Stochastic Block Model in a new critical regime and the Interacting Multiplicative Coalescent

Vitalii Konarovskyi^{*†}

*Leipzig University

[†]Institute of Mathematics of NAS of Ukraine

Bernoulli-IMS One World Symposium 2020

joint work with Vlada Limic





900

Stochastic Block Model

Stochastic Block Model G(n, p, q) is a random graph such that:

- consists of nm vertices divided into m subsets (m = 2);
- edges are drown independently;
- intra class edges appear with probability $p = p_n$;
- inter class edges appear with probability $q = q_n$.



We are interesting in the scaling limit as $n \to \infty$ and $p_n, q_n \to 0$.

◆□ ▶ ◆□ ▶ ◆三 ▶ ◆□ ▶ ◆□ ▶ ◆○ ◆

Stochastic Block Model

Stochastic Block Model G(n, p, q) is a random graph such that:

- consists of nm vertices divided into m subsets (m = 2);
- edges are drown independently;
- intra class edges appear with probability $p = p_n$;
- inter class edges appear with probability $q = q_n$.



We are interesting in the scaling limit as $n \to \infty$ and $p_n, q_n \to 0$.

◆□ ▶ ◆□ ▶ ◆三 ▶ ◆□ ▶ ◆□ ▶ ◆○ ◆

 $C_1(n)$ – the largest component of the SBM

It is well-known:

If p_n = q_n = a/m, then SBM is an Erdős-Rényi graph for which:
for a > 1, C₁(n) ~ Θ(n);
for a < 1, C₁(n) ~ Θ(ln n); (Erdős, Rényi '60, '61)
for a = 1, C₁(n) ~ Θ(n^{2/b}).

9 If p_n = a/mn, q_n = b/mn, then

a + (m - 1)b > m, C₁(n) ~ Θ(n);
a + (m - 1)b ≤ m, C₁(n) ~ o(n). (Bollobás, Janson, Riordan '07)

We are interesting in a novel critical regime: $q_n \ll p_n \sim \frac{a}{n}$.

▲ロト ▲御ト ▲ヨト ▲ヨト 三国 - のへで

 $C_1(n)$ – the largest component of the SBM

It is well-known:

• If
$$p_n = q_n = \frac{a}{mn}$$
, then SBM is an Erdős-Rényi graph for which:
• for $a > 1$, $C_1(n) \sim \Theta(n)$;
• for $a < 1$, $C_1(n) \sim \Theta(\ln n)$; (Erdős, Rényi '60, '61)
• for $a = 1$, $C_1(n) \sim \Theta(n^{2/3})$.
• If $p_n = \frac{a}{mn}$, $q_n = \frac{b}{mn}$, then
• $a + (m-1)b > m$, $C_1(n) \sim \Theta(n)$;
• $a + (m-1)b \le m$, $C_1(n) \sim O(n)$. (Bollobás, Janson, Riordan '07)

We are interesting in a novel critical regime: $q_n \ll p_n \sim rac{a}{n}$.

 $C_1(n)$ – the largest component of the SBM

It is well-known:

◆□▶ ◆舂▶ ◆臣▶ ◆臣▶ 三臣.

900

We are interesting in a novel critical regime: $q_n \ll p_n \sim rac{a}{n}$.

 $C_1(n)$ – the largest component of the SBM

It is well-known:

• If
$$p_n = q_n = \frac{a}{mn}$$
, then SBM is an Erdős-Rényi graph for which:
• for $a > 1$, $C_1(n) \sim \Theta(n)$;
• for $a < 1$, $C_1(n) \sim \Theta(\ln n)$; (Erdős, Rényi '60, '61)
• for $a = 1$, $C_1(n) \sim \Theta(n^{2/3})$.
• If $p_n = \frac{a}{mn}$, $q_n = \frac{b}{mn}$, then
• $a + (m-1)b > m$, $C_1(n) \sim \Theta(n)$;
• $a + (m-1)b \le m$, $C_1(n) \sim o(n)$. (Bollobás, Janson, Riordan '07)

<ロト < 団 > < 臣 > < 臣 > 三 の < で</p>

We are interesting in a novel critical regime: $q_n \ll p_n \sim \frac{a}{n}$.

