Phase transitions of Best-of-two and Best-of-three on stochastic block models

Nobutaka Shimizu1 | Takeharu Shiraga2

1The University of Tokyo, Tokyo, Japan
2Chuo University, Tokyo, Japan

Correspondence
Takeharu Shiraga, Chuo University, Tokyo, Japan.
Email: shiraga.076@g.chuo-u.ac.jp

Abstract
This is concerned with voting processes on graphs where each vertex holds one of two different opinions. In particular, we study the Best-of-two and the Best-of-three. Here at each synchronous round, each vertex updates its opinion to match the majority among the opinions of two random neighbors and itself (the Best-of-two) or the opinions of three random neighbors (the Best-of-three). In this study, we consider the Best-of-two and the Best-of-three on the stochastic block model $G(2n, p, q)$, which is a random graph consisting of two distinct Erdős–Rényi graphs $G(n, p)$ joined by random edges with a density $q ≤ p$. We prove phase transition results for these processes: there is a threshold $r^*$ such that, if $q/p > r^*$ then the process reaches consensus within $O(\log n)$ rounds and the process requires $\exp(\Omega(n))$ rounds if $q/p < r^*$. For the Best-of-two and Best-of-three, the thresholds are $r^* = \sqrt{5} - 2$ and $r^* = 1/7$, respectively.

KEYWORDS
consensus problem, distributed voting, random graph

1 | INTRODUCTION

This study considers voting processes on distributed graphs. Consider an undirected connected graph $G = (V, E)$ where each vertex $v \in V$ initially holds an opinion from a finite set. A voting process is defined by a local updating rule: each vertex updates its opinion according to the rule. Voting processes
appear as simple mathematical models in a wide range of fields, for example, social behaviors, physical phenomena, and biological systems [4, 36, 38]. In distributed computing, voting processes are known as a simple approach to consensus problems [24, 27].

1.1 Previous work

The synchronous pull voting (a.k.a. the voter model) is a simple and well-studied voting process [29, 39]. In the pull voting, at each synchronous and discrete time step, each vertex adopts the opinion of a randomly selected neighbor. Here, the main quantity of interest is the consensus time, which is the number of steps required to reach consensus (i.e., the configuration where all vertices hold the same opinion). Hassin and Peleg [29] showed that the expected consensus time is $O(n^3 \log n)$ for all non-bipartite graphs and for all initial opinion configurations, where $n$ is the number of vertices. Note that, for bipartite graphs, there is an initial opinion configuration that never reaches consensus.

The pull voting has been extended to develop voting processes where each vertex queries multiple neighbors at each step. The simplest multiple-choice voting process is the Best-of-two (two sample voting, or 2-Choices), where each vertex $v \in V$ randomly samples two neighbors (with replacement) and, if both hold the same opinion, adopts it. Doerr et al. [22] showed that, for complete graphs initially involving two possible opinions, the consensus time of the Best-of-two is $O(\log n)$ with high probability. Likewise, the Best-of-three (a.k.a. 3-Majority) is another simple multiple-choice voting process where each vertex adopts the majority opinion among those of three randomly selected neighbors. Several researchers have studied this model on complete graphs initially involving $k \geq 2$ opinions [7, 8, 10, 26]. For example, Ghaffari and Lengler [26] showed that the consensus time of the Best-of-three is $O(k \log n)$ if $k = O(n^{1/3}/\sqrt{\log n})$.

Several studies on multiple-choice voting processes on non-complete graphs have considered expander graphs with an initial bias, that is, the difference between the initial sizes of the largest and the second largest opinions. Cooper et al. [14] showed that, for any regular expander graph initially involving two opinions, the Best-of-two reaches consensus within $O(\log n)$ steps w.h.p. if the initial bias is $\Omega(n \lambda_2)$, where $\lambda_2$ is the second largest eigenvalue of the graph’s transition matrix. This result was later extended to general expander graphs, including Erdős–Rényi random graphs $G(n, p)$, under milder assumptions about the initial bias [15]. Recall that the Erdős–Rényi graph $G(n, p)$ is a graph with $n$ vertices where each vertex pair is joined by an edge with a probability $p$, independent of any other pairs. In [16], the authors studied the Best-of-two and the Best-of-three on regular expander graphs initially involving more than two opinions. In [3, 32], the authors studied multiple-choice voting processes on non-complete graphs with a random initial configuration.

Recently, the Best-of-two on richer classes of graphs involving two opinions have been studied. Previous works demonstrated interesting results that do not hold on complete graphs or expander graphs. Cruciani et al. [18] studied the Best-of-two on the core periphery network, which is a graph consisting of core vertices and periphery vertices. They showed that a phase transition can occur, depending on the density of edges between core and periphery vertices: either the process reaches consensus within $O(\log n)$ steps, or remains a configuration where both opinions coexist for at least $\Omega(n)$ steps. Cruciani et al. [19] studied the Best-of-two on the $(a, b)$-regular stochastic block model, which is a graph consisting of two $a$-regular graphs connected by a $b$-regular bipartite graph. Under certain assumptions including $bla = O(n^{-0.5})$, they showed that, starting from a random initial opinion configuration, the process reaches an almost clustered configuration (e.g., both communities are in almost consensus but...

---

1If a graph initially involves two possible opinions, this definition matches the rule described in Abstract.

2In this study “with high probability” (w.h.p.) means a probability of at least $1 - n^{-c}$ for a constant $c > 0$. 
the opinions are distinct) within $O(\log n)$ steps with constant probability, then stays in that configuration for at least $\Omega(n)$ steps w.h.p. They also proposed a distributed community detection algorithm based on this property.

1.2 | Our results

This study considers the stochastic block model, a well-known random graph model that forms multiple communities. This model has been well explored in a wide range of fields, including biology [11, 37], network analysis [5, 28], and machine learning [1, 2], where it serves as a benchmark for community detection algorithms. The study of the voting processes on the stochastic block model has a potential application in distributed community detection algorithms [6, 9, 19]. In this study, we focus on the following model which admits two communities of equal size.

**Definition 1.1** (Stochastic block model). For $n \in \mathbb{N}$ and $p, q \in [0, 1]$ with $q \leq p$, the stochastic block model $G(2n, p, q)$ is a graph with a vertex set $V = V_1 \cup V_2$, where $|V_1| = |V_2| = n$ and $V_1 \cap V_2 = \emptyset$. In addition, each pair $\{u, v\}$ of distinct vertices $u \in V_i$ and $v \in V_j$ forms an edge with probability $\theta_{ij}$, independent of any other edges, where

$$\theta_{ij} = \begin{cases} p & \text{if } i = j, \\ q & \text{otherwise.} \end{cases}$$

Note that $G(2n, p, q)$ is not connected w.h.p. if $p = o(\log n/n)$ [25]. Throughout this study, we assume $p = o(\log n/n)$, in which regime each community is connected w.h.p.

In this study, we first generate a random graph $G(2n, p, q)$, and then set an initial opinion configuration from $\{1, 2\}$. Let $A^{(0)}, A^{(1)}, \ldots$ be a sequence of random vertex subsets where $A^{(0)}$ is the set of vertices of opinion 1 at step $t$. For any $A \subseteq V$, the consensus time $T_{\text{cons}}(A)$ is defined as

$$T_{\text{cons}}(A) := \min \{t \geq 0 : A^{(t)} \in \{\emptyset, V\}, A^{(0)} = A\}.$$ 

We obtain two main results, described below.

**Result 1: phase transition**.\(^3\) Observe that, if $p = q = 1$, then $G(2n, 1, 1)$ is a complete graph and the consensus time of the Best-of-two is $O(\log n)$, from the results of [22]. On the other hand, the graph $G(2n, 1, 0)$ consists of two disjoint complete graphs, each of size $n$, meaning that, depending on the initial state, it may not reach consensus. This naturally raises the following question: Where is the boundary between these two phenomena? This motivated us to study the consensus times of the Best-of-two and the Best-of-three on $G(2n, p, q)$ for a wide range of $r := q/p$, and led us to propose the following answers.

**Theorem 1.2** (Phase transition of the Best-of-three on $G(2n, p, q)$). Consider the Best-of-three on $G(2n, p, q)$ for two functions $p = p(n)$ and $q = q(n)$ such that $p = o(\log n/n)$, $q \geq \log n/n^2$, and $r = r(n) = q/p$ has a limit $\bar{r} := \lim_{n \to \infty} r(n)$. Then, for any constant $\epsilon > 0$, the following holds.

\(^3\)In the preliminary version [43], we assumed that $r = q/p$ is a constant in Theorems 1.2 and 1.3. In this version, we relax this condition by considering the case of $r = r(n)$ such that there is the limit $\lim_{n \to \infty} r(n)$. The proof is given in Appendix A.
(i) If \( \tilde{r} \geq \frac{1}{2} + \varepsilon \), then \( G(2n, p, q) \) w.h.p. satisfies the following property: There are two positive constants \( C \) and \( C' \) such that, for any \( A \subseteq V \) satisfying \( |A| - |V \setminus A| = \Omega(n) \),

\[
\Pr \left[ T_{\text{cons}}(A) \leq C \left( \log \log n + \frac{\log n}{\log(np)} \right) \right] \geq 1 - O \left( n^{-C'} \right).
\]

(ii) If \( \tilde{r} \leq \frac{1}{2} - \varepsilon \), then \( G(2n, p, q) \) w.h.p. satisfies the following property: There are a set \( A \subseteq V \) with \( |A| - |V \setminus A| = \Omega(n) \) and two positive constants \( C \) and \( C' \) such that

\[
\Pr \left[ T_{\text{cons}}(A) \geq \exp(Cn) \right] \geq 1 - O \left( n^{-C'} \right).
\]

Theorem 1.3 (Phase transition of the Best-of-two on \( G(2n, p, q) \)). Consider the Best-of-two on \( G(2n, p, q) \) for two functions \( p = p(n) \) and \( q = q(n) \) such that \( p = o(\log n/n) \), \( q \geq \log n/n^2 \), and \( r = r(n) = q/p \) has a limit \( \tilde{r} := \lim_{n \to \infty} r(n) \). Then, for any constant \( \varepsilon > 0 \), the following holds.

(i) If \( \tilde{r} \geq \sqrt{5} - 2 + \varepsilon \), then \( G(2n, p, q) \) w.h.p. satisfies the following property: There are two positive constants \( C \) and \( C' \) such that, for any \( A \subseteq V \) satisfying \( |A| - |V \setminus A| = \Omega(n) \),

\[
\Pr \left[ T_{\text{cons}}(A) \leq C \left( \log \log n + \frac{\log n}{\log(np)} \right) \right] \geq 1 - O \left( n^{-C'} \right).
\]

(ii) If \( \tilde{r} \leq \sqrt{5} - 2 - \varepsilon \), then \( G(2n, p, q) \) w.h.p. satisfies the following property: There are a set \( A \subseteq V \) with \( |A| - |V \setminus A| = \Omega(n) \) and two positive constants \( C \) and \( C' \) such that

\[
\Pr \left[ T_{\text{cons}}(A) \geq \exp(Cn) \right] \geq 1 - O \left( n^{-C'} \right).
\]

Note that the upper bound \( T_{\text{cons}}(A) = O(\log \log n + \log n/\log(np)) \) is tight up to a constant factor if \( \log n/\log(np) \geq \log \log n \). To see this, observe that there is a set \( A \subseteq V \) such that \( T_{\text{cons}}(A) \) is at least half of the diameter. In addition, it is easy to see that the diameter of \( G(2n, p, q) \) is \( \Theta(\log n/\log(np)) \) w.h.p. [25].

We also note that the consensus time of the pull voting is \( \text{poly}(n) \) w.h.p. for any connected non-bipartite graph [29]. To the best of our knowledge, Theorems 1.2 and 1.3 provide the first nontrivial graphs where the consensus time of a multiple-choice voting process is exponentially slower than that of the pull voting.

Result II: worst-case analysis. One of the important issue in voting processes is symmetry breaking, that is, the number of iterations required to cause a small bias starting from the half-and-half state. Here, we are interested in the worst-case consensus time with respect to initial opinion configurations. To the best of our knowledge, all current results on the worst-case consensus time of multiple-choice voting processes deal with complete graphs [7, 10, 22, 26]. All previous work on non-complete graphs has involved some special bias settings (e.g., an initial bias [14–16], or a random initial opinion configuration [3, 19, 32]). This study presents the first worst-case analysis of non-complete graphs.

Theorem 1.4 (Worst-case analysis of the Best-of-three on \( G(2n, p, q) \)). Consider the Best-of-three on \( G(2n, p, q) \) for positive constants \( p \) and \( q \). If \( \frac{2}{p} > \frac{1}{7} \), then \( G(2n, p, q) \) w.h.p. satisfies the following property: There are two positive constants \( C \) and \( C' \) such that, for any \( A \subseteq V \),

\[
\Pr \left[ T_{\text{cons}}(A) \leq C \log n \right] \geq 1 - O \left( n^{-C'} \right).
\]
**Theorem 1.5** (Worst-case analysis of the Best-of-two on $G(2n, p, q)$). Consider the Best-of-two on $G(2n, p, q)$ for positive constants $p$ and $q$. If $\frac{2}{p} > \sqrt{5 - 2}$, then $G(2n, p, q)$ w.h.p. satisfies the following property: There are two positive constants $C$ and $C'$ such that, for any $A \subseteq V$,

$$\Pr \left[ T_{\text{cons}}(A) \leq C \log n \right] \geq 1 - O \left( n^{-C'} \right).$$

**Corollary 1.6.** For any constant $p > 0$, the Best-of-two and the Best-of-three on the Erdős–Rényi graph $G(n, p)$ reach consensus within $O(\log n)$ steps w.h.p. for all initial opinion configurations.

Recall that the Best-of-two and the Best-of-three on $G(n, p)$ have been studied before; however, previous works put aforementioned assumptions on the initial bias.

### 1.3 Strategy

**Known techniques and our technical contribution.** Consider a voting process on a graph $G = (V, E)$, where each vertex holds an opinion from $\{1, 2\}$, and let $A$ be the set of vertices holding opinion 1. In general, a voting process with two opinions can be seen as a Markov chain with the state space $\{1, 2\}^V$. For $A \subseteq V$, let $A'$ denote the set of vertices that hold opinion 1 in the next time step. Then, $|A'| = \sum_{v \in V} 1_{v \in A'}$ is the sum of independent random variables; thus, for any fixed $A \subseteq V$, the size $|A'|$ concentrates on $E[|A'|]$. If the underlying graph is a complete graph, the state space can be regarded as $\{0, \ldots, n\}$ (each state represents $|A|$). Therefore, $E[|A'|] = f(|A|) := |A| \left( 1 - \left( \frac{|A|}{n} \right)^2 \right) + (n - |A|) \left( \frac{|A|}{n} \right)^2 = n \left( 3 \left( \frac{|A|}{n} \right)^2 - 2 \left( \frac{|A|}{n} \right)^3 \right)$. Doerr et al. [22] exploited this idea for the Best-of-two and obtained the worst-case result for the consensus time on complete graphs. Somewhat interestingly, we also have $E[|A'||A| = f(|A|)$ in the Best-of-three.

Cooper et al. [14] extended this approach to the Best-of-two on regular expander graphs using the expander mixing lemma. Specifically, they proved that $E[|A'||A] = f(|A|) \pm O(\epsilon)$ for all $A \subseteq V$, where $\epsilon = e(n, \lambda_2) = o(n)$ is some function. Based on this approximation, they proved that the consensus time of the Best-of-two is $O(\log n)$ w.h.p. under the assumption of an initial bias of size $\Omega(\epsilon)$. In another paper, Cooper et al. [15] improved this technique and proved more sophisticated results that hold for general (i.e., not necessarily regular) expander graphs.

This study considers $G(2n, p, q)$ on the vertex set $V = V_1 \cup V_2$. Let $A_i := A \cap V_i$ for $A \subseteq V$ and $i = 1, 2$. We prove that $G(2n, p, q)$ w.h.p. satisfies $E[|A'_i|] = F_i(|A_1|, |A_2|) + O(\sqrt{n/p})$ for all $A \subseteq V$ in the Best-of-three, where $F_i : \mathbb{N}^2 \rightarrow \mathbb{N}$ is some function ($i = 1, 2$). The same concentration result holds for Best-of-two. Here, our key tool is the concentration method, specifically the Janson inequality (Lemma B.15) and the Kim-Vu concentration (Lemma B.16).

