A fourth moment theorem for estimating subgraph counts in large graphs

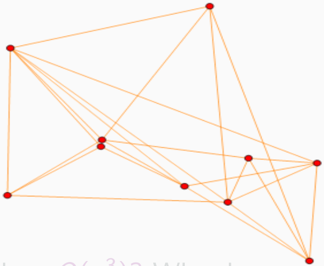
Dan Mikulincer

UW

Joint work with Nitya Mani

Subgraph Counting

Task: Given a large graph, count (estimate) number of triangles.

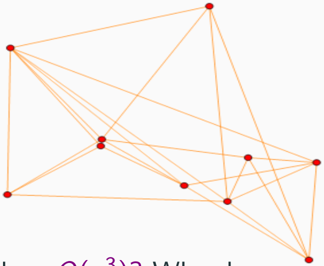


Can we do better than $O(n^3)$? Why do we care?

- Basic statistical task.
- Can carry latent geometric information about embeddings.
- Maybe you like triangles...

Subgraph Counting

Task: Given a large graph, count (estimate) number of triangles.

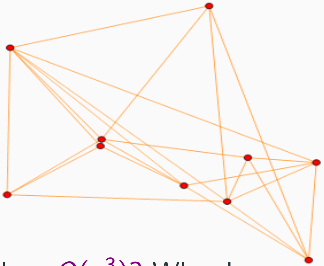


Can we do better than $O(n^3)$? Why do we care?

- Basic statistical task.
- Can carry latent geometric information about embeddings.
- Maybe you like triangles...

Subgraph Counting

Task: Given a large graph, count (estimate) number of triangles.

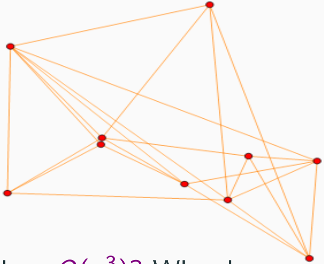


Can we do better than $O(n^3)$? Why do we care?

- Basic statistical task.
- Can carry latent geometric information about embeddings.
- Maybe you like triangles...

Subgraph Counting

Task: Given a large graph, count (estimate) number of triangles.

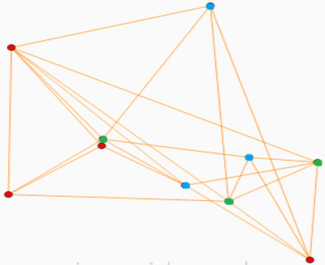


Can we do better than $O(n^3)$? Why do we care?

- Basic statistical task.
- Can carry latent geometric information about embeddings.
- Maybe you like triangles...

Subgraph Counting - Subsampling

Idea: Subsample the vertices and count monochromatic triangles.

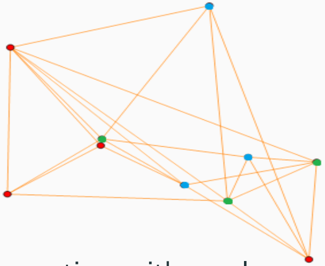


Algorithm: color the vertices with c colors.

- Count the number of number of blue triangles.
- In expectation should be proportional to the total count.
- Is this consistent? How effective?

Subgraph Counting - Subsampling

Idea: Subsample the vertices and count monochromatic triangles.

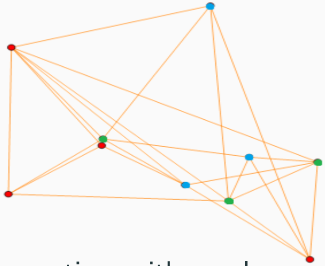


Algorithm: color the vertices with c colors.

- Count the number of number of blue triangles.
- In expectation should be proportional to the total count.
- Is this consistent? How effective?

Subgraph Counting - Subsampling

Idea: Subsample the vertices and count monochromatic triangles.



Algorithm: color the vertices with c colors.

- Count the number of number of blue triangles.
- In expectation should be proportional to the total count.
- Is this consistent? How effective?

Remarks About Consistency

Consistency: For a sequence of graphs G_n and their (random) monochromatic triangle counts $T(G_n)$, we say that $T(G_n)$ is a consistent estimator if, for some deterministic sequence a_n,

$$a_n T(G_n) \to \#\text{triangles}(G_n).$$

Theorem (Bhattacharya, Das, Mukherjee 20')

Under fairly general conditions on G_n, $T(G_n)$ is consistent.