Scaling limit of Erdős-Rényi Graphs

G(n,p) – a Erdős-Rényi random graph with n vertices and edges appearing with prob. $p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}$.

 $\operatorname{ER}_n(t) := \frac{1}{n^{2/3}}(C_1(n,t), C_2(n,t), \dots, C_k(n,t), 0, 0, \dots, n),$ where $C_k(n,t)$ is the k-th largest component.

Theorem.

For every $t \in \mathbb{R}$ the sequence $\text{ER}_n(t)$ converges in l_{\downarrow}^2 to $X^*(t)$ in distribution, where $X^*(t)$ is the ordered sequence of excursions of

$$W(s) - \frac{1}{2}s^2 + ts, \quad s \ge 0,$$

above past minima. $X^*(t)$, $t \in \mathbb{R}$, is called the *standard Multiplicative coalescent*, and is a Markov process in l^2_{\perp} .

Scaling limit of Erdős-Rényi Graphs

G(n,p) – a Erdős-Rényi random graph with n vertices and edges appearing with prob. $p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}$.

 $\operatorname{ER}_{n}(t) := \frac{1}{n^{2/3}}(C_{1}(n,t), C_{2}(n,t), \dots, C_{k}(n,t), 0, 0, \dots),$ where $C_{k}(n,t)$ is the k-th largest component.

Theorem.

For every $t \in \mathbb{R}$ the sequence $\text{ER}_n(t)$ converges in l_{\downarrow}^2 to $X^*(t)$ in distribution, where $X^*(t)$ is the ordered sequence of excursions of

$$W(s) - \frac{1}{2}s^2 + ts, \quad s \ge 0,$$

above past minima. $X^*(t)$, $t \in \mathbb{R}$, is called the *standard Multiplicative coalescent*, and is a Markov process in l^2_{\perp} .

Scaling limit of Erdős-Rényi Graphs

G(n,p) – a Erdős-Rényi random graph with n vertices and edges appearing with prob. $p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}$.

 $\operatorname{ER}_{n}(t) := \frac{1}{n^{2/3}}(C_{1}(n,t), C_{2}(n,t), \dots, C_{k}(n,t), 0, 0, \dots),$ where $C_{k}(n,t)$ is the k-th largest component.

Theorem. (Aldous '97, Anmerdariz '01, Limic '98,'19)

For every $t \in \mathbb{R}$ the sequence $\operatorname{ER}_n(t)$ converges in l_{\downarrow}^2 to $X^*(t)$ in distribution, where $X^*(t)$ is the ordered sequence of excursions of

$$W(s) - \frac{1}{2}s^2 + ts, \quad s \ge 0,$$

above past minima. $X^*(t)$, $t \in \mathbb{R}$, is called the *standard Multiplicative coalescent*, and is a Markov process in l^2_{\perp} .

Interacting Multiplicative Coalescent and Main Result

$$SBM_n(t,s) := \frac{1}{n^{2/3}} (C_1(n,t), C_2(n,t), \dots, C_k(n,t), 0, 0), \quad t \in \mathbb{R}, \ s \ge 0,$$

where $C_k(n,t)$ is the k-th largest component of SBM G(n, p, q) with

$$q = q_n(t) = \frac{s}{n^{4/3}}, \quad p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}.$$

For $s \ge 0$ and a fixed family of indep. r.v. $\xi_{i,j} \sim \text{Exp}(1)$, $i, j \ge 1$, define a random map $\text{RMM}_s : l_{\downarrow}^2 \times l_{\downarrow}^2 \to l_{\downarrow}^2$:

- consider coord. of $x, y \in l^2_{\downarrow}$ as a masses of corresponding vertices of a graph;
- ullet for every $i,j\geq 1$ draw an edge between x_i and y_j iff $\xi_{i,j}\leq sx_iy_j$;

• define $\text{RMM}_s(x, y)$ as a vector of the ordered masses of connected components.