**High-level proof sketch.** Consider the Best-of-three on $G(2n, p, q)$. Let $A^{(0)}, A^{(1)}, \ldots$ be a sequence of random vertex subsets determined by $A^{(t+1)} := (A^{(t)})'$ for each $t \geq 0$. We will show that the following arguments also work for the Best-of-two, which implies Theorem 1.3. Consider a stochastic process $\alpha^{(i)} = (\alpha_1^{(i)}, \alpha_2^{(i)}) \in [0, 1]^2$, where $\alpha_i^{(t)} = |A^{(i)} \cap V_i|/n$ for $i = 1, 2$. Our technical result in the previous paragraph approximates the stochastic process $\alpha^{(i)}$ by the deterministic process $\bar{a}^{(i)}$ defined as $a^{(t+1)} = H(a^{(t)})$ and $\bar{a}^{(0)} = a^{(0)}$ for some function $H : [0, 1]^2 \rightarrow [0, 1]^2$. The function $H$ induces a two-dimensional dynamical system, which we call the *induced dynamical system*. Figure 1 illustrates the induced dynamical system of the Best-of-three of $G(2n, p, q)$.

In terms of induced dynamical systems, we obtain two results concerning $\alpha^{(i)}$. First, we show that, for any initial configuration, the process reaches one of the zero areas (a neighbor of a fixed point of some function).
Figure 1 Four types of zero areas are illustrated. The points \((a_1, a_2) = (0, 0), (1, 1)\) correspond to the consensus point. Sink areas except for the consensus points do not exist if \(r > 1/7\).

\[ H \] within a constant number of steps. In general, it is quite difficult to predicate the orbit of a dynamical system since some dynamical systems exhibit so-called the chaos phenomenon. Moreover, some dynamical systems have a loop of a period of two or more (i.e., there might exist two distinct points \(a, b\) such that \(H(a) = b \) and \(H(b) = a\)). Therefore, the proof of the convergence of the sequence \((a^{(i)})\) generated by \(H\) is difficult in general. Fortunately, the induced dynamical system of the Best-of-two and the Best-of-three on \(G(2n, p, q)\) is competitive, which is a nice property of dynamical systems [30] (see Section 2.4 for definition).

Second, we characterize the behavior of \(\alpha^{(i)}\) starting from a point in a zero area. The zero areas depending only on \(r = q/p\) can be classified into four types using the Jacobian matrix: consensus, sink, saddle, and source areas (see Figure 1 for a description). In consensus areas, we show that the process reaches consensus within \(O(\log \log n + \log n/\log(np))\) steps. In sink areas, we show that the process remains there for at least \(2\Omega(n)\) steps, and also that sink areas only appear if \(r < 1/7\). In saddle and source areas, we show that the process escapes from there within \(O(\log n)\) steps if \(p\) and \(q\) are constants. Intuitively speaking, in these two kinds of areas, there are drifts towards outside and we can apply [22]. To this end, we show that \(\text{Var}[|A'_n|] = \Omega(n)\) in saddle and source areas if \(p\) is constant. This enables us to obtain the worst-case analysis result. Indeed, any previous works working on expander graphs did not investigate the worst-case due to the lack of variance evaluation.

1.4 Related work

The consensus time of the pull voting process is investigated via its dual process, called coalescing random walk [13, 17, 29]. Recently coalescing random walks have been extensively studied, including the relationship with properties of random walks such as the hitting and mixing times [31, 40].

Other studies have focused on voting processes with more general updating rules. Cooper and Rivera [17] studied the linear voting model, whose updating rule is characterized by a set of \(n \times n\) binary matrices. This model covers the synchronous pull and the asynchronous push/pull voting processes. However, it does not cover the Best-of-two and the Best-of-three. Schoenebeck and Yu [41] studied asynchronous voting processes whose updating functions are majority-like (including the
asynchronous Best-of-(2k + 1) voting processes). They gave upper bounds on the consensus times of such models on dense Erdős–Rényi random graphs.

**Organization.** As a preliminary, we introduce precise definition of our model and auxiliary results of the stochastic block model and dynamical systems in Section 2. In Section 3, we prove Theorems 1.2 to 1.5 using the auxiliary results. In Sections 4 and 5, we prove the auxiliary results of the stochastic block model and dynamical systems, respectively. Finally, we conclude in Section 6.

## 2 | PRELIMINARIES

In this section, we introduce auxiliary results for Theorems 1.2 to 1.5. For a \( \ell \in \mathbb{N} \), let \( [\ell] := \{1, 2, \ldots, \ell\} \). For a graph \( G = (V, E) \) and \( v \in V \), let \( N(v) \) be the set of vertices adjacent to \( v \). Denote the degree of \( v \in V \) by \( \deg(v) = |N(v)| \). For \( v \in V \) and \( S \subseteq V \), let \( \deg_S(v) = |S \cap N(v)| \).

### 2.1 | Our model

First, we describe the formal definition of voting processes in this study. To discuss the Best-of-two and the Best-of-three together, we consider the following general framework that contains these models as special cases.

**Definition 2.1** \(((f_1, f_2))\)-polynomial voting process.** Let \( G = (V, E) \) be a graph, where each vertex holds an opinion from \( \{1, 2\} \). Let \( f_1, f_2 : [0, 1] \to [0, 1] \) be polynomials. For the set \( A \) of vertices with opinion 1, let \( A' \) denote the set of vertices with opinion 1 after an update. In the \((f_1, f_2))\)-polynomial voting process, \( A' := \{v \in V : X_v = 1\} \) where \((X_v)_{v \in V}\) are independent binary random variables satisfying

\[
\Pr[X_v = 1] = \begin{cases} 
    f_1 \left( \frac{\deg_A(v)}{\deg(v)} \right) & (v \in A, \text{ i.e., } v \text{ has opinion 1}) \\
    f_2 \left( \frac{\deg_A(v)}{\deg(v)} \right) & (v \in V \setminus A, \text{ i.e., } v \text{ has opinion 2}).
\end{cases}
\]

**Example: Best-of-three.** For all \( A \subseteq V \) and \( v \in V \), it is straightforward to see that

\[
\Pr[X_v = 1] = \left( \frac{3}{3} \right) \left( \frac{\deg_A(v)}{\deg(v)} \right)^3 + \left( \frac{3}{2} \right) \left( \frac{\deg_A(v)}{\deg(v)} \right)^2 \left( 1 - \frac{\deg_A(v)}{\deg(v)} \right)
\]

in the Best-of-three. Hence the Best-of-three is a \((3x^2 - 2x^3, 3x^2 - 2x^3)\)-polynomial voting process.

**Example: Best-of-two.** In the Best-of-two, we have

\[
\Pr[X_v = 1] = \begin{cases} 
    1 - \left( 1 - \frac{\deg_A(v)}{\deg(v)} \right)^2 & (v \in A) \\
    \left( \frac{\deg_A(v)}{\deg(v)} \right)^2 & (v \in V \setminus A)
\end{cases}
\]

for all \( A \subseteq V \) and \( v \in V \). Hence the Best-of-two is an \((3(-2x), x^2)\)-polynomial voting process.

For a given initial vertex subset \( A^{(0)} \subseteq V \) of opinion 1, we are interested in the behavior of the Markov chain \((A^{(t)})_{t=0}^{\infty})\), that is, the sequence of random vertex subsets determined by \( A^{(t+1)} := (A^{(t)})' \) for each \( t \geq 0 \).
2.2 | Concentration result for the stochastic block model

This study considers polynomial voting processes on the stochastic block model $G(2n, p, q)$ defined in Definition 1.1. Consider a fixed stochastic block model $G(2n, p, q)$. Let $A_i := V_i \cap A$ for $A \subseteq V$ and $i = 1, 2$. Since $|A_i'| = \sum_{v \in V} X_v$ from Definition 2.1, the Hoeffding bound (Lemma B.9) implies that, for any fixed $A \subseteq V$ and any $i = 1, 2$,

$$|A_i'| = \mathbb{E}[|A_i'|] = O\left(\sqrt{n \log n}\right)$$

holds w.h.p. (here, the probability is over the randomness of the process). From Definition 2.1, for any fixed $A \subseteq V$, we have

$$\mathbb{E}[|A_i'|] = \sum_{v \in V_i} \mathbb{E}[X_v] = \sum_{v \in A_i} \left(\frac{\deg_A(v)}{\deg(v)}\right) + \sum_{v \in V \setminus A_i} \left(\frac{\deg_A(v)}{\deg(v)}\right).$$

In general, (2) is a random variable since $G(2n, p, q)$ is a random graph. Our key ingredient is the following concentration result for this random variable.

Definition 2.2 (f-good graph). Let $G = (V, E)$ be a graph on $2n$ vertices. Let $f : [0, 1] \rightarrow [0, 1]$ be a function, $V_1, V_2 \subseteq V$ be a partition of $V$ such that $|V_1| = |V_2| = n$, and $p, q \in [0, 1]$ be parameters satisfying $p \geq q > 0$ ($p = p(n)$ and $q = q(n)$ may depend on $n$). We say a graph $G = (V, E)$ is $f$-good for a partition $V_1, V_2$ and parameters $p, q$ if the graph satisfies the following properties.

(P1) It is connected and non-bipartite.

(P2) There is a positive constant $C_1$ such that, for all $A, S \subseteq V$ and $i \in \{1, 2\}$,

$$\left|\sum_{v \in S \cap V_i} f\left(\frac{\deg_A(v)}{\deg(v)}\right) - |S \cap V_i|f\left(\frac{|A_i|p + |A_{3-i}|q}{n(p + q)}\right)\right| \leq C_1 \sqrt{\frac{n}{p}}.$$

(P3) There is a positive constant $C_2$ such that, for all $A \subseteq V, S \in \{A, V \setminus A\}$, and $i \in \{1, 2\},$

$$\sum_{v \in S \cap V_i} f\left(\frac{\deg_A(v)}{\deg(v)}\right) \leq |S \cap V_i|f\left(\frac{|A_i|p + |A_{3-i}|q}{n(p + q)}\right) + C_2 |A| \sqrt{\frac{\log n}{np}}.$$

Theorem 2.3 (Main technical theorem). Consider a stochastic block model $G(2n, p, q)$ on a vertex set $V_1 \cup V_2$. Let $f : [0, 1] \rightarrow [0, 1]$ be a polynomial, and $p = o(\log n/n)$ and $q \geq \log n/n^2$ be functions. Then $G(2n, p, q)$ is $f$-good for a partition $(V_1, V_2)$ and parameters $p, q$ w.h.p.

Note that the proof of (P1) is not difficult since $p = o(\log n/n)$ and $q \geq \log n/n^2$ [25]. However, proving (P2) and (P3) is challenging; we show these in Section 4.

From Theorem 2.3, $G(2n, p, q)$ is $f_{1}$- and $f_{2}$-good for a partition $(V_1, V_2)$ and parameters $p, q$ w.h.p. Henceforth, we consider a polynomial voting process on a (fixed) $f_{1}$- and $f_{2}$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$. Let $r := \frac{q}{p}, \alpha_i := \frac{|A_i|}{n}$ and $a = \alpha_1 + \alpha_2$. From (2), (P2), and (P3), we have

$$\mathbb{E}[a_i^r] = \alpha_i f_1\left(\frac{\alpha_i + r\alpha_{3-i}}{1 + r}\right) + (1 - \alpha_i)f_2\left(\frac{\alpha_i + r\alpha_{3-i}}{1 + r}\right) + O\left(\min\left\{\frac{1}{\sqrt{np}}, a\sqrt{\frac{\log n}{np}}\right\}\right)$$

for all $A \subseteq V$ and $i = 1, 2$. Here, we note that the additive error $O\left(\min\left\{\frac{1}{\sqrt{np}}, a\sqrt{\frac{\log n}{np}}\right\}\right)$ depends on $|A|$. This property plays a key role in the proof of the consensus.
2.3 Induced dynamical system

Consider an \((f_1, f_2)\)-polynomial voting process on a fixed graph that is \(f_1\)- and \(f_2\)-good for a partition \((V_1, V_2)\) and parameters \(p, q\). Suppose that \(r = \frac{q}{p}\) is a constant. Define two functions \(H_1, H_2 : [0, 1]^2 \to [0, 1]\) as

\[
H_i(a_1, a_2) = a_1f_1 \left( \frac{a_i + ra_{3-i}}{1 + r} \right) + (1-a_i)f_2 \left( \frac{a_i + ra_{3-i}}{1 + r} \right) \quad \text{for } i = 1, 2.
\]

From (1) and (3), for all \(A \subseteq V\) and \(i \in \{1, 2\}\), it holds w.h.p. that

\[
|a'_i - H_i(a_1, a_2)| = O \left( \sqrt{\frac{1}{np}} + \sqrt{\frac{\log n}{n}} \right).
\]

Throughout this study, we use \(\alpha = (a_1, a_2)\) and \(\alpha' = (a_1', a_2')\) as vector-valued random variables. Equation (5) leads us to the dynamical system \(H\), where we define \(H : \mathbb{R}^2 \to \mathbb{R}^2\) as

\[
H : a \mapsto (H_1(a), H_2(a)).
\]

By combining (5) with the Lipschitz condition (see Appendix B.2), it is not difficult to show the following result.

**Theorem 2.4.** Consider an \((f_1, f_2)\)-polynomial voting process on an \(f_1\)- and \(f_2\)-good graph for a partition \((V_1, V_2)\) and parameters \(p, q\), starting with a vertex set \(A^{(0)} \subseteq V\) of opinion 1. Let \((A^{(t)})_{t=0}^{\infty}\) be a sequence of random vertex subsets defined by \(A^{(t+1)} := (A^{(t)})'\) for each \(t \geq 0\). Let \((\alpha^{(t)})_{t=0}^{\infty}\), where \(\alpha^{(t)} = (a_1^{(t)}, a_2^{(t)})\) and \(a_i^{(t)} = |A^{(t)} \cap V_i|/n\). Let \(H\) be the mapping defined in (4) and (6). Let \((a^{(t)})_{t=0}^{\infty}\) be the sequence defined as

\[
\begin{cases}
  a^{(0)} = \alpha^{(0)}, \\
  a^{(t+1)} = H(a^{(t)}).
\end{cases}
\]

Then, there is a constant \(C > 0\) such that

\[
\forall 0 \leq t \leq n^{\Theta(1)}, \forall A^{(0)} \subseteq V : \Pr \left[ \|a^{(t)} - a^{(0)}\|_\infty \leq C' \left( \frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} \right) \right] \geq 1 - n^{-\Theta(1)}.
\]

Broadly speaking, Theorem 2.4 approximates the behavior of \(\alpha^{(t)}\) by the orbit \(a^{(t)}\) of the corresponding dynamical system \(H\). We call the mapping \(H\) the induced dynamical system. Note that the mapping \(H\) of Theorem 2.4 is the induced dynamical system.

**Proof of Theorem 2.4.** From (5), for all \(A \subseteq V\), there is some positive constant \(C_1\) such that

\[
\|\alpha' - H(\alpha)\|_\infty \leq C_1 \left( \frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} \right)
\]

holds w.h.p. Since \(f_1\) and \(f_2\) are polynomials, the function \(H\) satisfies the Lipschitz condition. That is, there is a positive constant \(C_2\) such that \(\|H(x) - H(y)\|_\infty \leq C_2\|x - y\|_\infty\) holds for any \(x, y \in [0, 1]^2\) (see Appendix B.2). Then, we have

\[
\|\alpha^{(t)} - a^{(t)}\|_\infty = \|\alpha^{(t)} - H(\alpha^{(t-1)}) + H(\alpha^{(t-1)}) - H(a^{(t-1)})\|_\infty \\
\leq \|\alpha^{(t)} - H(\alpha^{(t-1)})\|_\infty + C_2\|\alpha^{(t-1)} - a^{(t-1)}\|_\infty
\]
In this subsection, for a map $d$ function, where $S$ is defined in 2.4. Therefore, we focus on the behavior $C$ for sufficiently large constant $\delta$. This suggests a dynamical system $A$ instead of $S$. Consider $\delta = (\delta_1, \delta_2) := (\alpha_1 - \alpha_2, \alpha_1 + \alpha_2 - 1)$, $\delta' = (\delta_1', \delta_2') := (\alpha_1' - \alpha_2', \alpha_1' + \alpha_2' - 1)$.

