Perfect factors in random (hyper)-graphs: A survey

Van H. Vu
Department of Mathematics
Rutgers
vanvu@math.rutgers.edu

Perfect factors:

Let H be a fixed graph on v vertices and G be a graph on n vertices where n divides v. (In the rest of the talk, we always assume this.)

We say that G contains a perfect H-factor if it contains a collection of n/v vertex-disjoint copies of H. (Each point of G appears in exactly one copy in the collection.)

Similar definition for r-uniform hyperpgraphs. (Both H and G are r-uniform.)

Examples. Perfect matchings. Perfect triangle factors.

Thresholds:

Let $G(n,p)$ be the Erdös-Rényi random graph with edge density p. Let Q be a monotone increasing graph property. We say that a function $f(n)$ is a threshold for Q if

- $p/f(n) \to \infty$ implies that $G(n,p)$ has Q with probability tending to 1.
- $p/f(n) \to 0$ implies that $G(n,p)$ has Q with probability tending to 0.

Notice that if $f(n)$ is a threshold, then so is $cf(n)$ for any positive constant c.

Similar for random hypergraphs.

The problem:

The property that $G(n,p)$ contains a perfect H-factor is clearly monotone increasing. We denote by $th_H(n)$ a threshold probability for this property. (Same notation for hypergraphs)

Problem. Determine $th_H(n)$ as n goes to infinity.

Lower bounds:

For a graph F on at least two vertices define

$$d(F) := \frac{|e(F)|}{v(F) - 1}$$

where $e(F)$ and $v(F)$ are the number of edges and vertices of F, respectively.

Define $d(H) := \max d(G)$ where the max is taken over all subgraphs with at least 2 vertices of H.

We say that H is *strictly balanced* if for any proper subgraph F of H with at least two vertices

$$d(H) > d(F).$$

In this talk we mostly focus on strictly balanced graphs (hypergraphs).

Lower bounds:

Fact 1. For any H,

$$th_H(n) = \Omega(n^{-1/d(H)}).$$

Proof. The probability that a fixed point is covered by a copy of the densest subgraph of H is $\Omega(1)$.

Fact 2. Let H be a strictly balanced graph with m edges. Then

$$th_H(n) = \Omega(n^{-1/d(H)}(\log n)^{1/m}).$$

Spencer (1990) showed $\Omega(n^{-1/d(H)}(\log n)^{1/m})$ is the threshold for the property that every vertex in G is covered by a copy of H. (The copies are not necessarily disjoint.)

Conjectures:

Question 1. (Alon-Yuster 1993) For a general H, is it true that

$$th_H(n) = O(n^{-1/d(H)+o(1)})?$$

Question 2. Let H be a strictly balanced graph with m edges. Is it true that

$$th_H(n) = O(n^{-1/d(H)}(\log n)^{1/m})?$$

More convenient form: H-degree of v is the number of copies of H covering v.

Question 2. (reformulated) For a strictly balanced H, the (expected) H-degree of a vertex at the threshold is $O(\log n)$.

Special cases:

H is one edge: $th_H(n) = \Theta(\log n/n)$ (Erdös-Rényi 1960s).

H is one edge in a 3-uniform hypergraph.

Conjecture. (Schmidt-Shamir 1983) (Perfect matching in random hypergraphs) $th_H(n) = O(n^{-2} \log n)$.

H is a triangle

Conjecture. (Erdös, after the Schmidt-Shamir's conjecture; Krivelevich 1997) $th_H(n) = O(n^{-2/3}(\log n)^{1/3})$.

Theorem. (Alon-Yuster 1993, Rucinsky 1993) If $\delta(H) < d(H)$, then $th_H(n) = O(n^{1/d(H)})$, where $\delta(H)$ is the minimum degree.

Remark. $d(F)$ is about half the average degree of F.

Basic ideas. Cover (say) $.75n$ vertices by disjoint copies of the densest subgraph of H. Match the rest.

Schmidt-Shamir problem. H is one (hyper)-edge. Thus, H-degree is just the degree.

Schmidt-Shamir (1983) (expected) degree is $n^{1/2}$ is sufficient.

Frieze-Janson (1995)$n^{1/3}$.

Kim (2002) $n^{1/6}$. In general $n^{\frac{1}{5+2/(r-1)}}$ for r-uniform. (converges to $n^{1/5}$).

Triangle factor problem.

Krivelevich noticed that the arguments in Alon-Yuster and Rucinski would imply that triangle degree $n^{1/2+o(1)}$ is sufficient.

Krivelevich (1995) Triangle degree $n^{1/15}$.

Kim (2002) Triangle degree $n^{1/18}$.