Theorem. (K., Limic '2

For every $t \in \mathbb{R}$ and $s \geq 0$ the process $\mathrm{SBM}_n(t,s)$ converges in l^2_{\downarrow} in distribution to $\mathrm{RMM}_s(X^*(t),Y^*(t))$, where X^*,Y^* are independent standard multiplicative coalescents that are independent of ξ

Interacting Multiplicative Coalescent and Main Result

$$SBM_n(t,s) := \frac{1}{n^{2/3}} (C_1(n,t), C_2(n,t), \dots, C_k(n,t), 0, 0), \quad t \in \mathbb{R}, \ s \ge 0,$$

where $C_k(n,t)$ is the k-th largest component of SBM G(n, p, q) with

$$q = q_n(t) = \frac{s}{n^{4/3}}, \quad p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}.$$

For $s \ge 0$ and a fixed family of indep. r.v. $\xi_{i,j} \sim \text{Exp}(1)$, $i, j \ge 1$, define a random map $\text{RMM}_s : l_{\downarrow}^2 \times l_{\downarrow}^2 \to l_{\downarrow}^2$:

- consider coord. of $x, y \in l^2_{\perp}$ as a masses of corresponding vertices of a graph;
- for every $i, j \ge 1$ draw an edge between x_i and y_j iff $\xi_{i,j} \le sx_iy_j$;
- define $\text{RMM}_s(x, y)$ as a vector of the ordered masses of connected components.

Theorem.

For every $t \in \mathbb{R}$ and $s \ge 0$ the process $\mathrm{SBM}_n(t,s)$ converges in l^2_{\downarrow} in distribution to $\mathrm{RMM}_s(X^*(t),Y^*(t))$, where X^*,Y^* are independent standard multiplicative coalescents that are independent of ξ

Interacting Multiplicative Coalescent and Main Result

$$SBM_n(t,s) := \frac{1}{n^{2/3}} (C_1(n,t), C_2(n,t), \dots, C_k(n,t), 0, 0), \quad t \in \mathbb{R}, \ s \ge 0,$$

where $C_k(n,t)$ is the k-th largest component of SBM G(n, p, q) with

$$q = q_n(t) = \frac{s}{n^{4/3}}, \quad p = p_n(t) = \frac{1}{n} + \frac{t}{n^{4/3}}.$$

For $s \ge 0$ and a fixed family of indep. r.v. $\xi_{i,j} \sim \text{Exp}(1)$, $i, j \ge 1$, define a random map $\text{RMM}_s : l_{\downarrow}^2 \times l_{\downarrow}^2 \to l_{\downarrow}^2$:

- consider coord. of $x, y \in l^2_{\perp}$ as a masses of corresponding vertices of a graph;
- for every $i, j \ge 1$ draw an edge between x_i and y_j iff $\xi_{i,j} \le sx_iy_j$;
- define RMM_s(x, y) as a vector of the ordered masses of connected components.

Theorem. (K., Limic '20)

For every $t \in \mathbb{R}$ and $s \geq 0$ the process $\mathrm{SBM}_n(t,s)$ converges in l^2_{\downarrow} in distribution to $\mathrm{RMM}_s(X^*(t),Y^*(t))$, where X^*,Y^* are independent standard multiplicative coalescents that are independent of ξ

Idea of Proof. ER_n and Multiplicative Coalescent

For $x \in l^2_{\perp}$, and independent $\xi_{i,j} \sim \text{Exp}(1)$ define



 $X^{x}(t)$ – ordered masses of connected componnents, $t \geq 0$.

The process $X^x(t)$, $t \ge 0$, is a Markov process in l_{\downarrow}^2 started from x. Moreover, it evolves according to the multiplicative coalescent dynamics:

> each pair of blocks of mass a and b merges at rate abinto a single block of mass a + b

Since for $x=\left(n^{-2/3},\ldots,n^{-2/3},0,0,\ldots
ight)$ and some $t_n o t$, $n o\infty$,

 $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_i x_j\right\}$ $X^x(t_n + n^{1/3}) \stackrel{d}{=} \operatorname{ER}_n(t) \stackrel{d}{\to} X^*(t), \quad n \to \infty.$

| ↓ □ ▶ | 4 厘 ▶ | 4 厘 ▶ | 4 厘 ▶ | 9 へ ()

Idea of Proof. ER_n and Multiplicative Coalescent

For $x \in l^2_{\perp}$, and independent $\xi_{i,j} \sim \text{Exp}(1)$ define



 $X^{x}(t)$ – ordered masses of connected componnents, $t \geq 0$.