Axes $\delta_1$ and $\delta_2$ correspond to the dotted lines of Figure 1. Note that $J$ satisfies $|\delta_1| + |\delta_2| \leq 1$. From (3), for any $A \subseteq V$ and any $i = 1, 2$, $E[\delta_i'] = T_i(\delta_1, \delta_2) + O \left( \frac{1}{\sqrt{np}} \right)$ holds, where

$$T_1(d_1, d_2) := H_1 \left( \frac{1 + d_1 + d_2}{2}, \frac{1 - d_1 + d_2}{2} \right) - H_2 \left( \frac{1 + d_1 + d_2}{2}, \frac{1 - d_1 + d_2}{2} \right),$$

$$T_2(d_1, d_2) := H_1 \left( \frac{1 + d_1 + d_2}{2}, \frac{1 - d_1 + d_2}{2} \right) + H_2 \left( \frac{1 + d_1 + d_2}{2}, \frac{1 - d_1 + d_2}{2} \right) - 1.$$  

This suggests a dynamical system $T(d) = (T_1(d), T_2(d))$. Here, we use $d = (d_1, d_2)$ as a specific point and $\delta = (\delta_1, \delta_2)$ as a vector-valued random variable. Note that the dynamical system $T$ is symmetric: specifically, $T_1(\pm d_1, \mp d_2) = \mp T_1(d_1, d_2)$ and $T_2(\pm d_1, \mp d_2) = \mp T_2(d_1, d_2)$ hold. To see this, observe $H_1(\alpha_2, \alpha_1) = H_3(\alpha_2, \alpha_1)$ (exchange $V_1$ and $V_2$) and $H_1(1 - \alpha_1, 1 - \alpha_2) = 1 - H_1(\alpha_1, \alpha_2)$ (consider $V \setminus A'$ instead of $A'$). Consider $\delta^{(i)} = (\delta_1^{(i)}, \delta_2^{(i)}) = (\alpha_1^{(i)} - \alpha_2^{(i)}, \alpha_1^{(i)} + \alpha_2^{(i)} - 1)$ and $(d^{(i)})_{i=0}^{\infty}$, where $d^{(0)} = \delta^{(0)}$ and $d^{(i+1)} = T(d^{(i)})$ for each $i \geq 0$. From Theorem 2.4, it holds w.h.p. that

$$\|\delta^{(i)} - d^{(i)}\|_\infty \leq C' \left( \frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} \right),$$

for sufficiently large constant $C > 0$, any $0 \leq t \leq n^{(1)}$ and any initial configuration $A^{(0)} \subseteq V$.

Let

$$S := \{(d_1, d_2) \in [0, 1]^2 \colon d_1 + d_2 \leq 1\}. \quad (10)$$

We will show in Lemma 3.1 that, $d^{(i+1)} \in S$ holds for any $d^{(i)} \in S$ in the Best-of-two and the Best-of-three. Therefore, we focus on the behavior $(d^{(i)})_{i=0}^{\infty}$ within $S$.

### 2.4 | Orbit convergence

In this subsection, for a map $T : S \to S$ and an initial point $x \in S$, we present a sufficient condition for the convergence of the orbit (i.e., the sequence $(T^n(x))_{n \geq 0}$). We call a point $x$ a fixed point if $T(x) = x$ holds.

**Theorem 2.5.** Let $T : S \to S$ be an injective and $C^1$ (i.e., differentiable and its derivative is continuous) function, where $S$ is defined in (10). Let $J = (j_{ij})_{i,j \in [2]}$ be the Jacobian matrix of $T$ at $x \in S$. Suppose that $J$ satisfies

\[ L \geq 0, \quad \text{and} \quad J \geq 0. \]
For any \( x \in S \), it hold that \( j_{11}, j_{22} \geq 0 \) and \( j_{12}, j_{21} \leq 0 \), and

(C2) For any \( x \in S \setminus \{(0, 1)\} \), the determinant satisfies \( \det J > 0 \).

Then, for any \( x \in S \), there is the limit \( \lim_{n \to \infty} T^n(x) \) and the limit is a fixed point of \( T \).

We will show that \( T \) of the Best-of-two (the Best-of-three) satisfies both (C1) and (C2). Roughly speaking, from (9) and Theorem 2.5, it holds w.h.p. that \( \mathcal{S}^{(0)} \) approaches around a fixed point after constant steps (see Section 3 for details).

To show Theorem 2.5, we introduce some previous techniques on competitive dynamical systems.

**Definition 2.6** (Competitive dynamical system). For two points \( x = (x_1, x_2) \) and \( y = (y_1, y_2) \), we write \( x \leq y \) if \( x_1 \leq y_1 \) and \( y_2 \leq x_2 \) hold. For \( S \subseteq \mathbb{R}^2 \), a map \( T : S \to S \) is competitive if \( T(x) \leq \kappa T(y) \) whenever \( x \leq \kappa y \).

See [30] for the background of competitive dynamical systems. For two points \( x = (x_1, x_2) \) and \( y = (y_1, y_2) \), we write \( x \leq y \) if \( x_1 \leq y_1 \) and \( x_2 \leq y_2 \). We write \( x < y \) if \( x_1 < y_1 \) and \( x_2 < y_2 \). The following known result provides a sufficient condition for the orbit convergence of a competitive dynamical system.

**Theorem 2.7** (Theorem 5.28 of [30]). Suppose that a competitive map \( T : S \to S \) satisfies \( x \leq y \) for any \( x, y \in S \) if \( T(x) < T(y) \). Then, for any \( x \in S \), the sequence \( (T^n(x))_{n \geq 0} \) converges to some fixed point of \( T \).

**Proof of Theorem 2.5.** It suffices to check the condition of Theorem 2.7 holds. First, we claim that \( T \) is competitive. Let \( T(x) = (T_1(x), T_2(x)) \) for \( x = (x_1, x_2) \in S \). From (C1), the function \( T_j \) is nondecreasing on \( x_j \) and is nonincreasing on \( x_{3-j} \). Therefore, for any \( (a, b), (c, d) \in S \) of \( (a, b) \leq \kappa (c, d) \), we have \( T(a, b) \leq \kappa T(c, d) \). In other words, \( T \) is competitive.

Second, we claim that the inverse \( T^{-1} \) satisfies \( T^{-1}(x) \leq T^{-1}(y) \) whenever \( x \leq y \). Let \( U := T^{-1} \) and \( U(x) = (U_1(x), U_2(x)) \) for \( x = (x_1, x_2) \in S \). By the inverse function theorem (Theorem B.3), the Jacobian matrix \( K \) of \( U \) at \( x \in S \setminus \{(0, 1)\} \) is the inverse of that of \( T \). Thus, from (C2), we have \( \frac{\partial U_1}{\partial x_1}(x) \geq 0 \) for any \( x \in S \setminus \{(0, 1)\} \). Hence the functions \( U_1(x_1, x_2) \) and \( U_2(x_1, x_2) \) are nondecreasing on both \( x_1 \) and \( x_2 \). Therefore for any two points \( (a, b), (c, d) \in S \) of \( (a, b) \leq (c, d) \), we have \( U_1(a, b) \leq U_1(c, d) \) and \( U_2(a, b) \leq U_2(c, d) \) (note that if \( (a, b) = (0, 1) \) then \( (c, d) \) must be \( (0, 1) \) and we are done).

For any points \( x, y \in S \) of \( T(x) < T(y) \), the second claim implies that \( x = T^{-1}(T(x)) \leq T^{-1}(T(y)) = y \). Therefore, we can apply Theorem 2.7.

### 2.5 Local dynamics around fixed points

Consider an \((f_1, f_2)\)-polynomial voting process on an \( f_1 \)- and \( f_2 \)-good graph for a partition \((V_1, V_2)\) and parameters \( p, q \). Let \( H \) be the induced dynamical system.

In this subsection, we focus on the behavior of \((\alpha(t))_{t=0}^{\infty}\) when the initial point \( \alpha(0) \) is around a fixed point of \( H \) (i.e., a point \( x \) such that \( H(x) = x \) holds). In this case, Theorem 2.4 does not provide enough information about the dynamics. In dynamical system theory, a common approach for the local behavior around fixed points is to consider the Jacobian matrix. In what follows, we will investigate the local dynamics from the viewpoint of the maximum singular value and eigenvalue of the Jacobian matrix. For the readability, we put the proofs of each statements in Section 5.

**Sink point.** We begin with defining the notion of sink points. Recall that the singular value of a matrix \( A \) is the positive square root of the eigenvalue of \( A^\top A \) (see Appendix B.1 for formal definition and
basic properties). For \( x \in [0, 1]^2 \) and \( r > 0 \), let \( B(x, r) := \{ y \in \mathbb{R}^2 : \|x - y\|_\infty < r \} \) denote the open ball of radius \( r \) with respect to the \( \ell^\infty \)-norm.

**Definition 2.8** (Sink point). Consider a dynamical system \( H \). A fixed point \( a^* \in \mathbb{R}^2 \) is a sink point if the Jacobian matrix \( J \) at \( a^* \) satisfies \( \sigma_{\max} < 1 \), where \( \sigma_{\max} \) is the largest singular value of \( J \).

**Proposition 2.9.** Consider an \( (f_1, f_2) \)-polynomial voting process on an \( f_1 \)- and \( f_2 \)-good graph for a partition \((V_1, V_2)\) and parameters \( p, q \) such that \( r = \frac{2}{p} \) is a constant. Let \( H \) be the induced dynamical system. Then, for any sink point \( a^* \) and any sufficiently small \( \epsilon = o(\sqrt{1/np}) \),

\[
\Pr \left[ \alpha' \notin B(a^*, \epsilon) \mid \alpha \in B(a^*, \epsilon) \right] \leq \exp(-\Omega(\epsilon^2 n))
\]

holds. In particular, let \( \tau := \inf \{ t \in \mathbb{N} : \alpha^{(t)} \notin B(a^*, \epsilon) \} \) be a stopping time. Then, \( \tau \geq \exp(\Omega(\epsilon^2 n)) \) holds w.h.p. conditioned on \( \alpha^{(0)} \in B(a^*, \epsilon) \) for any \( \epsilon \) satisfying \( \epsilon = o(\max\{1/\sqrt{np}, \sqrt{\log n/n}\}) \).

**Fast consensus.** We consider the case in which the initial opinion configuration \( A_0 \) is close to consensus. We first observe that, in the Best-of-two and Best-of-three, the Jacobian matrix at the consensus point (i.e., \( \alpha = (0, 0), (1, 1) \)) is the all-zero matrix.

**Proposition 2.10.** Consider an \( (f_1, f_2) \)-polynomial voting process on an \( f_1 \)- and \( f_2 \)-good graph for a partition \((V_1, V_2)\) and parameters \( p, q \) such that \( \frac{2}{p} \) is a constant. Suppose that the Jacobian matrix at the point \( \alpha = (0, 0) \) is the all-zero matrix. Then, there are positive constants \( C_1, C_2, \) and \( \delta > 0 \) such that

\[
\Pr \left[ T_{\text{cons}}(A) \leq C_1 \left( \log \log n + \frac{\log n}{\log np} \right) \right] \geq 1 - n^{-C_2}
\]

hold for all \( A \subseteq V \) satisfying \( |A| \leq \delta n \).

**Escape from a fixed point.** Let \( a^* \in \mathbb{R}^2 \) be a fixed point of the induced dynamical system \( H \). Let \( J \) be the Jacobian matrix of \( H \) at \( a^* \) and \( \lambda_1, \lambda_2 \) be its eigenvalues. Let \( u_i \) be the eigenvector of \( J \) corresponding to \( \lambda_i \). Suppose that \( u_1, u_2 \) are linearly independent. Then, we can rewrite \( J \) as

\[
J = U^{-1} \Lambda U,
\]

where \( \Lambda = \text{diag}(\lambda_1, \lambda_2) \) and \( U = (u_1 \ u_2)^{-1} \). For a fixed point \( a^* \in \mathbb{R}^2 \), let \( \beta = (\beta_1, \beta_2) \) be a vector-valued random variable defined as

\[
\beta = U(a^* - a^*).
\]

From the Taylor expansion of \( H \) at \( a^* \), we have \( E[\beta'] \approx \Lambda \beta \) if \( \|\beta\|_\infty \) is sufficiently small.

Recall that \( B(a, R) \) is the open ball of radius \( R \) (with respect to the \( \ell^\infty \)-norm) of center \( a \). If \( |\lambda_i| > 1 \) for some \( i \in [2] \), one may expect that \( a^{(r)} \notin B(a^*, \epsilon_0) \) holds for any \( A^{(0)} \subseteq V \) and for some constant \( \epsilon_0 > 0 \). We prove this under some assumptions.

**Assumption 2.11** (Basic assumptions). Consider an \( (f_1, f_2) \)-polynomial voting process on an \( f_1 \)- and \( f_2 \)-good graph for a partition \((V_1, V_2)\) and parameters \( p, q \), where \( p \geq q \geq 0 \) are constants. Let \( H \) be the induced dynamical system. Let \( a^* \) be a fixed point and \( J \) be the Jacobian matrix of \( H \) at \( a^* \). We assume that \( J \) satisfies
Let $\mathbf{a}^*$ be a fixed point satisfying Assumption 2.11. Suppose that the eigenvalues $\lambda_1, \lambda_2$ of the Jacobian matrix $J$ at $\mathbf{a}^*$ satisfy $|\lambda_i| \neq 1$ for all $i \in [2]$. Then, for some $t = O(\log n)$ and some constant $\epsilon' > 0$, it holds w.h.p. that $|\beta^{(t)}|_\infty > \epsilon'$, and $|\beta^{(t)}_j| \leq \epsilon'$ for any $j \in [2]$ of $|\lambda_j| \leq 1$.

We consider the case of $\lambda_i = 1$ for some $i$ as follows.

**Proposition 2.12.** Let $\mathbf{a}^*$ be a fixed point satisfying Assumption 2.11. Suppose that there is a constant $\epsilon^* > 0$ satisfying

(B1) There are two positive constants $\epsilon_1$ and $C$ such that $|E[\beta^{(t)}_j]| \geq (1 + \epsilon_1)|\beta| - \frac{C}{\sqrt{n}}$ holds for any $A \subseteq V$ of $|\beta| \leq \epsilon^*$ and $i \in [2]$ of $|\lambda_i| > 1$.

(B2) For any $i \in [2]$ of $|\lambda_i| \leq 1$ and any $A \subseteq V$ of $|\beta| \leq \epsilon^*$, it holds that $\Pr[|\beta^{(t)}_i| \leq \epsilon^*] \geq 1 - n^{-\Omega(1)}$.

Then, for some $t = O(\log n)$ and some constant $\epsilon' > 0$, it holds w.h.p. that $|\beta^{(t)}|_\infty > \epsilon'$, and $|\beta^{(t)}_j| \leq \epsilon'$ for any $j \in [2]$ of $|\lambda_j| \leq 1$.

## 3 | BEST-OF-TWO AND BEST-OF-THREE

This section provides the proofs of main results formulated in Theorems 1.2 to 1.5. For notational convenience, let $f^{Bo3}(x) := 3x_2 - 2x_3, f^{Bo2}(x) := x(2 - x)$, and $f^{Bo3}(x) := x^2$. Then, the Best-of-two is an $(f^{Bo3}, f^{Bo2})$-polynomial voting process, while the Best-of-three is an $(f^{Bo3}, f^{Bo3})$-polynomial voting process. Consider the Best-of-two on an $f^{Bo3}$-good graph and the Best-of-three on an $f^{Bo3}$-good graph. We consider the behavior of $\tilde{\delta}$ (see (8)).