Remarks About Consistency

Consistency: For a sequence of graphs G_n and their (random) monochromatic triangle counts $T(G_n)$, we say that $T(G_n)$ is a consistent estimator if, for some deterministic sequence a_n,

$$a_n T(G_n) \to \#\text{triangles}(G_n).$$

Theorem (Bhattacharya, Das, Mukherjee 20')

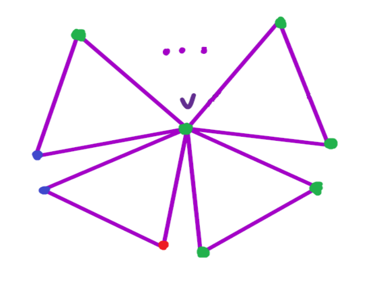
Under fairly general conditions on G_n, $T(G_n)$ is consistent.

Efficiency

For the efficiency of $T(G_n)$ we also need to bound the error.

Question

What is the asymptotic distribution of (a normalized) $T(G_n)$?



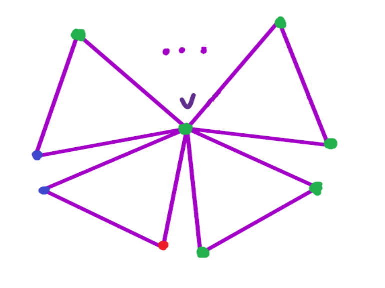
Sum of independent triangles triangles $\implies$ $T(G_n)$ is normal.

Efficiency

For the efficiency of $T(G_n)$ we also need to bound the error.

Question

What is the asymptotic distribution of (a normalized) $T(G_n)$?



Sum of independent triangles triangles $\implies$ $T(G_n)$ is normal.

Formal Setting

- $G_n = (V_n, E_n)$ graphs with $|V_n| \to \infty$, H fixed subgraph.
- Randomly color V_n with c colors and set

$$T(H, G_n) = \#\text{Monochromatic copies of } H.$$

- Normalize

$$Z(H, G_n) = \frac{T(H, G_n) - \mathbb{E}[T(H, G_n)]}{\sqrt{\mathrm{Var}(T(H, G_n))}}.$$

Question

Find necessary and sufficient conditions on G_n such that $Z(H, G_n) \to N(0, 1)$.

Formal Setting

- $G_n = (V_n, E_n)$ graphs with $|V_n| \to \infty$, H fixed subgraph.
- Randomly color V_n with c colors and set

$$T(H, G_n) = \#\text{Monochromatic copies of } H.$$

- Normalize

$$Z(H, G_n) = \frac{T(H, G_n) - \mathbb{E}\left[T(H, G_n)\right]}{\sqrt{\mathrm{Var}(T(H, G_n))}}.$$

Question

Find necessary and sufficient conditions on G_n such that $Z(H, G_n) \to N(0, 1)$.

Previous Results - Edges

Theorem (Bhattacharya, Diaconis, Mukherjee 17')

$Z(K_2, G_n) \to N(0,1)$ iff G_n does not have many 4-cycles.

Why 4-cycles?

$$\mathbb{E}\left[T(K_2, G_n)^4\right] = 3 \cdot \#\text{wedges} + \#\text{four-cycles}.$$

Not having many 4-cycles implies,

$$\mathbb{E}\left[Z(K_2, G_n)^4\right] \to 3 = \mathbb{E}\left[N(0,1)^4\right].$$

Previous Results - Edges

Theorem (Bhattacharya, Diaconis, Mukherjee 17')

$Z(K_2, G_n) \to N(0,1)$ iff G_n does not have many 4-cycles.

Why 4-cycles?

$$\mathbb{E}\left[T(K_2, G_n)^4\right] = 3 \cdot \#\text{wedges} + \#\text{four-cycles}.$$

Not having many 4-cycles implies,

$$\mathbb{E}\left[Z(K_2, G_n)^4\right] \to 3 = \mathbb{E}\left[N(0,1)^4\right].$$

Previous Results - Edges

Theorem (Bhattacharya, Diaconis, Mukherjee 17')

$Z(K_2, G_n) \to N(0,1)$ iff G_n does not have many 4-cycles.