The process $X^x(t)$, $t \ge 0$, is a Markov process in l^2_{\downarrow} started from x. Moreover, it evolves according to the multiplicative coalescent dynamics:

> each pair of blocks of mass a and b merges at rate abinto a single block of mass a + b

Since for $x=\left(n^{-2/3},\ldots,n^{-2/3},0,0,\ldots
ight)$ and some $t_n o t$, $n o\infty$

 $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_i x_j\right\}$ $X^x(t_n + n^{1/3}) \stackrel{d}{=} \operatorname{ER}_n(t) \stackrel{d}{\to} X^*(t), \quad n \to \infty.$

Idea of Proof. ER_n and Multiplicative Coalescent

For $x \in l^2_{\perp}$, and independent $\xi_{i,j} \sim \text{Exp}(1)$ define



 $X^{x}(t)$ – ordered masses of connected componnents, $t \geq 0$.

The process $X^x(t)$, $t \ge 0$, is a Markov process in l^2_{\downarrow} started from x. Moreover, it evolves according to the multiplicative coalescent dynamics:

> each pair of blocks of mass a and b merges at rate abinto a single block of mass a + b

Since for $x = \left(n^{-2/3}, \ldots, n^{-2/3}, 0, 0, \ldots\right)$ and some $t_n \to t$, $n \to \infty$,

 $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_ix_j\right\}$ $X^x(t_n + n^{1/3}) \stackrel{d}{=} \operatorname{ER}_n(t) \stackrel{d}{\to} X^*(t), \quad n \to \infty.$

 $x, y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi''_{i,j} \sim \operatorname{Exp}(1), i, j \geq 1$, are independent

$$Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t))$$
 for every $t \ge 0$ and $s \ge 0$.

 $x, y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi''_{i,j} \sim \operatorname{Exp}(1)$, $i, j \geq 1$, are independent



$Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t))$ for every $t \ge 0$ and $s \ge 0$.

◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□▶

 $x,y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi'''_{i,j} \sim \operatorname{Exp}(1)$, $i,j \geq 1$, are independent



$$Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t)) \text{ for every } t \geq 0 \text{ and } s \geq 0.$$

 $x, y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi''_{i,j} \sim \operatorname{Exp}(1)$, $i, j \geq 1$, are independent



 $Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t)) \text{ for every } t \geq 0 \text{ and } s \geq 0.$

 $x, y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi''_{i,j} \sim \operatorname{Exp}(1)$, $i, j \geq 1$, are independent



 $Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t)) \text{ for every } t \geq 0 \text{ and } s \geq 0.$

◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□▶

 $x, y \in l^2_{\downarrow}$ and $\xi'_{i,j}, \xi''_{i,j}, \xi'''_{i,j} \sim \operatorname{Exp}(1)$, $i, j \geq 1$, are independent



$$Z^{x,y}(t,s) \stackrel{d}{=} \operatorname{RMM}_s(X^x(t),Y^y(t))$$
 for every $t \ge 0$ and $s \ge 0$.



▲□▶ ▲□▶ ▲三▶ ▲三▶ ▲□▶ ▲□▶



Let $x = y = (n^{-2/3}, \dots, n^{-2/3}, 0, 0, \dots)$. For some $t_n \to t$, $n \to \infty$, $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_ix_j\right\}$

and for some $s_n \rightarrow s$, $n \rightarrow \infty$,

$$q_n = \frac{s}{n^{4/3}} = 1 - e^{s_n n^{-2/3} n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le s_n x_i y_j\right\}$$

Hence,

$$\text{SBM}_n(t,s) \stackrel{d}{=} Z^{x,y}(t_n + n^{1/3}, s_n) \stackrel{d}{=} \text{RMM}_{s_n}(X^x(t_n + n^{1/3}), Y^y(t_n + n^{1/3})).$$

<ロト <部ト <注下 < 한 - 한



Let $x = y = (n^{-2/3}, \dots, n^{-2/3}, 0, 0, \dots)$. For some $t_n \to t$, $n \to \infty$, $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_ix_j\right\}$

and for some $s_n \to s$, $n \to \infty$,

$$q_n = \frac{s}{n^{4/3}} = 1 - e^{s_n n^{-2/3} n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le s_n x_i y_j\right\}$$