Let $u := \frac{1-r}{1+r}$. Then $E[\delta^{(t)}_i] = T_i(\delta_1, \delta_2) + O\left(\frac{1}{1+\epsilon'}\right)$, where, in the Best-of-three,

$$T_1(d_1, d_2) := \frac{ud_1}{2} (3 - (ud_1)^2 - 3d_2^2), \quad T_2(d_1, d_2) := \frac{d_2}{2} (3 - 3(ud_1)^2 - d_1^2), \quad (12)$$

and in the Best-of-two,

$$T_1(d_1, d_2) := \frac{d_1}{2} \left( (2u + 1) - (ud_1)^2 - (2u + 1)d_2^2 \right), \quad T_2(d_1, d_2) := \frac{d_2}{2} \left( 3 - u(2 + u)d_1^2 - d_2^2 \right). \quad (13)$$

Note that $T : S \to S$, where $S$ is defined as (10). For notational convenience, $T(\mathbf{d})$ is denoted as $\mathbf{d}'$.

**Lemma 3.1.** Consider the Best-of-two or Best-of-three. For any $\mathbf{d} \in S$, it holds that $\mathbf{d}' \in S$.

**Proof.** In this study, we prove Lemma 3.1 for the Best-of-three. The case of the Best-of-two can be shown in the same way. If $(d_1, d_2) \in S$, we have $3d_2^2 + (ud_1)^2 \leq \max\{3, u^2\} \leq 3$ and $d_1^2 + 3(ud_1)^2 \leq \max\{1, 3u^2\} \leq 3$. Hence, from (12), we have $d_1'^2 \geq 0$ and $d_2'^2 \geq 0$. Let $x = \frac{1+u+u}{2}$ and $y = \frac{1+u-d_1}{2}$. Then, $(d_1, d_2) \in S$ implies that $\frac{1}{2} \leq x \leq 1$ and $0 \leq y \leq 1$. In addition, a simple calculation yields $T_1(d_1, d_2) +$
FIGURE 2 The induced dynamical system \( H \) of the Best-of-three. The horizontal and vertical axes correspond to \( \alpha_1 \) and \( \alpha_2 \), respectively. Points \( d_i^* \) are the fixed points of \( T \) in \( S \). In (A), the only sink point is \( d_4^* \), which is the consensus point

\[ T_2(d_1, d_2) = 3 \left( \frac{x+y}{1+r} \right)^2 - 2 \left( \frac{x+y}{1+r} \right)^3 \leq 1, \]

where \( r = 1 - \frac{x+y}{1+z} \leq 1 \) and the function \( f : z \mapsto 3z^2 - 2z^3 \) satisfies \( f(z) \leq f(1) = 1 \) for all \( 0 \leq z \leq 1 \). Therefore, \( d' \in S \).

From Lemma 3.1 and the symmetry of \( T \), it is sufficient to consider the case of \( S(0) \in S \).

3.1 Best-of-three

It is straightforward to check that the fixed points of (12) in \( S \) are \( d_1^*, d_2^*, d_3^*, d_4^* \), where

\[
d_i^* := \begin{cases} 
(0, 0) & \text{if } i = 1, \\
\left( \sqrt{\frac{3u-2}{u^3}}, 0 \right) & \text{if } i = 2 \text{ and } u \geq \frac{2}{3}, \\
\left( \frac{1}{4u^3}, \sqrt{\frac{4u-3}{4u}} \right) & \text{if } i = 3 \text{ and } u \geq \frac{3}{4}, \\
(0, 1) & \text{if } i = 4.
\end{cases}
\] (14)

The Jacobian matrix at \((d_1, d_2)\) of the dynamical system (12) is

\[
J = \frac{3}{2} \begin{pmatrix}
  u(1 - (ud_1)^2 - d_2^2) & -2ud_1d_2 \\
  -2ud_1d_2 & 1 - (ud_1)^2 - d_2^2 
\end{pmatrix}.
\] (15)

**Proposition 3.2** (Orbit convergence). Let \((d^{(i)})_{i=0}^{\infty}\) be a sequence such that \(d^{(0)} \in S \) and \(d^{(i+1)} = T(d^{(i)})\). Then \( \lim_{i \to \infty} d^{(i)} = d_i^* \) for some \( i \in \{1, 2, 3, 4\} \). Furthermore, let \( u < \frac{3}{4} \) and \( d^{(0)} = (d_1^{(0)}, d_2^{(0)}) \in S \) satisfy \( d_2^{(0)} > c \) for some constant \( c > 0 \). Then \( \lim_{i \to \infty} d^{(i)} = d_4^* \).
From Table 1, it is straightforward to check that the points $d_1^*$, $d_2^*$, $d_3^*$, and $d_4^*$ of $T$ (if $u \geq \frac{2}{3}$). From the first statement of Proposition 3.2, we have $\lim_{t \to \infty} d^{(i)}(t) = d_i^*$ for some $i \in \{1, 2, 4\}$. If $i = 4$, then the claim holds. Suppose that $i = 1$. Then, for any $e > 0$, there is $T \in \mathbb{N}$ such that $\|d^{(i)}(t) - d_1^*\|_\infty < e$ for all $t \geq T$. Recall that the assumption implies $d_2^{(0)} > 0$. From (12), it is easy to check that $d_2^* > d_2^*$ if $d_2 > 0$. Fix a sufficiently small constant $e > 0$ and a point $d$ such that $\|d - d_1^*\|_\infty < e$ and $d_2 > 0$ hold. From (12), we have $d_2^* \geq 1.5(d_2 - 3u^2 e^2 - e^2) > 1.49d_2$. Therefore, if $d_2^{(0)}$ satisfies $\|d^{(0)}(t) - d_1^*\|_\infty < e$ and $d_2^{(0)} > 0$, then either $d_1^{(0)} > e$ or $d_2^{(0)} > e$ holds for some $t = O(e^{-1}$. This contradicts the assumption that $\lim_{t \to \infty} d^{(i)}(t) = d_i^*$. Therefore, we have $\lim_{t \to \infty} d^{(i)}(t) \neq d_i^*$. Similarly, we can show that $\lim_{t \to \infty} d^{(i)}(t) \neq d_i^*$ (when $\frac{2}{3} \leq u < \frac{3}{4}$).

Now we focus on the behavior of $d^{(0)}$ when $d^{(0)}$ is around a fixed point. Table 1 shows the properties of eigenvalues of the Jacobian matrix at $d_i^*$ for each $i = 1, \ldots, 4$.

Recall that $B(x, r)$ is an open ball of radius $r$ with respect to the $\ell^\infty$-norm. For $d = (d_1, d_2) \in \mathbb{R}^2$, let $(d)_+ := (|d_1|, |d_2|) \in \mathbb{R}^2$.

**Proposition 3.3.** Consider the Best-of-three on an $f^{Bo_3}$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = qp < 1/7$ is a constant. Then there is a constant $e = e(r) > 0$ satisfying

\[
\Pr \left[ \left< \delta^t \right>_+ \notin B(d_2^*, e) \mid \left< \delta^t \right>_+ \in B(d_2^*, e) \right] \leq \exp(-\Omega(n)).
\]

In particular, $T_{\text{cons}}(A) = \exp(\Omega(n))$ w.h.p. for any $A$ satisfying $\left< \delta^t \right>_+ \in B(d_2^*, e)$.

**Proof.** From Table 1, it is straightforward to check that the points $d_2^*$ and $d_2^*$ are sink points if $r < \frac{1}{7}$ (or equivalently, $u > \frac{3}{4}$). Therefore, Proposition 3.3 immediately follows from Proposition 2.9.

**Proposition 3.4.** Consider the Best-of-three on an $f^{Bo_3}$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = qp$ is a constant. Then, for some constant $e = e(r) > 0$, $T_{\text{cons}}(A) \leq O(log log n + log n/ log(np)) holds w.h.p. for any $A \subseteq V$ satisfying min[|A|, 2n - |A|] \leq en.$

Note that $det J > 0$ for any $d \in S \setminus \{(0,1)\}$. Then, from the inverse function theorem (Theorem B.3), $T$ is injective on $S$.\n
---

**Table 1** Each $(c_1, c_2)$ represents the property of the eigenvalues $\lambda_1 \geq \lambda_2$ of the corresponding Jacobian matrix. In particular, each $c_2$ represents the sign of $\lambda_1 - 1$. For example, $(+, 0)$ means that $\lambda_1 > \lambda_2 = 1$ and $(+, -)$ means that $\lambda_1 > \lambda_2$.
Proof. Note that the Jacobian matrix at \( \mathbf{d}_1^* \) is the all-zero matrix and the same holds for \(-\mathbf{d}_2^*\). Let \( \varepsilon > 0 \) be sufficiently small constant. If \( A \) satisfies \( |A| \leq e n \), apply Proposition 2.10. If \( A \) satisfies \( |A| \geq (2 - \varepsilon)n \), apply Proposition 2.10 for \( V \setminus A \).

Proposition 3.5. Consider the Best-of-three on an \( f^{Bo3} \) and \( f^{Bo3}(1 - f^{Bo3}) \)-good graph for a partition \((V_1, V_2)\) and parameters \( p, q \) such that \( p \) and \( q \) are constants. If \( q/p > 17 \) and \( |\delta_2^{(0)}| = o(1) \), then it holds w.h.p. that \( |\delta_2^{(t)}| > \kappa \) for some \( t = O(\log n) \) and some constant \( \kappa > 0 \).

Proof. Suppose that \( u < \frac{3}{4} \) (or equivalently, \( r > \frac{1}{7} \)) and \( p \geq q > 0 \) are constants. From Proposition 3.2, we may assume

\[
\delta^{(0)} \in \bigcup_{i \in \{1, 2\}} B(d_i^*, \varepsilon_2)
\]

for a sufficiently small constant \( \varepsilon_2 > 0 \). We use Propositions 2.12 and 2.13.

First, we check the condition (A2) of Assumption 2.11. Note that, for fixed \( A \subseteq V \), the variance \( \text{Var}[|A_i'|] \) can be written as \( \text{Var}[|A_i'|] = \sum_{i \in V} g \left( \frac{\deg_i(v)}{\deg(v)} \right) \) for any \( A \subseteq V \), where \( g(x) := f^{Bo3}(x)(1 - f^{Bo3}(x)) \). Therefore, from the property (P2) of \( g \)-goodness, there are two positive constants \( C_1 \) and \( C_2 > 0 \) such that

\[
\forall A \subseteq V, \forall i \in \{1, 2\} : \left| \text{Var}[|A_i'|] - n \cdot g \left( \frac{|A_i|p + |A_3 - i|q}{n(p + q)} \right) \right| \leq C_2 \sqrt{\frac{n}{p}}. \tag{17}
\]

Using \( z_i := \frac{a_i + a_{i-1}}{1 + r} \), we can rewrite \( \text{Var}[\alpha_i'] = \text{Var}[|A_i'|]/n^2 \) as

\[
\text{Var}[\alpha_i'] = \frac{z^2_i(3 - 2z_i)(1 - z_2)^2(2z_i + 1)}{n} \pm O \left( \frac{1}{\sqrt{n^3p}} \right).
\]

Note that \( \text{Var}[\alpha_i'] = \Omega(n^{-1}) \), if \( a_1 < 1 - \varepsilon_3 \) or \( a_2 < 1 - \varepsilon_3 \) for some constant \( \varepsilon_3 > 0 \). Therefore, the statement (A2) holds for every \( \delta \) satisfying \( \delta \in \bigcup_{i \in \{1, 2\}} B(d_i^*, \varepsilon_2) \) with sufficiently small constant \( \varepsilon_2 < 1 \) mentioned in (16).

We consider two cases: \( u \neq \frac{2}{3} \) and \( u = \frac{2}{3} \).

The case of \( u \neq \frac{2}{3} \): A straightforward calculation of the Jacobian matrix implies that both \( d_1^* \) and \( d_2^* \) satisfy the conditions (A1) and (A3) of Assumption 2.11. Moreover, the condition of Proposition 2.12 holds (see Table 1). Therefore, we can apply Proposition 2.12.

Suppose that \( u < \frac{2}{3} \). Then the fixed point \( d_2^* \) does not exist and thus we may assume \( \delta^{(0)} \in B(d_1^*, \varepsilon_2) \). From Proposition 2.12, we have \( |\delta_2^{(t)}| > \varepsilon_2 \) for some \( t = O(\log n) \) (note that, here, \( \beta = \delta \) and the eigenvalues satisfy \( 0 \leq \lambda_1 < 1 < \lambda_2 \)).

Suppose that \( u > \frac{2}{3} \). Both eigenvalues of \( J_1 \) are strictly larger than 1. Hence, for \( d_1^* \), Proposition 2.12 implies that either \( |\delta_1^{(t)}| > \varepsilon_2 \) or \( |\delta_2^{(t)}| > \varepsilon_2 \) holds for some \( t = O(\log n) \) if \( \delta^{(0)} \in B(d_1^*, \varepsilon_2) \). If the former holds with \( |\delta_1^{(t)}| = o(1) \), then \( \delta^{(t+T)} \in B(d_1^*, \varepsilon_2) \) holds for some constant \( T = T(\varepsilon_2) \) since \( d_1^* > d_1 \) holds whenever \( 0 < d_1 < \sqrt{\frac{3a_2 - 2}{4t}} \) and \( d_2 = 0 \). Note that, at the point \( d_2^* \), the Jacobian matrix \( J_2 \) has eigenvalues \( \lambda_1, \lambda_2 \) satisfying \( 0 < \lambda_1 < 1 < \lambda_2 \). Moreover, we have \( \beta = \delta - d_2^* \)
from (11). Thus, Proposition 2.12 yields that \(|\delta_2^{(t')}| > \epsilon_2\) holds for some \(t' = O(\log n)\) and any \(\delta^{(0)} \in B(d_2^*, \epsilon_2)\).

**The case of \(u = \frac{2}{3}\):** In this case, \(d_1^* = d_2^* = (0, 0)\). We claim that this point satisfies (B1) and (B2) and then apply Proposition 2.13.

Let \(\epsilon_2 > 0\) be the sufficiently small constant mentioned in (16). The Jacobian matrix \(J = J_2\) has eigenvalues \(1 + 3\frac{2}{3}\). Suppose that \(\|\delta^{(0)}\|_\infty \leq \epsilon_2\) for a sufficiently small constant \(\epsilon_2 > 0\). Then, \(|E[\delta_2^2]| = \left|\frac{2}{3}(3 - 3(u\delta_1)^2 - \delta_2^2)\right| + O(n^{-0.5}) \geq 1.49|\delta_2| - O(n^{-0.5})\). This verifies the assumption (B1). Now we check that the assumption (B2) holds. Note that (B2) is equivalent to

\[
\Pr \left[ |\delta_1^{(t)}| \leq \epsilon_2 \mid |\delta_1| \leq \epsilon_2 \right] \geq 1 - n^{-O(1)}.
\]

For any \(\delta\) of \(|\delta_1| \leq \epsilon_2\),

\[
|E[\delta_1^{(t)}] = \left|\frac{u\delta_1}{2}\right| \left|3 - (u\delta_1)^2 - 3\delta_2^2\right| \pm O(n^{-0.5}) \leq |\delta_1| \left(1 - \frac{4}{27}\delta_1^2\right) + O(n^{-0.5}).
\]

Therefore, from the Hoeffding inequality (Lemma B.9), if \(|\delta_1| \leq \epsilon_2\), it holds w.h.p. that

\[
|\delta_1^{(t)}| \leq |\delta_1| - \frac{4}{27}|\delta_1|^3 + C\sqrt{\frac{\log n}{n}}
\]

for sufficiently large constant \(C > 0\) and large \(n\). If \(|\delta_1|^3 \geq \frac{27C}{4}\sqrt{\log n}\), we have \(|\delta_1^{(t)}| \leq |\delta_1| \leq \epsilon_2\) holds w.h.p. If \(|\delta_1|^3 < \frac{27C}{4}\sqrt{\log n}\), we have \(|\delta_1^{(t)}| = O\left(\sqrt{\frac{\log n}{n}}\right) \leq \epsilon_2\) holds w.h.p.

Thus, from Proposition 2.12, we have \(|\delta_2^{(t)}| > \epsilon_2\) w.h.p. for some \(t = O(\log n)\). This completes the proof of Proposition 3.5.

Here, we prove Theorems 1.2 and 1.4 using Propositions 3.2 to 3.5.