Why 4-cycles?

$$\mathbb{E}\left[T(K_2, G_n)^4\right] = 3 \cdot \#\text{wedges} + \#\text{four-cycles}.$$

Not having many 4-cycles implies,

$$\mathbb{E}\left[Z(K_2, G_n)^4\right] \to 3 = \mathbb{E}\left[N(0,1)^4\right].$$

Previous Results - Edges

Theorem (Bhattacharya, Diaconis, Mukherjee 17')

$Z(K_2, G_n) \to N(0,1)$ iff G_n does not have many 4-cycles.

Why 4-cycles?

$$\mathbb{E}\left[T(K_2, G_n)^4\right] = 3 \cdot \#\text{wedges} + \#\text{four-cycles}.$$

Not having many 4-cycles implies,

$$\mathbb{E}\left[Z(K_2, G_n)^4\right] \to 3 = \mathbb{E}\left[N(0,1)^4\right].$$

Previous Results - General

Theorem (Bhattacharya, Fang, Yan 20')

Suppose that $c \geq 5$, $Z(K_3, G_n) \to N(0,1)$ *iff*

$$\mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3.$$

Theorem (Das, Himwich, Mani 23')

Suppose that $c \geq 32$, *then for any* H, $Z(H, G_n) \to N(0,1)$ *iff*

$$\mathbb{E}\left[Z(H, G_n)^4\right] \to 3.$$

Previous Results - General

Theorem (Bhattacharya, Fang, Yan 20')

Suppose that $c \geq 5$, $Z(K_3, G_n) \to N(0,1)$ *iff*

$$\mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3.$$

Theorem (Das, Himwich, Mani 23')

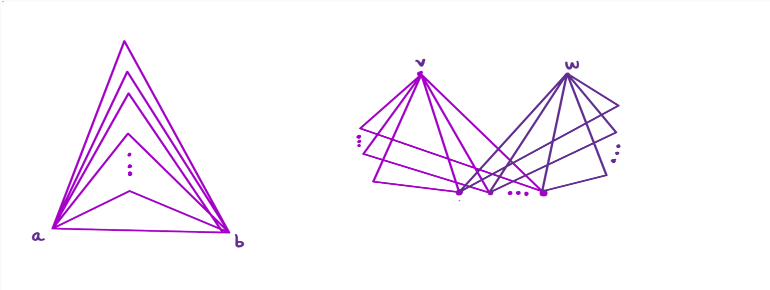
Suppose that $c \geq 32$*, then for any* H, $Z(H, G_n) \to N(0,1)$ *iff*

$$\mathbb{E}\left[Z(H, G_n)^4\right] \to 3.$$

Small Number of Colors

In fact, when $c = 2, 3, 4$ there exists graphs G_n such that

$$\mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3 \text{ and } Z(K_3, G_n) \nrightarrow N(0, 1).$$



Small Number of Colors

For the rest of the talk we focus on $c = 2$ and $H = K_3$.

Goal: Necessary and sufficient conditions for $Z(K_3, G_n) \to N(0,1)$.

A useful representation:

Let $X := \{X_v\}_{v \in V(G_n)}$ be *i.i.d.* Radamacher($\frac{1}{2}$), then

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w .$$

Question

Let $P : \mathbb{R}^{|V|} \to \mathbb{R}$ be a (quadratic, homogeneous) polynomial.
When is $P(X)$ close to $N(0,1)$?

Small Number of Colors

For the rest of the talk we focus on $c = 2$ and $H = K_3$.

Goal: Necessary and sufficient conditions for $Z(K_3, G_n) \to N(0,1)$.

A useful representation:

Let $X := \{X_v\}_{v \in V(G_n)}$ be *i.i.d.* Radamacher($\frac{1}{2}$), then

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w.$$

Question

Let $P : \mathbb{R}^{|V|} \to \mathbb{R}$ be a (quadratic, homogeneous) polynomial.
When is $P(X)$ close to $N(0,1)$?

Small Number of Colors

For the rest of the talk we focus on $c = 2$ and $H = K_3$.

Goal: Necessary and sufficient conditions for $Z(K_3, G_n) \to N(0,1)$.