Hence,

$$\text{SBM}_n(t,s) \stackrel{d}{=} Z^{x,y}(t_n + n^{1/3}, s_n) \stackrel{d}{=} \text{RMM}_{s_n}(X^x(t_n + n^{1/3}), Y^y(t_n + n^{1/3})).$$

《曰》 《聞》 《臣》 《臣》 三臣



Let $x = y = (n^{-2/3}, \dots, n^{-2/3}, 0, 0, \dots)$. For some $t_n \to t$, $n \to \infty$, $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_ix_j\right\}$

and for some $s_n \to s$, $n \to \infty$,

$$q_n = \frac{s}{n^{4/3}} = 1 - e^{s_n n^{-2/3} n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le s_n x_i y_j\right\}$$

Hence,

$$\operatorname{SBM}_n(t,s) \stackrel{d}{=} Z^{x,y}(t_n + n^{1/3}, s_n) \stackrel{d}{=} \operatorname{RMM}_{s_n}(X^x(t_n + n^{1/3}), Y^y(t_n + n^{1/3})).$$

<ロト <部ト <注下 < 한 - 한



Let $x = y = (n^{-2/3}, \dots, n^{-2/3}, 0, 0, \dots)$. For some $t_n \to t$, $n \to \infty$, $p_n = \frac{1}{n} + \frac{t}{n^{4/3}} = 1 - e^{(t_n + n^{1/3})n^{-2/3}n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le (t_n + n^{1/3})x_ix_j\right\}$

and for some $s_n \to s$, $n \to \infty$,

$$q_n = \frac{s}{n^{4/3}} = 1 - e^{s_n n^{-2/3} n^{-2/3}} = \mathbb{P}\left\{\xi_{i,j} \le s_n x_i y_j\right\}$$

Hence,

$$\text{SBM}_n(t,s) \stackrel{d}{=} Z^{x,y}(t_n + n^{1/3}, s_n) \stackrel{d}{=} \text{RMM}_{s_n}(X^x(t_n + n^{1/3}), Y^y(t_n + n^{1/3})).$$

<ロト <部ト <注下 < 한 - 한

Idea of proof. Continuity of RMM

The convergence

 $\mathrm{SBM}_n(t,s) \stackrel{d}{=} \mathrm{RMM}_{s_n}(X^x(t_n+n^{1/3}),Y^y(t_n+n^{1/3})) \to \mathrm{RMM}_s(X^*(t),Y^*(t)),$

follows from

$$\left(\operatorname{ER}_n(t) \stackrel{d}{=}\right) \quad X^x(t_n + n^{1/3}) \stackrel{d}{\to} X^*(t), \quad Y^x(t_n + n^{1/3}) \stackrel{d}{\to} Y^*(t), \quad n \to \infty$$

and

Proposition (K., Li

Let $x^n o x$, $y^n o y$ in l^2_{\downarrow} and $s_n o s$ in $[0,\infty)$. Then

 $\operatorname{RMM}_{s_n}(x_n, y_n) \to \operatorname{RMM}_s(x, y)$ in l^2_{\downarrow} in probability

as $n \to \infty$.

Idea of proof. Continuity of RMM

The convergence

 $\mathrm{SBM}_n(t,s) \stackrel{d}{=} \mathrm{RMM}_{s_n}(X^x(t_n+n^{1/3}),Y^y(t_n+n^{1/3})) \to \mathrm{RMM}_s(X^*(t),Y^*(t)),$

follows from

$$\left(\mathbb{ER}_n(t) \stackrel{d}{=}\right) \quad X^x(t_n + n^{1/3}) \stackrel{d}{\to} X^*(t), \quad Y^x(t_n + n^{1/3}) \stackrel{d}{\to} Y^*(t), \quad n \to \infty$$

and

Proposition (K., Limic '20) Let $x^n \to x$, $y^n \to y$ in l^2_{\downarrow} and $s_n \to s$ in $[0, \infty)$. Then $\operatorname{RMM}_{s_n}(x_n, y_n) \to \operatorname{RMM}_s(x, y)$ in l^2_{\downarrow} in probability as $n \to \infty$.

<ロト < 団 > < 臣 > < 臣 > 三 の < で</p>

V. Konarovskyi, V. Limic Stochastic Block Model in a new critical regime and the Interacting Multiplicative Coalescent. *arXiv:2003.10958*

417(17.2000.10500

Thank you for your attention!