**Proof of Theorem 1.2.** From Theorem 2.3, \(G(2n, p, q)\) is \(f^{Bo3}\)-good. If \(r > \frac{1}{2}\) and \(A^{(0)} \subseteq V\) satisfies \(|A^{(0)}| - n = \Omega(n)\), then we have \(|d_2^{(0)}| = |\delta_2^{(0)}| > \kappa\) for some constant \(\kappa > 0\). Next, for any constant \(\epsilon > 0\), Proposition 3.2 implies \(\langle d^{(l)} \rangle_+ \in B(d_1^*, \epsilon)\) for some constant \(l = l(\epsilon)\). From (9), we have \(\langle \delta^{(l)} \rangle_+ \in B(d_1^*, \epsilon)\) for sufficiently large \(n\). Let \(r\) be the constant mentioned in Proposition 3.4. Then, from Proposition 3.4, it holds w.h.p. that \(T_{cons}(A^{(0)}) \leq l + T_{cons}(A^{(0)}) \leq O(\log \log n + \log n/\log(np))\).

If \(r < \frac{1}{2}\), Proposition 3.3 yields \(T_{cons}(A^{(0)}) \geq \exp(\Omega(n))\) w.h.p. for any \(A^{(0)} \subseteq V\) with \(\delta^{(0)} \in B(d_2^*, \epsilon)\), where \(\epsilon > 0\) is the constant from Proposition 3.3. This completes the proof of (ii).

**Remark: for \(r = r(n)\).** The proof above assumes that \(r = q/p\) is a constant. In fact, our proofs also work for the case \(r = r(n)\), with small additional discussions. See Appendix A for details.

**Proof of Theorem 1.4.** From Theorem 2.3, \(G(2n, p, q)\) is both \(f^{Bo3}\)- and \(f^{Bo3}(1 - f^{Bo3})\)-good. If \(|\delta^{(0)}| = o(1)\), then Proposition 3.5 yields that \(|\delta^{(0)}| > \kappa\) for some constant \(\kappa > 0\) and some \(t = O(\log n)\). Then, from Theorem 1.2, we have \(T_{cons}(A^{(0)}) \leq O(\log \log n + \log n/\log(np))\). Thus, \(T_{cons}(A^{(0)}) \leq t + T_{cons}(A^{(0)}) \leq O(\log n)\).
A straightforward calculation yields that $s(\delta) \geq s(\delta')$ and $|s(\delta') - s(\delta)| = o(1)$.

Consider the Best-of-two on an $f^2_1$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = q/p < \sqrt{5} - 2$ is a constant. Then, there is a constant $\epsilon = \epsilon(r) > 0$ satisfying

$$\Pr \left[ \delta' \not\in B(d^*_2, \epsilon) \right| \delta \in B(d^*_2, \epsilon) \leq \exp(-\Omega(n)).$$

In particular, $T_{\text{cons}}(A) = \exp(\Omega(n))$ holds w.h.p. for any $A \subseteq V$ satisfying $(\delta)_+ \in B(d^*_2, \epsilon)$.

Consider the Best-of-two on an $f^3_1$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = q/p > \sqrt{5} - 2$ and $|\delta_{02}^0| = o(1)$, then it holds w.h.p. that $|\delta_{02}^0| > \kappa$ for some $t = O(\log n)$ and some constant $\kappa > 0$. 

### Table 2

Each $(c_1, c_2)$ represents the property of the eigenvalues $\lambda_1 \geq \lambda_2$ of the corresponding Jacobian matrix. In particular, each $c_i$ represents the sign of $\lambda_i - 1$.

| Points | $0 < u < \frac{1}{2}$ | $\frac{1}{2} < u < \frac{\sqrt{5} - 1}{2}$ | $u = \frac{\sqrt{5} - 1}{2}$ | $\frac{\sqrt{5} - 1}{2} < u \leq 1$ |
|--------|----------------------|--------------------------|--------------------------|---------------------------------|
| $d_1^+$ | $(+, -)$             | $(+, 1)$                  | $(+, +)$                  | $(+, +)$                        |
| $d_2^+$ | $\text{Undefined}$   | $(+, 1)$                  | $(+, -)$                  | $(-, -)$                        |
| $d_3^+$ | $\text{Undefined}$   | $\text{Undefined}$        | $(1, -)$                  | $(+, +)$                        |
| $d_4^+$ | $(-, -)$             | $(-, -)$                  | $(-, -)$                  | $(-, -)$                        |

### 3.2 Best-of-two

The induced dynamical system (13) of the Best-of-two has the same form as that of the Best-of-three. A straightforward calculation yields that $d' = d \in S$ holds if and only if $d \in \{d_i^*\}_{i=1}^4$, where

$$d_i^* := \begin{cases} (0, 0) & \text{if } i = 1, \\ \left( \frac{2u - 1}{u^2}, 0 \right) & \text{if } i = 2 \text{ and } u \geq \frac{1}{2}, \\ \left( \frac{u^2 + u - 1}{(u+1)^2}, \frac{1}{u(u+1)^2} \right) & \text{if } i = 3 \text{ and } u \geq \frac{5 - \sqrt{5}}{2}, \\ (0, 1) & \text{if } i = 4. \end{cases}$$

The Jacobian matrix $J$ at $(d_1, d_2)$ is

$$J = \frac{1}{2} \begin{pmatrix} 2u + 1 - 3(u d_1)^2 - (2u + 1)d_2^2 & -2(2u + 1)d_1 d_2 \\ -2u(u + 2)d_1 d_2 & 3 - u(2 + u)d_1^2 - 3d_2^2 \end{pmatrix}. \tag{19}$$

See Table 2 for the eigenvalues of $J$ at each $d_i^*$.

**Proposition 3.6.** For any sequence $(d_i^{(0)})_{i=0}^\infty$, $\lim_{i \to \infty} (d_i^{(0)})_+ = d_i^*$ for some $i \in \{1, 2, 3, 4\}$. Furthermore, if $u < \frac{\sqrt{5} - 1}{2}$ and there is a positive constant $\kappa > 0$ such that the initial point $d_i^{(0)} = (d_1^{(0)}, d_2^{(0)}) \in S$ satisfies $|d_2^{(0)}| > \kappa$, then $\lim_{i \to \infty} (d_i^{(0)})_+ = d_i^*$.

**Proposition 3.7.** Consider the Best-of-two on an $f^2_1$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = q/p < \sqrt{5} - 2$ is a constant. Then, there is a constant $\epsilon = \epsilon(r) > 0$ satisfying

$$\Pr \left[ \delta' \not\in B(d^*_2, \epsilon) \right| \delta \in B(d^*_2, \epsilon) \leq \exp(-\Omega(n)).$$

In particular, $T_{\text{cons}}(A) = \exp(\Omega(n))$ holds w.h.p. for any $A \subseteq V$ satisfying $(\delta)_+ \in B(d^*_2, \epsilon)$.

**Proposition 3.8.** Consider the Best-of-two on an $f^3_1$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = q/p$ is a constant. Then, for some constant $\epsilon = \epsilon(r) > 0$, $T_{\text{cons}}(A) \leq O(\log \log n + \log n/ \log(np))$ holds w.h.p. for any $A \subseteq V$ satisfying $\min(|A|, 2n-|A|) \leq \epsilon n$.

**Proposition 3.9.** Consider the Best-of-two on an $f^2_1$-good graph for a partition $(V_1, V_2)$ and parameters $p, q$ such that $r = q/p > \sqrt{5} - 2$ and $|\delta_{02}^{(0)}| = o(1)$, then it holds w.h.p. that $|\delta_{02}^{(t)}| > \kappa$ for some $t = O(\log n)$ and some constant $\kappa > 0$. 


We omit proofs of Propositions 3.6 to 3.8 since they are substantially the same as those of Propositions 3.2 to 3.4.

**Proof of Theorems 1.3 and 1.5.** The proof of Theorems 1.3 and 1.5 is the same as that of Theorems 1.2 and 1.4, respectively, except for the threshold and using Propositions 3.6 to 3.8 instead of Propositions 3.2 to 3.4.

### 4 PROOF OF THE $F$-GOODNESS OF THE STOCHASTIC BLOCK MODEL (THEOREM 2.3)

In this section we show Theorem 2.3. In Section 4.1, we show that the property (P2) can be obtained from Lemma 4.1.

**Lemma 4.1.** For a vertex set $V$ with $|V|=N$, let $(I_e)_{e \in \binom{V}{2}}$ be independent binary random variables. Let $p := \max_{e \in \binom{V}{2}} E[I_e]$. Suppose that $Np \geq 1$. For $\ell + 1$ vertex subsets $S_0, S_1, \ldots, S_\ell$, let

$$W(S_0; S_1, \ldots, S_\ell) := \prod_{s \in S_0} \deg_{S_i}(s),$$

$$\hat{W}(S_0; S_1, \ldots, S_\ell) := \prod_{s \in S_0} E[\deg_{S_i}(s)],$$

where $\deg_S(v) = \sum_{s \in S \setminus \{v\}} f_{\{v,s\}}$ for $S \subseteq V$ and $v \in V$.

Then there are two positive constants $C_1$ and $C_2$ depending only on $\ell$ such that the following holds with probability $1 - N^{-C_1}$:

$$\forall S_0, S_1, \ldots, S_\ell : \left| W(S_0; S_1, \ldots, S_\ell) - \hat{W}(S_0; S_1, \ldots, S_\ell) \right| \leq C_2 N(Np)^{\ell-1/2}.$$

Our proof of Lemma 4.1 consists of three parts. First, we give a concentration of $W$ (Lemma 4.3). Next, we give an upper bound on the discrepancy between $E[W]$ and $\hat{W}$ (Lemma 4.4). Finally, we present Lemma 4.2 playing a key role in showing Lemmas 4.3 and 4.4.

#### 4.1 Reduction to $W$

**Proof of (P2) of Theorem 2.3 via Lemma 4.1.** Let $f(x) = \sum_{j=0}^{\ell} c_j x^j$. For notational convenience, let $x_v = \frac{\deg_S(v)}{\deg(v)}$, $\bar{x}_v = E[\deg_{S}(v)]/E[\deg(v)]$, and $\hat{x} = \frac{A_0 p + A_1 q}{n(p+q)}$. Then from the triangle inequality, it holds that

$$\left| \sum_{v \in S \cap V_i} (f(x_v) - f(\hat{x})) \right| \leq \sum_{v \in S \cap V_i} (f(x_v) - f(\bar{x}_v)) + \sum_{v \in S \cap V_i} (f(\bar{x}_v) - f(\hat{x})) \right| \leq C_2 |S \cap V_i|/n \leq C_2.$$

For the second term of the right side of (20), there are two positive constants $C_1$ and $C_2$ such that

$$\left| \sum_{v \in S \cap V_i} (f(\bar{x}_v) - f(\hat{x})) \right| \leq \sum_{v \in S \cap V_i} |f(\bar{x}_v) - f(\hat{x})| \leq C_1 \sum_{v \in S \cap V_i} |\bar{x}_v - \hat{x}| \leq C_2.$$
The second inequality follows from the Lipschitz condition of $f$ (c.f. Appendix B.2). The third inequality holds since $\mathbf{E}[\deg(v)] = (n - 1)p + nq$ and $|A_i| - 1)p + |A_{3-i}|q \leq \mathbf{E}[\deg_A(v)] \leq |A_i|p + |A_{3-i}|q$ for any $v \in V_i$.

For the first term of the right side of (20), since

$$
\left( \frac{\deg_A(v)}{\deg(v)} \right)^j - \left( \frac{\mathbf{E}[\deg_A(v)]}{\mathbf{E}[\deg(v)]} \right)^j = \frac{\left( \mathbf{E}[\deg(v)]^j - \deg(v)^j \right) \left( \frac{\deg_A(v)}{\deg(v)} \right)^j + \left( \deg_A(v)^j - \mathbf{E}[\deg_A(v)]^j \right)}{\mathbf{E}[\deg(v)]^j}
$$

for any $j$ and $v \in V$, we have

$$
\left| \sum_{v \in S \cap V_i} (f(x_v) - f(\bar{x}_v)) \right| = \left| \sum_{j=1}^\ell c_j \sum_{v \in S \cap V_i} (x_v)^j - (\bar{x}_v)^j \right| 
\leq \sum_{j=1}^\ell \frac{|c_j|}{(|v| - 1)p^j} \left( \sum_{v \in S \cap V_i} \left( \mathbf{E}[\deg(v)]^j - \deg(v)^j \right) (x_v)^j + \sum_{v \in S \cap V_i} \left( \deg_A(v)^j - \mathbf{E}[\deg_A(v)]^j \right) \right).
$$

Note that $\mathbf{E}[\deg(v)] = (n - 1)p + nq \geq (n - 1)p$ for any $v \in V$. Since

$$
\left| \sum_{v \in S \cap V_i} \left( \mathbf{E}[\deg(v)]^j - \deg(v)^j \right) (x_v)^j \right| \leq \max_{U \subseteq V} \left| \sum_{u \in U} \left( \mathbf{E}[\deg(u)]^j - \deg(u)^j \right) \right| 
\leq \max_{U \subseteq V} \left| \mathbf{W}(U; V, \ldots, V) - \mathbf{W}(U; V, \ldots, V) \right|
$$

and $\sum_{v \in S \cap V_i} \left( \deg_A(v)^j - \mathbf{E}[\deg_A(v)]^j \right) = W(S \cap V; A, \ldots, A) - \mathbf{W}(S \cap V; A, \ldots, A)$, we obtain the claim from Lemma 4.1. Note that, for any $S \subseteq V$, $a \in \mathbb{R}^V$ and $x \in [0, 1]^V$, $\left| \sum_{s \subseteq a} a_s x_s \right| \leq \max_{U \subseteq V} \left| \sum_{u \in U} a_u \right|$ since $\sum_{s \subseteq a, s \subseteq a} a_s x_s \leq \sum_{s \subseteq a} a_s x_s \leq \sum_{s \subseteq a} a_s x_s$.

Lemma 4.2 will be used in Sections 4.2 and 4.3.