A useful representation:

Let $X := \{X_v\}_{v \in V(G_n)}$ be *i.i.d.* Radamacher($\frac{1}{2}$), then

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w.$$

Question

Let $P : \mathbb{R}^{|V|} \to \mathbb{R}$ be a (quadratic, homogeneous) polynomial.
When is $P(X)$ close to $N(0,1)$?

Two Deep Results - Invariance Principle

Theorem (Invariance Principle – Mossel, O'Donnell, Oleszkiewicz 10')

Let $P : \mathbb{R}^n \to \mathbb{R}$ *be a low-degree multi-linear polynomial, with* **low influences** *then for* $G \sim N(0, \mathrm{I}_n)$ *and* $X \sim \mathrm{Rademacher}^{\otimes n}$,

$$P(X) \simeq P(G).$$

Influences: The 'influence' of variable i is

$$\mathrm{Influence}_i(P) := \mathbb{E}\left[(\partial_i P(X))^2\right].$$

In other words, "$\mathrm{Influence}_i(P) = \#$ of monomials containing x_i".

Two Deep Results - Invariance Principle

Theorem (Invariance Principle – Mossel, O'Donnell, Oleszkiewicz 10')

Let $P : \mathbb{R}^n \to \mathbb{R}$ *be a low-degree multi-linear polynomial, with* **low influences** *then for* $G \sim N(0, \mathrm{I}_n)$ *and* $X \sim \mathrm{Rademacher}^{\otimes n}$,

$$P(X) \simeq P(G).$$

Influences: The 'influence' of variable i is

$$\mathrm{Influence}_i(P) := \mathbb{E}\left[(\partial_i P(X))^2\right].$$

In other words, "$\mathrm{Influence}_i(P) = \#$ of monomials containing x_i".

Two Deep Results - The Fourth Moment Theorem

Theorem (Fourth Moment Theorem – Nualart and Peccati 05')

Let $P : \mathbb{R}^n \to \mathbb{R}$ *be a low-degree, multi-linear, homogeneous, polynomial then for* $G \sim N(0, \mathrm{I}_n)$*, with* $\mathbb{E}\left[P(G)^2\right] = 1$*.*

$$\mathbb{E}\left[P(G)^4\right] \simeq 3 \implies P(G) \simeq N(0,1).$$

For a polynomial with Gaussian variables a **necessary and sufficient** condition for normal approximation is convergence of the 4th moment.

Two Deep Results - The Fourth Moment Theorem

Theorem (Fourth Moment Theorem – Nualart and Peccati 05')

Let $P : \mathbb{R}^n \to \mathbb{R}$ *be a low-degree, multi-linear, homogeneous, polynomial then for* $G \sim N(0, \mathrm{I}_n)$, *with* $\mathbb{E}\left[P(G)^2\right] = 1$.

$$\mathbb{E}\left[P(G)^4\right] \simeq 3 \implies P(G) \simeq N(0,1).$$

For a polynomial with Gaussian variables a **necessary and sufficient** condition for normal approximation is convergence of the 4th moment.

Monochromatic Triangles - A First Sufficient Condition

Recall:

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w.$$

So, $\mathrm{Influence}_v(Z(K_3, G_n)) \propto \#$ of triangles containing v.

Vertex v is ' influential' if it appears in many triangles.

Invariance Principle + Fourth Moment Theorem imply:

$$\text{No influential vertices} + \mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3$$
$$\implies Z(K_3, G_n) \to N(0, 1).$$

Monochromatic Triangles - A First Sufficient Condition

Recall:

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w.$$

So, $\mathrm{Influence}_v(Z(K_3, G_n)) \propto \#$ of triangles containing v.

Vertex v is ' influential' if it appears in many triangles.

Invariance Principle + Fourth Moment Theorem imply:

No influential vertices $+ \mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3$

$$\implies Z(K_3, G_n) \to N(0, 1).$$

Monochromatic Triangles - A First Sufficient Condition

Recall:

$$Z(K_3, G_n) \propto \sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w.$$

So, $\mathrm{Influence}_v(Z(K_3, G_n)) \propto \#$ of triangles containing v.

Vertex v is ' influential' if it appears in many triangles.