Methods. Concentration, Second moment method (via Robinson-Wormald), Coupling.

Variants:

Factors in random regular (hyper)-graphs: Cooper-Frieze-Molloy-Reed (Moment method works very well here, exact results)

Fractional matching/factor: Krivelevich.

Nearly-perfect factors: Some vertices maybe covered more than once. Bohman-Frieze-Ruczinko-Thoma.

Factors in pseudo-random graphs: Krivelevich-Sudakov-Szabo.

Recent developments:

Johansson (2005) (Expected) Triangle degree n^ϵ is sufficient.

Johansson (2005) For 3-uniform hypergraph, degree n^ϵ is sufficient.

Theorem (Kahn-V. 2007). If H is strictly balanced, (expected) H-degree $C \log n$ is sufficient, for some sufficiently large C.

Corollary. For strictly balanced H, $th_H(n) = \Theta(n^{-1/d(H)}(\log n)^{1/m})$.

Corollary. The threshold for perfect matching in random 3-uniform hypergraph is $n^{-2} \log n$.

Corollary. The threshold for triangle factor in $G(n,p)$ is $n^{-2/3}(\log n)^{1/3}$.

Theorem (Kahn-V. 2007). For a general H, the threshold is $O(n^{-1/d(H)+o(1)})$.

A stronger version

The number of perfect H-factors in K_n, the complete graph on n vertices, is

$$\frac{n!}{(n/v)!}\left(\frac{1}{|Aut(H)|}\right)^{n/v} = n^{\frac{v-1}{v}n+O(n/\log n)} \tag{1}$$

where $Aut(H)$ is the automorphism group of H.

By linearity of expectations, the expectation of the number of perfect H-factors in $G(n,p)$ is

$$n^{\frac{v-1}{v}n+O(n/\log n)}p^{mn/v} = n^{o(n)}(n^{v-1}p^m)^{n/v}. \tag{2}$$

Theorem. (Kahn-V. 2007) Let H be strictly balanced. For any constant C_1 there is a constant C_2 such that the following holds. For any $p > C_2 n^{-1/d(H)}(\log n)^{1/m}$, the number of perfect H-factors in $G(n,p)$ is $(n^{v-1}p^m)^{(1-o(1))n/v}$ with probability at least $1 - n^{-C_1}$.

Method:

The approach: Start with the complete graph. Remove random edges. Control the number of perfect factors as a function of time.

$G_i = K_n \backslash \{e_1, \ldots, e_i\}$.

Let r_i be the ratio between the number of perfect H-factors in G_{i-1} containing the edge e_i and F_{i-1}. Notice that if G_{i-1} has a perfect H-factor, then $E(r_i | e_1, \ldots, e_{i-1}) = \frac{nm/v}{M_{i-1}}$, since the next random edge e_i is chosen uniformly from the set $E_n \backslash \{e_1, \ldots, e_{i-1}\}$ with M_{i-1} elements and any perfect H-factor has nm/v edges.

F_i be number of perfect factors in G_i:

$$F_i = F_0 \frac{F_1}{F_0} \frac{F_2}{F_1} \ldots \frac{F_i}{F_{i-1}} = F_0 (1 - r_1)(1 - r_2) \ldots (1 - r_i).$$

$$L_i = \log F_i = L_0 + \sum_{j=1}^{i} \log(1 - r_i).$$

Assume that all r_i are small ($\leq g(n) \to 0$)

$$-\sum_{i=1}^{T} r_i \geq \sum_{i=1}^{T} \log(1-r_i) \geq -\sum_{i=1}^{T} r_i - \sum_{i=1}^{T} r_i^2 \geq -(1+g(n))\sum_{i=1}^{T} r_i.$$

By a martingale argument:

$$L_T \geq L_0-(1+g(n))\frac{nm}{v}\log\frac{n^2}{M_T}+O(n) = L_0-(1+g(n))\frac{nm}{v}\log\frac{1}{p}+O(n).$$

Key tool: Concentration of polynomials:

Let $Y = Y(t_1, \cdots, t_n)$ be a polynomial, where the t_i are independent indicator random variables. Let $E_j(Y)$ be the maximum among the expectations of partial derivatives of order at most j of Y.

Lemma. (Kim-V. 1998, V. 2000) The following holds for any positive constant ϵ. Let Y be a multi-linear, homogeneous, normal polynomial of fixed degree d such that $E(Y)$ is significantly larger than E_j, then

$$\mathbf{P}(|Y - E(Y)| \geq \epsilon E(Y)) = n^{-\omega(1)}.$$

Work in process:

Sharp threshold.

Stopping time version.

Random regular graphs (via Sandwich Theorem).