**Lemma 4.2.** Let $V$ be a set of $|V|=N$ vertices and fix $l+1$ subsets $S_0, S_1, \ldots, S_l \subseteq V$. For any $s = (s_0, s_1, \ldots, s_l) \in S_0 \times S_1 \times \cdots \times S_l$, define

$$
U(s) := \{ s_i : i \in \{0\} \cup [l] \}.
$$

Consider $\sum_{s \in S} p^{\left| F(s) \right|}$, where $p \in [1/N, 1]$, $S \subseteq S_0 \times S_1 \times \cdots \times S_l$ and

$$
F : S_0 \times S_1 \times \cdots \times S_l \rightarrow \binom{l}{2}.
$$

Suppose that the following three conditions hold for any $s \in S$: (1) $\left| F(s) \right| \leq k$, (2) $F(s) \subseteq \binom{U(s)}{2}$, and (3) the graph $G(s) = (U(s), F(s))$ is connected. Let $L \subseteq \{0\} \cup [l]$ be a set of indices such that $S_i \cap S_j = \emptyset$ for any $i, j \in L (i \neq j)$. Then

$$
\sum_{s \in S} p^{\left| F(s) \right|} \leq B_{l+1} N(Np)^k \prod_{i \in L} |S_i|^{N/|L|},
$$

where $B_l$ denotes the $l$th Bell number.
The $l$th Bell number $B_l$ is the number of possible partitions of a set with $l$ labeled elements. It is known that $B_l < \left(\frac{0.7921}{\ln(l+1)}\right)^l$ for all positive integer $l$ [20].

\[ \text{Lemma 4.3.} \quad \text{Consider the same setting as that of Lemma 4.1. Then, there are two positive constants } C_1 \text{ and } C_2 \text{ depending only on } \ell \text{ such that the following holds with probability } 1 - N^{-C_1}: \]

\[ \forall S_0, S_1, \ldots, S_\ell : |W(S_0; S_1, \ldots, S_\ell) - E[W(S_0; S_1, \ldots, S_\ell)]| \leq C_2 N(Np)^{\ell-1/2}. \]

\[ \text{Proof.} \quad \text{For } \ell + 1 \text{ vertex subsets } S_0, S_1, \ldots, S_\ell, \text{ let} \]

\[ S := \{(s_0, s_1, \ldots, s_\ell) : s_0 \in S_0, s_i \in S_i \setminus \{s_0\} \text{ for every } i \in [\ell]\} \]

and for any $s = (s_0, s_1, \ldots, s_\ell) \in \prod_{i=0}^\ell S_i = S_0 \times S_1 \times \cdots \times S_\ell$, let

\[ F(s) := \{ (s_0, s_i) : i \in [\ell] \} \setminus \{s_0\}. \]

For example, $F((a, b, a, c, d, b, f, a)) = \{(a, b), (a, c), (a, d), \{a, f\}\}$. Then,

\[ W(S_0; S_1, \ldots, S_\ell) = \sum \prod_{s_0 \in S_0} \left( \sum_{s_i \in S_i \setminus \{s_0\}} I_{(s_0, s_i)} \right) = \sum \prod_{s_0 \in S_0} I_{(s_0, s_i)} = \sum \prod_{s_0 \in S_0} I_{s_0}. \quad (21) \]

\[ \text{Lower bound on } W. \text{ First, we claim the following statement: there are two positive constants } C_3 \text{ and } C_4 \text{ such that} \]

\[ \Pr[\forall S_0, S_1, \ldots, S_\ell : W(S_0; S_1, \ldots, S_\ell) \geq E[W(S_0; S_1, \ldots, S_\ell)] - C_4 N(Np)^{\ell-1/2}] \geq 1 - N^{-C_3}. \quad (22) \]

Janson’s inequality (Lemma B.15) can be applied to (21) to obtain (22). Then,

\[ \Pr[\exists S_0, S_1, \ldots, S_\ell : W(S_0; S_1, \ldots, S_\ell) \leq E[W(S_0; S_1, \ldots, S_\ell)] - l] \leq (2^N)^{\ell+1} \exp\left(-\frac{l^2}{2V(S_0; S_1, \ldots, S_\ell)}\right) \leq \exp\left((\ell + 1)N - \frac{l^2}{2V(S_0; S_1, \ldots, S_\ell)}\right), \quad (23) \]

where

\[ V(S_0; S_1, \ldots, S_\ell) = \sum_{s_0 \in S_0} \mathbb{E}\left[ \prod_{s_0 \in S_0} I_{s_0} \prod_{s_0 \in S_0} I_{s_0} \right]. \]

Thus, it is sufficient to show that $V(S_0; S_1, \ldots, S_\ell) = O(N(Np)^{2\ell-1})$. Since $\max_{e \in \{\ell\}} E[I_e] = p$, it holds that

\[ V(S_0; S_1, \ldots, S_\ell) = \sum_{s_0 \in S_0} \mathbb{E}\left[ \prod_{s_0 \in S_0} I_{s_0} \prod_{s_0 \in S_0} I_{s_0} \right] \leq \sum_{s_0 \in S_0} p^{\mid F(s) \cup F(s') \mid}. \quad (24) \]
Lemma 4.2 can be applied to bound (24). The proof of Lemma 4.2 is provided in Section 4.4. Consider $2\ell + 2$ vertex subsets $S', S_1', \ldots, S_{2\ell + 1}'$ where $S_i' := S_{i \text{ mod } (\ell + 1)}$. For any $i \in \{0\} \cup \{2\ell + 1\}$, let

$$S := \{(s_0, s_1, \ldots, s_{2\ell + 1}) \in S \times S : F((s_0, \ldots, s_{\ell})) \cap F((s_{\ell + 1}, \ldots, s_{2\ell + 1})) \neq \emptyset\} \subseteq \prod_{i=0}^{2\ell + 1} S_i',$$

$$F(s) := F((s_0, \ldots, s_{\ell})) \cup F((s_{\ell + 1}, \ldots, s_{2\ell + 1})) \text{ for any } s = (s_0, s_1, \ldots, s_{2\ell + 1}) \in \prod_{i=0}^{2\ell + 1} S_i'.$$

Then for any $s \in S$, $G(s) = (U(s), F(s))$ is a connected graph and $|F(s)| \leq 2\ell - 1$. Thus, the following can be obtained for any $i_* \in \{0\} \cup \{\ell\}$ by applying Lemma 4.2 with $l = 2\ell + 1, k = 2\ell - 1$ and $L = \{i_*\}$:

$$\sum_{s \in S} p^{[F(s)\cup F(s')]\setminus F(s')} = \sum_{s \in S} p^{[F(s)]} \leq B_{2(\ell+1)}N(Np)^{2\ell-1} \frac{|S_{i_*}|}{N} = B_{2(\ell+1)}|S_{i_*}|(Np)^{2\ell-1}.$$  (25)

Equations (24) and (25) imply the following statement: for any $\ell + 1$ vertex subsets $S_0, S_1, \ldots, S_{\ell}$ and any $i_* \in \{0\} \cup \{\ell\}$,

$$\nabla(S_0; S_1, \ldots, S_{\ell}) \leq B_{2(\ell+1)}|S_{i_*}|(Np)^{2\ell-1} \leq B_{2(\ell+1)}N(Np)^{2\ell-1}.$$  (26)

Thus, the statement formulated in (22) can be obtained by substituting $l = C_4N(Np)^{\ell-1/2}$ with $C_4 = \sqrt{2(\ell + 1 + C_3)}$ to (23).

**Upper bound on $W$.** To complete the proof of Lemma 4.3, we combine statement (22) and the following claim: there are two positive constants $C_5$ and $C_6$ such that

$$\Pr\left[\forall S_0, S_1, \ldots, S_{\ell} : W(S_0; S_1, \ldots, S_{\ell}) \leq E[W(S_0; S_1, \ldots, S_{\ell})] + C_6N(Np)^{\ell-1/2}\right] \geq 1 - N^{-C_5}.$$  (27)

To show the claim, we consider the following expression of $W$. For any $S_0, S_1, \ldots, S_{\ell}$, let $W_0 := W(S_0; S_1, \ldots, S_{\ell})$ and let $W_i := W(V; V, \ldots, V, S_i, S_{i+1}, \ldots, S_{\ell})$ for each $i \in \{\ell + 1\}$. Since $W_{i+1} - W_i = W(V; V, \ldots, V, V\backslash S_i, S_{i+1}, \ldots, S_{\ell})$ for any $i \in \{0\} \cup \{\ell + 1\}$ and $\sum_{i=0}^{\ell}(W_{i+1} - W_i) = W_{\ell+1} - W_0$, we have

$$W(S_0; S_1, \ldots, S_{\ell}) = W(V; V, \ldots, V) - \sum_{i=0}^{\ell} W(V; V, \ldots, V, V\backslash S_i, S_{i+1}, \ldots, S_{\ell}).$$  (28)

We can apply (22) to the second term of the right side of (28). Now, we try to get an upper bound on $W(V; V, \ldots, V)$. For the notational convenience, let $Y = W(V; V, \ldots, V)$. Let $S_i = V$ for every $i \in \{0\} \cup \{\ell\}$ and let

$$E := \{F(s) : s \in S\}.$$
From (21), we have

\[ Y = \sum_{s \in S} \prod_{e \in F(s)} I_e = \sum_{F \in \mathcal{E}} \left| \{ s \in S : F(s) = F \} \right| \prod_{e \in F} I_e. \]

Thus, the following can be obtained by applying the Kim-Vu inequality (Lemma B.16) to \( Y \):

\[
\Pr \left[ |Y - E[Y]| \geq \sqrt{\ell! \max_{A \subseteq \binom{V}{2}} E[Y_A] \max_{A \subseteq \binom{V}{2}: A \neq \emptyset} E[Y_A] (8\lambda)^{\ell} \right] \leq 2 \exp(2 + 2(\ell - 1) \log N - \lambda),
\]

where

\[ Y_A = \sum_{F \in \mathcal{E}: F_{\sigma} \supseteq A} \left| \{ s \in S : F(s) = F \} \right| \prod_{e \in F \setminus A} I_e = \sum_{s \in S: F(s) \supseteq A} \prod_{e \in F(s) \setminus A} I_e. \]

Now, we give an upper bound on \( E[Y_A] \). Since \( \max_{e \in \binom{V}{2}} E[I_e] = p \), it holds that

\[ E[Y_A] = \sum_{s \in S: F(s) \supseteq A} \left( \prod_{e \in F(s) \setminus A} I_e \right) \leq \sum_{s \in S: F(s) \supseteq A} p^{\left| F(s) \setminus A \right|} = \sum_{s \in S: F(s) \supseteq A} p^{\left| F(s) \setminus |A| \right|}. \]

If \( A = \emptyset \), the following can be obtained by applying Lemma 4.2 with \( l = k = \ell \) and \( L = \emptyset \):

\[ E[Y_A] = E[Y] \leq \sum_{s \in S} p^{\left| F(s) \right|} \leq B_{\ell+1} N(Np)^\ell. \quad (30) \]

Note that \( |F(s)| \leq \ell \) and \( G(s) = (U(s), F(s)) \) is a connected graph for any \( s \in S \subseteq \prod_{i=0}^{\ell} S_i \).

Now we consider the case \( |A| = \kappa \geq 1 \). Let \( V(A) \) be the set of vertices induced by the edge set \( A \subseteq \binom{V}{2} \). If \( F(s) \supseteq A \) for some \( s \in \prod_{i=0}^{\ell} V \), the graph \( G' = (V(A), A) \) is a star graph; hence, \( |V(A)| = |A| + 1 = \kappa + 1 \). Let \( V(A) = \{ a_0, a_1, \ldots, a_\kappa \} \). Now consider \( (\ell + 1) + (\kappa + 1) \) vertex subsets \( S'_0, S'_1, \ldots, S'_{\ell + \kappa + 1} \), where \( S'_i = S_i \) for any \( 0 \leq i \leq \ell' \) and \( S'_i = \{ a_{i-(\ell+1)} \} \) for any \( \ell' + 1 \leq i \leq \ell + \kappa + 1 \). Let

\[ S := \left\{ (s_0, s_1, \ldots, s_{\ell + \kappa + 1}) \in S \times \prod_{i=0}^{\kappa} \{ a_i \} : F((s_0, \ldots, s_\ell)) \supseteq A \right\} \subseteq \prod_{i=0}^{\ell + \kappa + 1} S'_i. \]

\[ F(s) := F((s_0, \ldots, s_\ell)) \text{ for any } s = (s_0, s_1, \ldots, s_{\ell + \kappa + 1}) \in \prod_{i=0}^{\ell + \kappa + 1} S'_i. \]

Note that for any \( s \in S \), the graph \( G(s) = (U(s), F(s)) \) is connected and \( |F(s)| \leq \ell \). Thus, the following can be obtained by applying Lemma 4.2 with \( l = \ell + \kappa + 1, k = \ell \) and \( L = \{ \ell + 1, \ell + 2, \ldots, \ell + \kappa + 1 \} \) (note that \( a_i \neq a_j \) for any \( i \neq j \) and \( \prod_{i=\ell+1}^{\ell+\kappa+1} |S'_i| = \prod_{i=\ell+1}^{\ell+\kappa+1} |\{ a_{i-(\ell+1)} \}| = 1 \):

\[ E[Y_A] \leq \sum_{s \in S: F(s) \supseteq A} p^{\left| F(s) \setminus |A| \right|} \leq \frac{1}{p^\kappa} \sum_{s \in S} p^{\left| F(s) \right|} \leq \frac{1}{p^\kappa} B_{\ell+\kappa+2} N(Np)^\ell \sum_{i=\ell+1}^{\ell+\kappa+1} |S'_i|^{\ell+\kappa+1} \leq B_{2(\ell+1)}(Np)^{\ell - \kappa}. \quad (31) \]
Combining (30) and (31), we have
\[
\max_{A \subseteq \binom{V}{2} : |A| \geq 1} \mathbb{E}[Y_A] \leq \max_{A \subseteq \binom{V}{2} : |A| \geq 1} B_{2(\ell+1)}(Np)^{\ell-1} = B_{2(\ell+1)}(Np)^{\ell-1},
\]
\[
\max_{A \subseteq \binom{V}{2}} \mathbb{E}[Y_A] = \max_{A \subseteq \binom{V}{2} : |A| \geq 1} \mathbb{E}[Y_A], \mathbb{E}[d_\ell(V)] \leq B_{2(\ell+1)}N(Np)^{\ell}.
\]

Thus, from (29) with \( \lambda = (2(\ell - 1) + C_7/2) \log N \) and \( C_8 = \sqrt{\lambda}B_{2(\ell+1)}(16(\ell - 1 + C_7/2))^{\ell} \), we can obtain
\[
\Pr[|Y - \mathbb{E}[Y]| \geq C_8 \sqrt{N}(\log N)^{\ell}(Np)^{\ell-1/2}] \leq 2e^2/N^{C_7}. \tag{32}
\]

Combining (28), (22), and (32), the following holds with probability at least \( 1 - 2e^2/N^{C_7} - 1/N^{C_7} \):
\[
\forall S_0, S_1, \ldots, S_{\ell} : 
W(S_0; S_1, \ldots, S_{\ell}) \leq \mathbb{E}[W(S_0; S_1, \ldots, S_{\ell})] + C_9 \sqrt{N}(\log N)^{\ell}(Np)^{\ell-1/2} + (\ell + 1)C_4N(Np)^{\ell-1/2}.
\]

This supports statement (27). Combining statements (22) and (27) completes the proof of Lemma 4.3.

\[
\square
\]

4.3 | Discrepancy between the expected value and the ideal value

**Lemma 4.4.** Consider the settings of Lemma 4.1. Then, for any vertex subsets \( S_0, S_1, \ldots, S_{\ell} \) and any \( i_0 \in \{0\} \cup \{\ell\} \), there is a positive constant \( C \) such that
\[
\left| \mathbb{E}[W(S_0; S_1, \ldots, S_{\ell})] - \hat{W}(S_0; S_1, \ldots, S_{\ell}) \right| \leq C|S_{i_0}|(Np)^{\ell-1}.
\]

**Proof of Lemma 4.4.** We show that
\[
\sum_{s \in S} \prod_{e \in [\ell]} \mathbb{E}[\deg_{S_i}(s)] \leq \mathbb{E}[W(S_0; S_1, \ldots, S_{\ell})] \leq \sum_{s \in S} \prod_{e \in [\ell]} \mathbb{E}[\deg_{S_i}(s)] + B_{\ell+1}|S_{i_0}|(Np)^{\ell-1}
\]
for any \( i_0 \in \{0\} \cup \{\ell\} \). The first inequality follows directly from the FKG inequality (Lemma B.14) since \( \deg_{S_i}(s) \) is a monotone increase function on \( (I_e)_{e \in \binom{V}{2}} \) for every \( i \). Now we show the second inequality. We write each element \( s \in S \) as \( s = (s_0, s_1, \ldots, s_{\ell}) \). Since \( \mathbb{E}[W(S_0; S_1, \ldots, S_{\ell})] = \sum_{s \in S} \mathbb{E}\left[ \prod_{e \in [\ell]} I_{\{s_0, s_1\}} \right] \), we have
\[
\mathbb{E}[W(S_0; S_1, \ldots, S_{\ell})] = \sum_{s \in S} \mathbb{E}\left[ \prod_{e \in [\ell]} I_{\{s_0, s_1\}} \right] + \sum_{s \in S} \mathbb{E}\left[ \prod_{e \in [\ell]} I_{\{s_0, s_1\}} \right].
\]
For the first term, since \( s_i \neq s_j \) for any \( i, j \in \mathcal{E} \) \((i \neq j)\) if \(|F(s)|=\mathcal{E}\), we obtain
\[
\sum_{s \in S: |F(s)|=\mathcal{E}} \mathbf{E} \left[ \prod_{e \in \mathcal{E}} I_{\{s_{t(e)}\}} \right] = \sum_{s \in S: |F(s)|=\mathcal{E}} \prod_{e \in \mathcal{E}} \mathbf{E} \left[ I_{\{s_{t(e)}\}} \right] 
\leq \sum_{s \in S: |F(s)|=\mathcal{E}} \prod_{e \in \mathcal{E}} \mathbf{E} \left[ I_{\{s_{t(e)}\}} \right] = \sum_{s \in S: |F(s)|=\mathcal{E}} \mathbf{E} \left[ \deg \{s\} \right].
\]

For the second term, from Lemma 4.2,
\[
\sum_{s \in S: |F(s)| \leq \mathcal{E}-1} \mathbf{E} \left[ \prod_{e \in F(s)} I_e \right] \leq \sum_{s \in S: |F(s)| \leq \mathcal{E}-1} p^{|F(s)|} \leq B_{\mathcal{E}+1} |S| (Np)^{\mathcal{E}-1}.
\]

Note that \( G(s) = (U(s), F(s)) \) is a connected graph for any \( s \in S \).