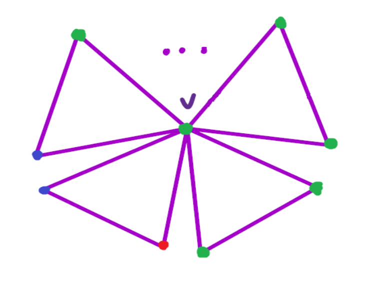
Invariance Principle + Fourth Moment Theorem imply:

No influential vertices $+ \, \mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3$

$$\implies Z(K_3, G_n) \to N(0,1).$$

Monochromatic Triangles - A First Sufficient Condition

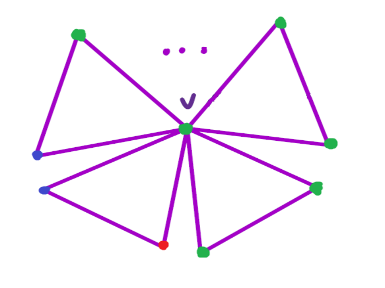
This condition is not necessary.



v is clearly an influential vertex. Already seen CLT for this graph.

Monochromatic Triangles - A First Sufficient Condition

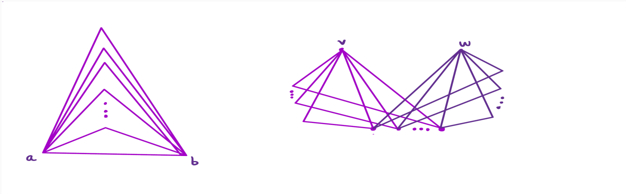
This condition is not necessary.



v is clearly an influential vertex. Already seen CLT for this graph.

Influential Edges

Recall the bad example.



$\{a, b\}$ is an **influential edge**, appear together in many triangles.

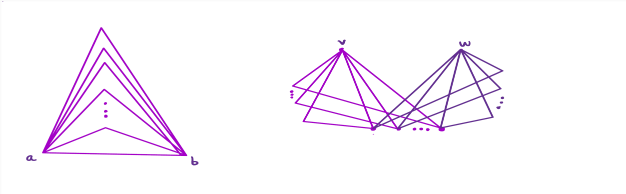
Theorem (Mani, M. 24')

Suppose that G_n has no influential edges and that $\mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3$. Then,

$$Z(K_3, G_n) \to N(0, 1).$$

Influential Edges

Recall the bad example.



$\{a, b\}$ is an **influential edge**, appear together in many triangles.

Theorem (Mani, M. 24')

Suppose that G_n has no influential edges and that $\mathbb{E}\left[Z(K_3, G_n)^4\right] \to 3$. Then,

$$Z(K_3, G_n) \to N(0, 1).$$

Influential Edges

The no influential edge condition is in fact necessary.

Theorem (Mani, M. 24')

Suppose that G_n contains an influential edge. Then

$$Z(K_3, G_n) \nrightarrow N(0,1).$$

Influential edges completely characterize the fourth moment theorem for $Z(K_3, G_n)$.

Influential Edges

The no influential edge condition is in fact necessary.

Theorem (Mani, M. 24')

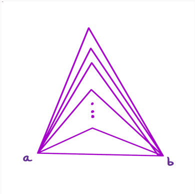
Suppose that G_n contains an influential edge. Then

$$Z(K_3, G_n) \nrightarrow N(0,1).$$

Influential edges completely characterize the fourth moment theorem for $Z(K_3, G_n)$.

Proof Sketch

Suppose that $\{a, b\}$ is an influential edge.



In the above picture,

$$P(X)|X_a \neq X_b \sim \text{constant.}$$

A bit less immediate is that, by the LLN,

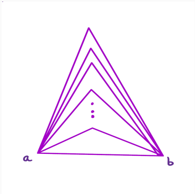
$$P(X)|X_a = X_b \to \text{constant.}$$

So, in general

$$P(X) \simeq \text{Bernoulli} + \tilde{P}(X) \neq^{???} N(0,1).$$

Proof Sketch

Suppose that $\{a, b\}$ is an influential edge.



