A Counterexample to the DeMarco-Kahn Upper Tail Conjecture

Lutz Warnke

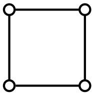
Georgia Tech

Joint work with Matas Šileikis, Czech Academy of Sciences
(who created most of these slides)

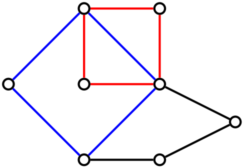
Random Structures & Algorithms 2019, Zürich

Subgraph counts: concentration

- Fixed ("small") graph H: v_H vertices, e_H edges
- X_H = number of copies of H in $\mathbb{G}(n,p)$



$H = C_4$



$X_{C_4} = 3$

- Expected number of copies:

$$\mu_H := \mathbb{E}X_H = \Theta(n^{v_H} p^{e_H}), \quad n \to \infty.$$

- **Basic concentration result** (Bollobás 1981):
  $\min_{G \subseteq H} \mu_G \to \infty$ implies $X_H \sim \mu_H$ with high probability

Subgraph counts: large deviations

- *How fast* $\mathbb{P}(|X_H - \mu_H| \geq \varepsilon\mu_H) \to 0$ *for a fixed* $\varepsilon > 0$*?*
- **Lower tail problem**: asymptotics of

$$-\log \mathbb{P}(X_H \leq (1-\varepsilon)\mu_H)$$

- **Upper tail problem**: asymptotics of

$$-\log \mathbb{P}(X_H \geq (1+\varepsilon)\mu_H)$$

- *Symmetric special case edge* $H = K_2$:
  X_H binomial $\to$ each tail $\Theta\big(\varepsilon^2 \binom{n}{2} p\big)$ by Chernoff bounds

The asymmetry of tails: triangle example

- **Triangle lower tail** (Janson–Łuczak–Ruciński 1988):

$$\mathbb{P}\left(X_{K_3} \le \tfrac{1}{2}\mu_{K_3}\right) = \exp\Big(- \Theta(\min\{n^3p^3, n^2p\})\Big)$$

- Vu 2004: if $K_m \subseteq \mathbb{G}(n,p)$ with $m = Cnp$ (large enough C),

$$X_{K_3} \ge \tbinom{m}{3} \ge 2\tbinom{n}{3}p^3$$

- Since K_m has $\binom{m}{2} = \Theta(n^2p^2)$ edges, **triangle upper tail**:

$$\begin{aligned}\mathbb{P}\left(X_{K_3} \ge 2\mu_{K_3}\right) &\ge \mathbb{P}\left(K_m \subseteq \mathbb{G}(n,p)\right)\\ &= p^{\binom{m}{2}} = \exp\Big(-\Theta(n^2p^2\log(1/p))\Big)\end{aligned}$$

- Asymmetry of both tails when $np \gg \log n$ and $p \to 0$:

$$(np)^2\log(1/p) = o(\min\{(np)^3, n^2p\})$$

Upper tail: clustered-type bounds

Theorem (Janson–Oleszkiewicz–Ruciński 2004)

There is a function $M_H^* = M_H^*(n,p)$ *such that*

$$\exp\Big(-O(M_H^*\log(1/p))\Big) \le \mathbb{P}\left(X_H \ge 2\mu_H\right) \le \exp\Big(-\Omega(M_H^*)\Big)$$

- *Lower bound*: boost expectation of X_H by planting graph F which contains $100\mu_G$ copies of *some* subgraph $G \subseteq H$.
- Lower bound is essentially $p^{e(F)}$, hence the $\log(1/p)$
- *Upper bound*: careful moment calculation + optimization
- If H is k-regular (e.g., clique, cycle), then $M_H^* \asymp n^2p^k$

Another lower bound for upper tail: disjoint copies

- Recall: for triangle K_3, a 'clustered' construction gives

$$\mathbb{P}\left(X_{K_3} \geq 2\mu_{K_3}\right) \geq \exp\Big(-O(n^2p^2\log(1/p))\Big)$$

- DeMarco and Kahn 2011: for p not too big

$$\mathbb{P}\left(X_{K_3} \geq 2\mu_{K_3}\right) \geq \mathbb{P}\left(2\mu_{K_3} \text{ vertex-disjoint triangles}\right)$$
$$= \exp\Big(-\Theta(n^3p^3)\Big)$$

- *Disjoint bound* better than *clustered bound* when $p \leq \frac{\log n}{n}$.
- In general: counting disjoint copies of some *subgraph* gives

$$\mathbb{P}\left(X_H \geq 2\mu_H\right) \geq \exp\Big(-\Theta(\min_{G\subseteq H}\mu_G)\Big)$$