4.4 Proof of Lemma 4.2

To complete the proof of Lemma 4.1, we show Lemma 4.2 in this section.

**Proof of Lemma 4.2.** It is easy to see that
\[
\forall s \in S : |U(s)| - 1 \leq |F(s)| \leq k
\]
since \( G(s) = (U(s), F(s)) \) is a connected graph from the assumption. Hence,
\[
\sum_{s \in S} p^{|F(s)|} \leq \sum_{s \in S} p^{|U(s)|-1} = \sum_{s \in S: |U(s)| \leq k+1} p^{|U(s)|-1} \leq \sum_{s \in \prod_{i=0}^{\mathcal{E}} S_i : |U(s)| \leq k+1} p^{|U(s)|-1}. \tag{33}
\]

To estimate above, we introduce the following notations. For any \((l+1)\)-dimensional vector \( s = (s_0, s_1, \ldots, s_l) \in S_0 \times S_1 \times \cdots \times S_l \), let
\[
R(s) := \{ \{j \in \{0\} \cup \{I\} : s_j = s_i\} : i \in \{0\} \cup \{I\} \}.
\]
For example, \( R((a, b, a, c, d, b, f, a)) = \{\{0, 2, 7\}, \{1, 5\}, \{3\}, \{4\}, \{6\}\} \). Note that \( R(s) \) is a partition of \( \{0\} \cup \{I\} \). From the definition, we have \(|R(s)| \leq |U(s)|\). For example, \(|U((a, b, a, c, d, b, f, a))| = |\{a, b, c, d, f\}| = 5 = |R((a, b, a, c, d, b, f, a))| \). Let \( R_i \) be the family of all partitions of \( \{0\} \cup \{I\} \). For example,
\[
R_2 = \{\{0\}, \{1\}, \{2\}\}, \{\{0\}, \{1, 2\}\}, \{\{1\}, \{0, 2\}\}, \{\{2\}, \{0, 1\}\}, \{\{0, 1, 2\}\}\}.
\]

Note that \(|R_i| = B_{i+1}\). Then,
\[
\sum_{s \in \prod_{i=0}^{\mathcal{E}} S_i : |U(s)| \leq k+1} p^{|U(s)|} = \sum_{R \in R_i} \sum_{s \in \prod_{i=0}^{\mathcal{E}} S_i : |R|=R} p^{|U(s)|} 
= \sum_{R \in R_i} p^{|R|} \left\{ s \in \prod_{i=0}^{\mathcal{E}} S_i : R(s) = R \right\}. \tag{34}
\]
From the definition of $R(s)$, for any $r \in R(s)$, $s_i = s_j$ for any $i, j \in r$. Thus,

$$\left\{ s \in \prod_{i=0}^{l} S_i : R(s) = R \right\} = \sum_{s \in \prod_{i \in S_i : R(s) = R}^{l} 1 \leq \sum_{s_0 \in S_0} \sum_{s_1 \in S_1} \cdots \sum_{s_l \in S_l} \prod_{i \in R} \prod_{i \in r} \prod_{i \in S_i} I_{s_i = s_j} \leq \prod_{i \in R} \prod_{i \in r} \prod_{i \in S_i} \right| \prod_{i \in S_i} \right| R \right| \right.$$ (35)

For example, consider four vertex subsets: $S_0, S_1, S_2, S_3$. Let $R = \{\{0, 1\}, \{2\}, \{3\}\} \in R_3$ and $l = \{i^*\} \subseteq \{0\} \cup \{3\}$ where $i^* \in \{0\} \cup \{3\}$. Then (35) means that

$$\left\{ s \in \prod_{i=0}^{3} S_i : R(s) = R \right\} = \left| \{(s_0, s_1, s_2, s_3) \in S_0 \times S_1 \times S_2 \times S_3 : s_0 = s_1, s_0 \neq s_2, s_0 \neq s_3, s_2 \neq s_3\} \right| \leq \sum_{s_0 \in S_0} \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} \sum_{s_3 \in S_3} I_{s_0 = s_1} \leq |S_0 \cap S_1||S_2||S_3| = \prod_{i \in \{0, 1\} \cup \{2\}} \prod_{i \in r} \prod_{i \in S_i} \right| \prod_{i \in S_i} \right| R \right| \right.$$ (35)

For an index $i \in \{0\} \cup \{3\}$, let $r_i$ be the element of $R$ such that $r_i \ni i$. Now let us consider the set $L$ described in the statement (of Lemma 4.2). First we assume that there are $i, j \in L$ with $i \neq j$ such that both $i$ and $j$ in the same $r = r_i = r_j \in R$. In this case, since $S_i \cap S_j = \emptyset$ from the definition of $L$, we have

$$\prod_{i \in R} \prod_{i \in r_i} \prod_{i \in S_i} = \prod_{i \in L} \prod_{j \in r_j} \prod_{j \in S_j} = 0. \quad (36)$$

Now we assume that $r_i \neq r_j$ for any $i, j \in L$. Since $|\{r_i : i \in L\}| = |L|$ and $R = \{r_i : i \in L\} \cup R \setminus \{r_i : i \in L\}$, we have

$$\prod_{i \in R} \prod_{i \in r_i} \prod_{i \in S_i} = \prod_{i \in L} \prod_{j \in r_j} \prod_{j \in S_j} \leq \left( \prod_{i \in L} |S_i| \right)^N_{|L| - |L|} \quad (37)$$

Finally, by combining (33) to (37), we obtain

$$\sum_{s \in S} p_{[F(s)]} \leq \frac{1}{p} \sum_{R \in R_3 : |R| \leq k+1} p_{[R]} \left| \left\{ s \in \prod_{i=0}^{l} S_i : R(s) = R \right\} \right| \leq \frac{1}{p} \sum_{R \in R_3 : |R| \leq k+1} p_{[R]} N_{|R|} \prod_{i \in L} |S_i| \leq \frac{1}{p} \left( \prod_{i \in L} |S_i| \right) N_{|L|} \left( Np \right)^{k+1} \sum_{R \in R_3 : |R| \leq k+1} 1 \leq |R| \left( \prod_{i \in L} |S_i| \right) N_{|L|} \left( Np \right)^{k} \quad B_{i+1} \left( \prod_{i \in L} |S_i| \right) N_{|L|} \left( Np \right)^{k}.$$

Note that the third inequality follows since $Np \geq 1$ from the assumption.

**Proof of Lemma 4.1.** The proof can be obtained by combining Lemmas 4.3 and 4.4.
**Lemma 4.5.** Suppose that $0 \leq q \leq p = o(\log n / n)$. Then, there are two positive constants $C_1$ and $C_2$ such that $G(2n, p, q)$ satisfies the following with probability $1 - O(n^{-C_1})$:

$$\forall v \in V : |\deg(v) - n(p + q)| \leq C_2 \sqrt{np \log n}.$$  

**Proof.** The following can be obtained by applying the Chernoff bound (Lemma B.8):

$$\Pr \left[ \exists v \in V : |\deg(v) - n(p + q)| > t \right]$$

$$\leq \sum_{v \in V} \left( \exp \left(-\frac{t^2}{3E[\deg(v)]} \right) + \exp \left(-\frac{t^2}{2E[\deg(v)]} \right) \right)$$

$$\leq n \left( 2 \exp \left(-\frac{t^2}{6np} \right) + \exp \left(-\frac{t}{3} \right) \right)$$

$$\leq 2 \exp \left( \log n - \frac{t^2}{6np} \right) + \exp \left( \log n - \frac{t}{3} \right).$$

Note that $E[\deg(v)] = (n - 1)p + nq$ for any $v \in V$ and $E[\deg(v)] \leq n(p + q) \leq 2np$. Thus we obtain the claim with $t = C_2 \sqrt{np \log n}$ since $t = C_2 \sqrt{np \log n} \geq C \log n$ for some constant $C$. $\blacksquare$

**Lemma 4.6.** Suppose that $0 \leq q \leq p = o(\log n / n)$. Let $S(A) = \{ S \cap U : S \in \{A, V \setminus A, V\}, U \in \{V_1, V_2, V\} \}$ for $A \subseteq V$. For any constant $\ell$, there are two positive constants $C_3$ and $C_4$ such that $G(2n, p, q)$ satisfies the following with probability $1 - O(n^{-C_1})$:

$$\forall A \subseteq V, \forall S_0, \ldots, S_{\ell - 1} \in S(A) :$$

$$\left| W(S_0; S_1, \ldots, S_{\ell - 1}, A) - \hat{W}(S_0; S_1, \ldots, S_{\ell - 1}, A) \right| \leq C_2 |A| \sqrt{\log n(p)\ell^{-1/2}}.$$  

**Proof.** **Lower bound.** First we claim the following statement: there are two positive constants $C_3$ and $C_4$ such that the following holds with probability $1 - n^{-C_1}$:

$$\forall A \subseteq V, \forall S_0, \ldots, S_{\ell - 1} \in S(A) :$$

$$W(S_0; S_1, \ldots, S_{\ell - 1}, A) \geq \hat{W}(S_0; S_1, \ldots, S_{\ell - 1}, A) - C_4 |A| \sqrt{\log n(p)\ell^{-1/2}}. \tag{38}$$

The following can be obtained from Janson’s inequality (Lemma B.15) and (26) with a constant $C_5$ and $C_6 = \sqrt{2(C_5 + 1)B_{2(\ell + 1)}}$:

$$\Pr \left[ \exists A \subseteq V, \exists S_0, \ldots, S_{\ell - 1} \in S(A) : W(S_0; S_1, \ldots, S_{\ell - 1}, A) \right]$$

$$\leq E[W(S_0; S_1, \ldots, S_{\ell - 1}, A)] - C_6 |A| \sqrt{\log N(p)\ell^{-1/2}}$$

$$\leq \left( \frac{N}{|A|} \right) |S(A)| \exp \left( -\frac{2(C_5 + 1)B_{2(\ell + 1)}|A| \log N(p)2\ell^{-1}}{2B_{2(\ell + 1)}|A| (Np)^{2\ell-1}} \right)$$

$$\leq 9^\ell \exp \left( |A| \log N - (C_5 + 1)|A| \log N \right) \leq 9^\ell / N^{C_5}. \tag{39}$$

Thus, statement (38) can be supported by combining (39) and Lemma 4.4.
Upper bound. Now we show the following statement: there are two positive constants $C_7$ and $C_8$ such that the following holds with probability $1 - n^{-C_7}$:

$$\forall A \subseteq V, \forall S_0, \ldots, S_{\ell-1} \in S(A) :$$

$$W(S_0; S_1, \ldots, S_{\ell-1}, A) \leq \hat{W}(S_0; S_1, \ldots, S_{\ell-1}, A) + C_8|A| \sqrt{\log np}^{\ell-1/2}. \quad (40)$$

From the discussion of (28),

$$W(S_0; S_1, \ldots, S_{\ell-1}, A) = \sum_{i=0}^{\ell-1} W(V; V, \ldots, V, V \setminus S_i, S_{i+1}, \ldots, S_{\ell-1}, A) \quad (41)$$

since $W_{\ell} = W_0 = \sum_{i=0}^{\ell-1} (W_{i+1} - W_i)$. Thus we consider an upper bound on $W(S_0; S_1, \ldots, S_{\ell-1}, A)$. Let $d_{\text{max}} := \max_{v \in V} \deg(v)$. Since $\sum_{v \in V} \deg_A(v) = \sum_{a \in A} \deg(v)$, we have

$$\sum_{v \in V} \deg(v)^{\ell-1} \deg_A(v) \leq d_{\text{max}}^{\ell-1} \sum_{v \in V} \deg(v) \leq d_{\text{max}}^{\ell-1} \sum_{a \in A} \deg(a) \leq d_{\text{max}}^{\ell} |A|.$$  

From Lemma 4.5, it holds w.h.p. that

$$d_{\text{max}}^{\ell} = \left(n(p + q) + O(\sqrt{np \log n})\right)^{\ell} = (n(p + q))^{\ell} \left(1 + O\left(\sqrt{\frac{\log n}{np}}\right)\right).$$

The second equality holds since $(\log n)/(np) = o(1)$ and $\ell$ is a constant. Hence,

$$W(V; V, \ldots, V, A) \leq d_{\text{max}}^{\ell} |A| = |A|(n(p + q))^{\ell} \left(1 + O\left(\sqrt{\frac{\log n}{np}}\right)\right) \leq \hat{W}(V; V, \ldots, V, A) + O(|A| \sqrt{\log np}^{\ell-1/2}). \quad (42)$$

Note that $\hat{W}(V; V, \ldots, V, A) = |A|(n-1)p + nq)^{\ell}$. Thus, statement (40) can be supported by applying (38) and (41) to (42). Combining statements (38) and (40) completes the proof of Lemma 4.6.  

Proof of (P3) of Theorem 2.3. Let $d_{\text{min}} := \min_{v \in V} \deg(v)$. Then, for any $j \in [\ell]$,

$$\sum_{s \in \text{SL}_i} \left(\frac{\deg_A(s)}{\deg(s)}\right)^j \leq d_{\text{min}}^{-j} \sum_{s \in \text{SL}_i} \deg_A(s)^j = d_{\text{min}}^{-j} W(S \cap V_i; A, \ldots, A).$$

From Lemma 4.5, it holds w.h.p. that

$$d_{\text{min}}^{-j} = \left(n(p + q) - O(\sqrt{np \log n})\right)^{-j} = \left(1 + O\left(\sqrt{\frac{\log n}{np}}\right)\right)(n(p + q))^{-j}.$$
The second equality holds since \((\log n)/(np) = o(1)\) and \(j \in [\ell]\) is a constant. Thus, from Lemma 4.6,

\[
\sum_{s \in \mathcal{V}_j} \left( \frac{\deg_A(s)}{\deg(s)} \right) = \left( 1 + O \left( \sqrt{\frac{\log n}{np}} \right) \right) \sum_{j=1}^{\mathcal{H}} \left( S \cap \mathcal{V}_j : A, \ldots, A \right) + O \left( |A| \sqrt{\frac{\log n}{np}} \right) 
\]

\[
\leq \left| S \cap \mathcal{V}_j \right| \left( \frac{|A_1|p + |A_3|q}{n(p + q)} \right) + O \left( |A| \sqrt{\frac{\log n}{np}} \right) .
\]

Note that \(\sum_{i=1}^{\mathcal{H}} \left( |A_1|p + |A_3|q \right)/(n(p + q)) = \sum_{i=1}^{\mathcal{H}} \left( |A_1|p + |A_3|q \right)/(n(p + q)) \leq |A|/(p+q) \leq |A| \). Thus, we obtain the claim.

\[\Box\]

5 | PROOF OF LOCAL DYNAMICS AROUND FIXED POINTS

Consider an \((f_1, f_2)\)-polynomial voting process on an \(f_1\)- and \(f_2\)-good graph \(G\) for a partition \((V_1, V_2)\) and parameters \(p, q\). Throughout this section, the randomness is the generation of \(A'\) at each step. Moreover, we assume that \(q/p\) is a constant. Let \(H = (H_1, H_2) : [0, 1]^2 \rightarrow [0, 1]^2\) be the induced dynamical system.

