[Equivalently: $\log R \cdot O((\log \chi)^{\log_6 3})$]
Communication Complexity:

Let $f: X \times Y \rightarrow \{0,1\}$ be a Boolean function. Alice knows $x$ in $X$, Bob knows $y$ in $Y$, and they wish to compute $f(x,y)$ by communicating as few bits as possible among them.

$D(f)$ = the deterministic communication complexity of $f$ = the minimum number of bits enabling them to compute $f(x,y)$ (worst $x,y$, best protocol).
Let $M$ be the matrix of $f$, that is, $M_{x,y} = f(x,y)$

Viewing $M$ as the adjacency matrix of a bipartite graph $H$ with vertex classes $X$ and $Y$ the following holds:

$$bc(H) \leq 2^{D(f)}$$

Therefore:

$$\log_2 (\text{rank}(M)) \leq \log_2 (bc(H)) \leq D(f)$$
Let $M$ be the matrix of $f$, that is, $M_{x,y} = f(x,y)$

$$\log_2 (\text{Rank}(M)) \leq D(f)$$

The log-rank conjecture [Lovász and Saks (87)]:

$$D(f) \leq (\log_2 (\text{Rank}(M)))^{O(1)}$$

This is equivalent to the conjecture that for every graph with chromatic number $\chi$ and adjacency matrix of rank $R$

$$\log \chi \leq (\log R)^{O(1)}$$
Best known bounds:

**Lovett (14):** For every function $f$ with matrix $M$ of rank $R$: $D(f) \leq O(\sqrt{R})$

**Kushilevitz, following Nisan Wigderson (95):**
There are function $f$ with

$$D(f) \geq (\log R)^{\log_3 6}$$
III Bipartite Decomposition of Random Graphs

Fact: For every graph $G$ with $n$ vertices and independence number $\alpha$:

$$bc(G) \leq n - \alpha$$
Conjecture [Erdős, 88]):
For almost every graph $G$ with $n$ vertices and independence number $\alpha$:
$$bc(G) = n - \alpha$$

That is: for the random graph $G = G(n, 0.5)$:
$$bc(G) = n - \alpha(G)$$
with high probability (whp)

Conjecture [Chung and Peng (14)]: For the random graph $G = G(n, p)$ with any $\Omega(1) \leq p < 0.5$
$$bc(G) = n - (1 + o(1))\alpha(G)$$
whp
Thm [Chung and Peng (14)]: For any fixed $p$ satisfying $\Omega(1) \leq p \leq 0.5$

$$bc(G) \geq n - O\left((\log n)^{3+\epsilon}\right)$$

$$= n - O\left((\alpha(G)^{3+\epsilon})\right)$$

for any positive $\epsilon$. 
New Results

Thm 1: For $G = G(n, 0.5)$ and “most” values of $n$

$$bc(G) \leq n - \alpha(G) - 1$$

whp, whereas for infinitely many exceptional values of $n$

$$bc(G) \leq n - \alpha(G) - 2$$

with probability bounded away from zero. Thus, Erdős’ conjecture is (slightly) false.

Thm 2: There exists a constant $c > 0$ so that for $2/n \leq p \leq c$

$$bc(G) = n - \Theta(\alpha(G)) = n - \Theta\left(\frac{\log(np)}{p}\right)$$

whp.
A more precise statement for $G = G(n, 0.5)$:

Let $\beta(G)$ denote the largest number of vertices in an induced complete bipartite subgraph of $G$. Clearly, for every $G$: $bc(G) \leq n - \beta(G) + 1$

**Thm 1’**: (i) If $G = G(n, 0.5)$ and $\alpha(G)$ is concentrated in a single point, then so is $\beta(G)$ and $\beta(G) = \alpha(G) + 2$ whp. In this case $bc(G) \leq n - \alpha(G) - 1$ whp.

(ii) If $\alpha = \alpha(G)$ is concentrated in 2 points, then so is $\beta = \beta(G)$, and $\beta \in \{\alpha + 1, \alpha + 2, \alpha + 3\}$ whp, and each of the three possibilities holds with probability bounded away from 0 and 1. In this case $bc(G) \leq n - \alpha - 2$ with probability bounded away from 0.
Let \( f(k) = \binom{n}{k} 2^{-\binom{k}{2}} \) be the expected number of independent sets of size \( k \) in \( G = G(n, 0.5) \). The independence number \( \alpha = \alpha(G) \) of \( G \) is, whp, essentially the largest \( k \) for which \( f(k) \geq 1 \), which is \((2+o(1)) \log_2 n\).

For this value of \( k \), \( n=\Theta(k2^{k/2}) \), and the expected number of induced complete bipartite subgraphs of \( G \) of size \( k+2 \) is \( g(k)=\Theta(f(k)) \)
When for the largest $k$ with $f(k) \geq 1$, $f(k)$ is far larger than 1, but $f(k+1)$ is far smaller than 1, the second moment method implies that $\alpha(G) = k$ whp, and in this case $g$ also satisfies $g(k)>>1>>g(k+1)$ and $\beta(G) = k + 2$ whp.

When $f(k)=\Theta(1)$ then $g(k)=\Theta(1)$, and then by the (two-dimensional) Stein-Chen method, the two variables $X$=number of independent sets of size $k$ and $Y$=number of induced complete bipartite graphs of size $k+2$ are approximately independent Poisson random variables with expectations $f(k)$ and $g(k)$. 
Therefore, the events \((X=0, Y=0), (X=0, Y>0), (X>0, Y=0)\) and \((X>0, Y>0)\) all hold with probability bounded away from 0 and 1.

These correspond to \((\alpha = k - 1, \beta = k + 1)\), 
\((\alpha = k - 1, \beta = k + 2)\), \((\alpha = k, \beta = k + 1)\) and 
\((\alpha = k, \beta = k + 2)\), respectively.
Some computation can be saved by applying the one-dimensional Stein-Chen method, using the following:

**Lemma:** if $X$ and $Y$ are non-negative integers, and $X+Y=0$, then $X=Y=0$.

Thus, it suffices to show that $X$, $Y$ and $X+Y$ are approximately Poisson variables with the right expectations.
V Open Problems

Conjecture 1: For $G = G(n,0.5)$

\[ bc(G) = n - O(\alpha(G)) \]

with high probability.

Conjecture 2 (Chung-Peng): For $G = G(n,p)$

with $\Omega(1) \leq p < 0.5$

\[ bc(G) = n - (1 + o(1))\alpha(G) \]

with high probability.
Problem: Find a purely combinatorial proof of the Graham-Pollak Theorem:
$bc(K_n) = n - 1$. 