DeMarco–Kahn Upper Tail Conjecture

Conjecture (DeMarco–Kahn 2011)

$$-\log \mathbb{P}\left(\mathbf{X_H} \geq \mathbf{2}\mu_{\mathbf{H}}\right) \asymp \min\left\{\min_{\mathbf{G}\subseteq\mathbf{H}} \mu_{\mathbf{G}},\ \mathbf{M}^*_{\mathbf{H}} \log(\mathbf{1}/\mathbf{p})\right\}$$

- Sparse mechanism: many *disjoint* copies of some $G \subseteq H$
- Dense mechanism: many *clustered* copies of some $G \subseteq H$

- Consistent with key principle of Large deviation theory: "Deviation happens in most likely of all unlikely ways"

DeMarco–Kahn Upper Tail Conjecture: known cases

Conjecture (DeMarco–Kahn 2011)

$$-\log \mathbb{P}\left(X_H \geq 2\mu_H\right) \asymp \min\Big\{\min_{G \subseteq H} \mu_G,\ M_H^* \log(1/p)\Big\}$$

- $H = K_k$ (DeMarco–Kahn 2012)
- $H = K_{1,k}$ (Šileikis–Warnke 2019+)
- $H = C_k$ (Raz 2019+)
- Numerous *tremendous* results for $p \geq n^{-\alpha_H}$ (Chatterjee–Dembo, Lubetzky-Zhao, Dembo–Cook, ...)

Disproof of DeMarco–Kahn Upper Tail Conjecture

Conjecture (DeMarco–Kahn 2012)

For every H there is $H_0 \subseteq H$ such that

$$-\log \mathbb{P}\left(X_H \geq 2\mu_H\right) \asymp \min\{\mu_{H_0},\ M_H^* \log(1/p)\}$$

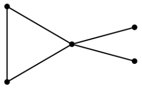
Theorem (Šileikis–Warnke 2019+)

For infinitely many connected H there is $c = c(H) \in (0,1)$ s.t.

$$-\log \mathbb{P}\left(X_H \geq 2\mu_H\right) \leq C \min\{(\mu_{H_0})^c \log \mu_{H_0},\ M_H^* \log(1/p)\}$$

- DeMarco–Kahn conjecture fails when $\mu_{H_0} \to \infty$ slowly

The smallest counterexample (that we have)



- Close to the threshold the DeMarco–Kahn conjecture says

$$-\log \mathbb{P}\left(X_H \geq 2\mu_H\right) \asymp \mu_{K_3} \asymp (np)^3$$

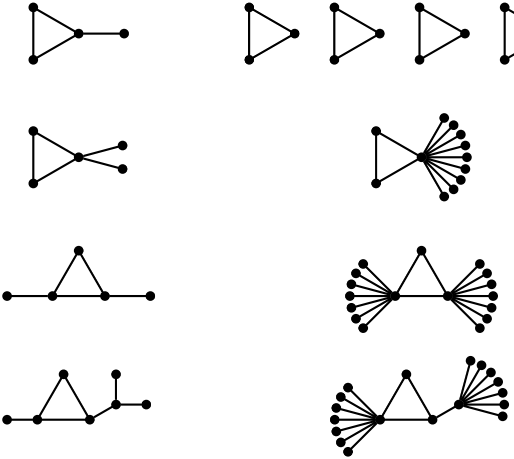
(i) Expose edges on small subset of vertices, choose some vertex v in some triangle

(ii) Expose remaining edges; suppose v has $> (np)^{5/2}$ neighbours

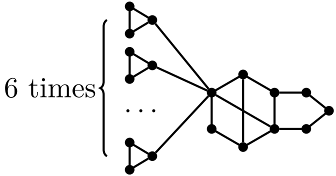
- This gives $> \binom{(np)^{5/2}}{2} \geq 2\mu_H$ copies of H
- Since $(np)^{5/2} \gg np$, Stirlings' formula implies:

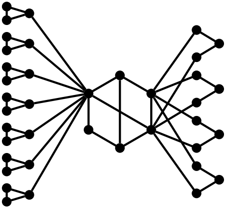
$$\mathbb{P}\left(\mathrm{Bin}(n,p) \geq (np)^{5/2}\right) \geq \exp\left(-\Theta\left((np)^{5/2}\log(np)\right)\right)$$

DeMarco–Kahn Conjecture: zoo of counterexamples

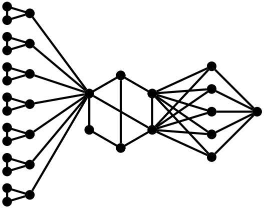


DeMarco–Kahn Conjecture: tricky counterexamples





p just above threshold



Slightly higher p