In the above picture,

$$P(X)|X_a \neq X_b \sim \text{constant.}$$

A bit less immediate is that, by the LLN,

$$P(X)|X_a = X_b \to \text{constant.}$$

So, in general

$$P(X) \simeq \text{Bernoulli} + \tilde{P}(X) \neq^{???} N(0, 1).$$

Proof Sketch

Suppose $\{a, b\}$ is the only influential edge and let

$$M := \#\{\text{triangles containing } \{a, b\}\}.$$

By the Invariance Principle

$$\sum_{\substack{\{u,v,w\} \\ \text{triangle}}} X_v X_u + X_v X_w + X_u X_w = M X_a X_b + P(X) \quad \simeq M X_a X_b + P(G).$$

and

$$Z(K_3, G_n) \simeq \frac{M X_a X_b + P(G)}{\sigma} = \text{Bernoulli} + \tilde{P}(G),$$

with $\tilde{P}$ quadratic and $\mathrm{Var}(\tilde{P}(G)) \leq 1$.

Proof Sketch

$$Z(K_3, G_n) = \text{Bernoulli} + \tilde{P}(G), \tilde{P} \text{ quadratic.}$$

Only two possibilities for the tails of $\tilde{P}(G)$.

Case I: $\tilde{P}(G)$ has strictly exponential tails. In that case, $\text{Bernoulli} + \tilde{P}(G) \neq N(0,1)$.

Case II: $\tilde{P}(G)$ has better than exponential tails. In that case, $\tilde{P}(G)$ is actually sub-Gaussian. Moreover,

$$\tilde{P}(G) \stackrel{\text{law}}{=} \sum \lambda_i (G_i^2 - 1),$$

with $\max_i \lambda_i = o(1)$. So, by the CLT

$$\text{Bernoulli} + \tilde{P}(G) = \text{Gaussian Mixture} \neq N(0,1).$$

Proof Sketch

$$Z(K_3, G_n) = \mathrm{Bernoulli} + \tilde{P}(G), \tilde{P} \text{ quadratic.}$$

Only two possibilities for the tails of $\tilde{P}(G)$.

Case I: $\tilde{P}(G)$ has strictly exponential tails. In that case, $\mathrm{Bernoulli} + \tilde{P}(G) \neq N(0,1)$.

Case II: $\tilde{P}(G)$ has better than exponential tails. In that case, $\tilde{P}(G)$ is actually sub-Gaussian. Moreover,

$$\tilde{P}(G) \overset{\text{law}}{=} \sum \lambda_i (G_i^2 - 1),$$

with $\max_i \lambda_i = o(1)$. So, by the CLT

$$\mathrm{Bernoulli} + \tilde{P}(G) = \text{Gaussian Mixture} \neq N(0,1).$$

Proof Sketch

$$Z(K_3, G_n) = \text{Bernoulli} + \tilde{P}(G), \tilde{P} \text{ quadratic.}$$

Only two possibilities for the tails of $\tilde{P}(G)$.

Case I: $\tilde{P}(G)$ has strictly exponential tails. In that case, $\text{Bernoulli} + \tilde{P}(G) \neq N(0,1)$.

Case II: $\tilde{P}(G)$ has better than exponential tails. In that case, $\tilde{P}(G)$ is actually sub-Gaussian. Moreover,

$$\tilde{P}(G) \stackrel{\text{law}}{=} \sum \lambda_i (G_i^2 - 1),$$

with $\max_i \lambda_i = o(1)$. So, by the CLT

$$\text{Bernoulli} + \tilde{P}(G) = \text{Gaussian Mixture} \neq N(0,1).$$

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Upper Bounds:

- $Z(H, G_n)$ is a polynomial of degree at most r.
- $Z(H, G_n)$ is not homogeneous and the fourth moment theorem does not apply in general.
- Can still prove a a fourth moment theorem for polynomials of the form $Z(H, G_n)$.
- Same upper bounds involving influential edges hold.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Upper Bounds:

- $Z(H, G_n)$ is a polynomial of degree at most r.
- $Z(H, G_n)$ is not homogeneous and the fourth moment theorem does not apply in general.
- Can still prove a a fourth moment theorem for polynomials of the form $Z(H, G_n)$.
- Same upper bounds involving influential edges hold.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Upper Bounds:

- $Z(H, G_n)$ is a polynomial of degree at most r.
- $Z(H, G_n)$ is not homogeneous and the fourth moment theorem does not apply in general.
- Can still prove a a fourth moment theorem for polynomials of the form $Z(H, G_n)$.
- Same upper bounds involving influential edges hold.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Upper Bounds:

- $Z(H, G_n)$ is a polynomial of degree at most r.
- $Z(H, G_n)$ is not homogeneous and the fourth moment theorem does not apply in general.
- Can still prove a a fourth moment theorem for polynomials of the form $Z(H, G_n)$.
- Same upper bounds involving influential edges hold.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Upper Bounds:

- $Z(H, G_n)$ is a polynomial of degree at most r.
- $Z(H, G_n)$ is not homogeneous and the fourth moment theorem does not apply in general.
- Can still prove a a fourth moment theorem for polynomials of the form $Z(H, G_n)$.
- Same upper bounds involving influential edges hold.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Lower Bounds:

- Can separate the influential vertices.
- Leads to a mixture with a degree r Gaussian form.

$$Z(H, G_n) \simeq \mathrm{Bernoulli} P_{r-2}(G) + \cdots + P_r(G).$$

- There is a taxonomy of different tail behaviors for $P_r(G)$, as $e^{-t^{2\ell}}$ for $\ell = 1, \frac{1}{2}, \ldots, \frac{1}{r}$.
- Could not find a dichotomy with the central limit theorem... Lower bound only applies in certain special cases.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Lower Bounds:

- Can separate the influential vertices.
- Leads to a mixture with a degree r Gaussian form.

$$Z(H, G_n) \simeq \mathrm{Bernoulli} P_{r-2}(G) + \cdots + P_r(G).$$

- There is a taxonomy of different tail behaviors for $P_r(G)$, as $e^{-t^{2\ell}}$ for $\ell = 1, \frac{1}{2}, \ldots, \frac{1}{r}$.
- Could not find a dichotomy with the central limit theorem... Lower bound only applies in certain special cases.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Lower Bounds:

- Can separate the influential vertices.
- Leads to a mixture with a degree r Gaussian form.

$$Z(H, G_n) \simeq \mathrm{Bernoulli} P_{r-2}(G) + \cdots + P_r(G).$$

- There is a taxonomy of different tail behaviors for $P_r(G)$, as $e^{-t^{2\ell}}$ for $\ell = 1, \frac{1}{2}, \ldots, \frac{1}{r}$.
- Could not find a dichotomy with the central limit theorem... Lower bound only applies in certain special cases.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Lower Bounds:

- Can separate the influential vertices.
- Leads to a mixture with a degree r Gaussian form.

$$Z(H, G_n) \simeq \mathrm{Bernoulli} P_{r-2}(G) + \cdots + P_r(G).$$

- There is a taxonomy of different tail behaviors for $P_r(G)$, as $e^{-t^{2\ell}}$ for $\ell = 1, \frac{1}{2}, \ldots, \frac{1}{r}$.
- Could not find a dichotomy with the central limit theorem... Lower bound only applies in certain special cases.

Different Sub-Graphs

What about *non-triangle* monochromatic counts? Suppose $|H| = r$.

Lower Bounds:

- Can separate the influential vertices.
- Leads to a mixture with a degree r Gaussian form.

$$Z(H, G_n) \simeq \mathrm{Bernoulli} P_{r-2}(G) + \cdots + P_r(G).$$

- There is a taxonomy of different tail behaviors for $P_r(G)$, as $e^{-t^{2\ell}}$ for $\ell = 1, \frac{1}{2}, \ldots, \frac{1}{r}$.
- Could not find a dichotomy with the central limit theorem... Lower bound only applies in certain special cases.

Further Questions

Question

Is it true that when G_n contains an influential edge, then $Z(H, G_n)$ is non-Gaussian?

In particular, if a polynomial in Gaussian variables has approximately sub-Gaussian tails, is it approximately Gaussian?

Question

Let $c < c'$. Suppose that a CLT holds for $Z_c(H, G_n)$. Does this imply that a CLT holds for $Z_{c'}(H, G_n)$ as well?

Further Questions

Question

Is it true that when G_n contains an influential edge, then $Z(H, G_n)$ is non-Gaussian?

In particular, if a polynomial in Gaussian variables has approximately sub-Gaussian tails, is it approximately Gaussian?

Question

Let $c < c'$. Suppose that a CLT holds for $Z_c(H, G_n)$. Does this imply that a CLT holds for $Z_{c'}(H, G_n)$ as well?

Thank You