5.1 | Proof of Proposition 2.9

Let \(B_2(x, r)\) denote an open ball of radius \(r\) (with respect to the \(\ell^2\)-norm) centered at \(x\). Let \(a^*\) be the sink point. From the property of the singular value (Proposition B.2) and Taylor expansion, there are two positive constants \(r\) and \(K\) such that, for any \(x \in B(a^*, r)\),

\[\|H(x) - x^*\|_2 = \|H(x) - H(x^*)\|_2 \leq \sigma_{\max} \|x - x^*\|_2 + O_{x \rightarrow x^*}(\|x - x^*\|_2^2) < (1 - K)r.\]

Let \(\epsilon > 0\) be such that \(\epsilon < r\) and \(\epsilon = o(1/\sqrt{np})\). From the Hoeffding inequality (Lemma B.9), for any \(A \subseteq V\) of \(\alpha \in B(a^*, \epsilon)\),

\[
\Pr[\|\alpha' - a^*\|_2 \geq \epsilon] \leq \Pr[\|\alpha' - E[\alpha']\|_2 + \|E[\alpha'] - H(\alpha)\|_2 + \|H(\alpha) - a^*\|_2 \geq \epsilon] 
\]

\[
\leq \Pr[\|\alpha' - E[\alpha']\|_2 \geq K\epsilon - O(1/\sqrt{np})] 
\]

\[
\leq \exp \left[ \|\alpha' - E[\alpha']\|_2 \geq \frac{K\epsilon}{\sqrt{2}} - O \left( \frac{1}{\sqrt{np}} \right) \right] 
\]

Fix an initial set \(A^{(0)} \subseteq V\) such that \(\alpha^{(0)} \in B(a^*, \epsilon)\). For any \(T \geq 0\), from the union bound over the time \(t = 1, \ldots, T\), we obtain

\[
\Pr \left[ \exists t \in [T] : \alpha^{(t)} \notin B(a^*, \epsilon) \right] \leq T \exp(-\Omega(\epsilon^2 n)).
\]

Suppose that \(\epsilon = o(\max \{ \sqrt{\log n/n}, \sqrt{1/np} \})\). If we set \(T = \exp(D\epsilon^2 n)\) for some constant \(D > 0\), the stopping time \(\tau = \min \{ t : \alpha^{(t)} \notin B(a^*, \epsilon) \}\) satisfies \(\tau \geq \exp(\Omega(\epsilon^2 n))\) w.h.p. Note that \(T_{\text{cons}}(A^{(0)}) \geq \tau\). This completes the proof.
5.2 Proof of Proposition 2.10

This subsection is devoted to prove Proposition 2.10. We begin with the following result which is of independent interest.

**Proposition 5.1.** Suppose that there are two positive constants $C$ and $\delta$ and a function $\epsilon = \epsilon(n) = o(1)$ such that

$$E[|A'|] \leq \frac{C|A|^2}{n} + \epsilon|A|$$

holds for all $A \subseteq V$ satisfying $|A| \leq \delta n$. Then, there are positive constants $\delta', C'$, and $C''$ such that

$$Pr \left[ T_{cons}(A) \leq C' \left( \log \log n + \frac{\log n}{\log e^{-1}} \right) \right] \geq 1 - n^{-C'}$$

holds for any $A \subseteq V$ satisfying $|A| \leq \delta' n$.

**Proof of Proposition 5.1.** Note that we may assume $\epsilon(n) = \Omega(\sqrt{\log n/n})$: If $\epsilon = o(\sqrt{\log n/n})$, we have $\log n / \log e^{-1} = O(1)$ and we will obtain the claim by applying Proposition 5.1 with letting $\epsilon = \sqrt{\log n/n}$.

Take a positive constant $\delta'$ such that

$$10 \left( \frac{CM^2}{n} + \epsilon M \right) \leq M, \quad (43)$$

$$\delta' \leq \min \left\{ \delta, \frac{1}{16C} \right\} \quad (44)$$

hold for any $0 \leq M \leq \delta' n$. This constant can be $\delta' > 0$ since $\epsilon = o(1)$. Thus, the inequality (43) holds if the ratio $\frac{M}{n}$ is sufficiently small.

Consider $A^{(0)}, A^{(1)}, \ldots$ given by the voting model such that $|A^{(0)}| \leq \delta n$. To exploit the assumption of the expectation, we first claim that $|A^{(i)}| \leq \delta' n \leq \delta n$ holds w.h.p. for all $t = 0, \ldots, n^{o(1)}$. Let $B^{(i)}$ be the event that $|A^{(i)}| \leq \delta n$ for all $i = 0, \ldots, t$. Note that $B^{(0)}$ holds. Consider $Pr[B^{(i+1)}|B^{(i)}]$. If $E[|A'|] \geq \log n$, from the Chernoff bound ((ii) of Lemma B.7), for any $A \subseteq V$ such that $|A| \leq \delta' n$, we obtain

$$Pr \left[ |A'| \geq 10 \left( \frac{C|A|^2}{n} + \epsilon|A| \right) \right] \leq Pr \left[ |A'| \geq 10E[|A'|] \right] \leq \exp \left( -\frac{10}{3} \log n \right) \leq n^{-3}.$$

Then, for $|A| \leq \delta' n$, it holds with probability $1 - O(n^{-3})$ that

$$|A'| \leq 10 \left( \frac{C|A|^2}{n} + \epsilon|A| \right) \leq |A| \leq \delta' n. \quad (45)$$

Here, we used (43) with $M = |A|$. If $E[|A'|] \leq \log n$, from the Chernoff bound ((iii) of Lemma B.7), we have

$$|A'| \leq 6 \log n = o(\delta' n) \quad (46)$$

with probability at least $1 - O(n^{-3})$. From (45) and (46), we obtain $Pr \left[ B^{(i+1)} \mid B^{(i)} \right] \geq 1 - O(n^{-3})$ for each $t$ and thus $B^{(i)}$ holds for $t = n^{0.01}$ with probability $1 - O(n^{-2.99})$. 

Next, consider $|A^{(t)}|$. Note that, if $|A| \leq \delta'n$, then
\[
E[|A'|] \leq \begin{cases} 
\frac{2C|A|^2}{n} & \text{if } \frac{\epsilon}{2} n \leq |A| \leq \delta'n, \\
2\epsilon|A| & \text{if } 0 \leq |A| \leq \frac{\epsilon}{2}.
\end{cases}
\]

In particular, consider the following two cases.

**Case I:** $\frac{\epsilon}{2} n \leq |A^{(t)}| \leq \delta'n$. From the Chernoff bound ((iii) of Lemma B.7), we have
\[
\Pr \left[ |A^{(t+1)}| \geq \frac{12C|A^{(t)}|^2}{n} \mid B^{(t)} \right] \leq 2^{-\Omega(\log n)}.
\]

Here, we used $|A^{(t)}| \geq \frac{\epsilon}{2} n = \Omega(\sqrt{n \log n})$. Hence, conditioned on $B^{(t)}$ and $|A^{(i)}| \geq \frac{\epsilon}{2} n$ ($i = 0, \ldots, t$), it holds w.h.p. that
\[
|A^{(t)}| \leq \frac{12C|A^{(t-1)}|^2}{n} \leq \frac{n}{12C} \left( \frac{12C|A^{(0)}|}{n} \right)^2 \leq 0.75^t \frac{n}{12C}.
\]

Here, we used (44). Therefore, for some $\tau_1 = O(\log \log n)$, $|A^{(r_1)}| \leq \frac{\epsilon}{2}$ holds w.h.p.

**Case II:** $0 \leq |A^{(t)}| \leq \frac{\epsilon}{2} n$. We claim that $E[|A^{(r_2)}|] \leq n^{-\Omega(1)}$ for some $\tau_2 = O(\log n / \log \epsilon^{-1})$, when conditioned on $|A^{(0)}| \leq \frac{\epsilon}{2} n$. Note that this completes the proof of Proposition 5.1 since $\Pr[A^{(r_2)} \neq \emptyset] \leq E[|A^{(r_1)}|] = n^{-\Omega(1)}$ from the Markov inequality.

To show the claim, we exploit the property that $E[|A'|] \leq 2\epsilon|A|$ if $|A| \leq \frac{\epsilon n}{C}$. Before using this, we show that $|A^{(t)}| \leq \frac{\epsilon n}{C}$ holds for all $t = 1, \ldots, n^{(t)}$. Conditioned on $|A| \leq \frac{\epsilon n}{C}$, we have $E[|A'|] \leq 2\epsilon|A| \leq O(\epsilon^2 n)$ and thus, from the Chernoff bound ((iii) of Lemma B.7), for any $A \subseteq V$ such that $|A| \leq \frac{\epsilon n}{C}$, we obtain $\Pr \left[ |A'| \geq \frac{\epsilon n}{C} \right] \leq 2^{-\Omega(\epsilon n)} = n^{-\Omega(1)}$. Therefore, $|A^{(t)}| \leq \frac{\epsilon}{2} n$ holds for all $t = 0, \ldots, n^{(1)}$. Let $C^{(i)}$ be the event that $|A^{(i)}| \leq \frac{\epsilon}{2} n$ holds for all $i = 0, \ldots, t$. Then, from the tower property of the conditional expectation, we have
\[
E[|A^{(r_2)}||C^{(r_2)}] \leq E[E[|A^{(r_2)}||A^{(r_1)}||C^{(r_1)}]|C^{(r_2)}]
\]
\[
\leq E[2\epsilon|A^{(r_1)}||C^{(r_2)}]
\]
\[
\leq (2\epsilon)^{r_2} \cdot \frac{\epsilon n}{C}
\]
\[
\leq n^{-\Omega(1)}
\]
for some $\tau_2 = O(\log n / \log \epsilon^{-1})$. This implies the aforementioned claim, which completes the proof of Proposition 5.1.

**Proof of Proposition 2.10.** It is sufficient to check the condition of Proposition 5.1 for $\epsilon = \epsilon(n) = \Theta \left( \sqrt{\frac{\log n}{np}} \right)$. Using (P3) and the Taylor expansion, there is a constant $C = C(H)$ such that
\[
E[|A'|] = nH_2(\alpha_1, \alpha_2) \pm O \left( |A| \sqrt{\frac{\log n}{np}} \right) \leq nC \left( (\alpha_1 + \alpha_2)^2 + |A| \sqrt{\frac{\log n}{np}} \right)
\]
\[
= C|A|^2/n + C|A| \sqrt{\frac{\log n}{np}}
\]
holds if $||\alpha|| \leq \delta$ for sufficiently small constant $\delta$. 

\[\blacksquare\]
5.3  Proofs of Propositions 2.12 and 2.13

Let \( a^* \) be a fixed point satisfying Assumption 2.11. Recall the random variable \( \beta \) defined in (11). From the definition (11), each element \( \beta_i \) of \( \beta \) can be rewritten as

\[
\beta_i = \sum_{j=1}^{2} u_{ij} \alpha_j - (Ua^*)_i,
\]

where \( U = (u_{ij}) \). Each element \( u_{ij} \) does not depend on \( n \). Hence, the Hoeffding bound (Lemma B.9) implies that

\[
\Pr[|\beta'_i - \mathbb{E}[\beta'_i]| \geq t] \leq \exp\left(-\Omega\left(t^2n\right)\right).
\]

From (3) and the Taylor expansion, it holds w.h.p. for any \( A \subseteq V \) that

\[
\mathbb{E}[\beta'] = U(H(\alpha) - a^*) + O\left(\frac{1}{\sqrt{np}}\right) \cdot 1
\]

\[
= UJ(\alpha - a^*) + \left(O_{a \rightarrow a^*}(\|\alpha - a^*\|_2^2) + O\left(\frac{1}{\sqrt{np}}\right)\right) \cdot 1
\]

\[
= \Lambda\beta + \left(O_{\|\beta\| \rightarrow 0}(\|\beta\|_\infty^2) + O\left(\frac{1}{\sqrt{np}}\right)\right) \cdot 1.
\]

Hence, the \( i \)th element \( \beta'_i \) of \( \beta = (\beta_1 \beta_2)^T \) satisfies

\[
|\mathbb{E}[\beta'_i]| = |\lambda_i||\beta_i| + O_{\|\beta\| \rightarrow 0}(\|\beta\|_\infty^2) + O\left(\frac{1}{\sqrt{np}}\right).
\]

It is convenient to consider the behavior of \( \beta \) instead of \( \alpha \). Note that \( \alpha \rightarrow a^* \) implies \( \beta \rightarrow 0 \) and vice versa since the matrix \( U \) is nonsingular. By substituting \( t = \Theta\left(\sqrt{\frac{\log n}{n}}\right) \) to (48), it holds w.h.p. for a sufficiently large constant \( C > 0 \) that

\[
||\beta'_i| - |\lambda_i||\beta_i|| | \leq C||\beta||_\infty^2 + C\sqrt{\frac{\log n}{n}}.
\]

**Proof of Proposition 2.13.** Suppose that the fixed point \( a^* \) satisfies the condition of Proposition 2.13 and Assumption 2.11. Let \( I_{>1} := \{i \in [2] : |\lambda_i| > 1\} \) and \( I_{\leq 1} := [2] \setminus I_{>1} \). Fix a sufficiently large constant \( K > 0 \) and let \( e^* \) be the constant mentioned in Proposition 2.13. Let

\[
\mathcal{A}_1 = \left\{A \subseteq V : \|\beta\|_\infty \leq e^* \text{ and } |\beta_j| < K\sqrt{\frac{\log n}{n}} \text{ for all } j \in I_{>1}\right\},
\]

\[
\mathcal{A}_2 = \left\{A \subseteq V : \|\beta\|_\infty \leq e^* \text{ and } |\beta_j| \geq K\sqrt{\frac{\log n}{n}} \text{ for some } j \in I_{>1}\right\},
\]

\[
\mathcal{A}_3 = \left\{A \subseteq V : \|\beta\|_\infty > e^* \text{ and } |\beta_j| \leq e^* \text{ for all } j \in I_{\leq 1}\right\}.
\]
We claim that, for each $i = 1, 2$ and any $A \in A_i$, there is some $\tau = O(\log n)$ satisfying
\[
\Pr \left[ A^{(\tau)} \in A_{i+1} \left| A^{(0)} = A \right. \right] \geq 1 - n^{-\Omega(1)}.
\]
This completes the proof of Proposition 2.13.

**Case I: $A^{(0)} \in A_1$.** Let $f(A) := \lfloor n \max \{ |\beta_i| : i \in I_{>1} \} \rfloor$ and $m = K\sqrt{n \log n}$.
We use Corollary B.13 to show that $A^{(\tau)} \in A_2$ for some $\tau = O(\log n)$.

From (47) and (A2), we have $\text{Var}[\beta_i] = \sum_{k \in [2]} u^2_{\theta} \text{Var}[\alpha_j] = \Omega(n^{-1})$ for any $A \in A_1$.
Note that, for every $i \in [2]$, there is $j \in [2]$ such that $u_{ij} \neq 0$; otherwise, it contradicts the fact that $U$ is nonsingular.
Thus, from Corollary B.11, it holds that, for any constant $h > 0$, there is a positive constant $C_1 < 1$ such that
\[
\Pr[f(A') < h\sqrt{n}] < C_1 \quad \text{for any } A \subseteq V \text{ with } f(A) \leq m.
\]
This verifies the condition $(1')$ of Corollary B.13.

Now we check the condition $(2')$. Let $z \in [2]$ be the least index satisfying $|\beta_z| = \max\{ |\beta_i| : i \in [2] \}$.
Suppose that $A \in A_1$ satisfies $f(A) = \lfloor n |\beta_z| \rfloor \geq h\sqrt{n}$ for sufficiently large constant $h > \frac{100C}{\epsilon_1}$ (recall that the constant $\epsilon_1$ is mentioned in (B1)). Then, from (B1), we have
\[
|E[\beta'_z]| \geq (1 + 0.99\epsilon_1)|\beta_z| + 0.01\epsilon_1|\beta_z| - \frac{C}{\sqrt{n}} \geq (1 + 0.99\epsilon_1)|\beta_z|.
\]
Thus, from the Hoeffding inequality (Lemma B.9), we obtain
\[
\Pr[f(A') < (1 + 0.98\epsilon_1)f(A)] \leq \Pr \left[ f(A') < \frac{1 + 0.98\epsilon_1}{1 + 0.99\epsilon_1} E[f(A')] \right] \leq \exp \left( -\Omega \left( \frac{(f(A))^2}{n} \right) \right)
\]
holds for every $A \in A_1$ satisfying $f(A) \geq h\sqrt{n}$.

Finally, we check the condition $(3')$ of Corollary B.13. From (B2), for any $A \in A_1$, it holds that
\[
\Pr[A' \notin A_1 \text{ and } f(A') < m] \leq \Pr[\exists j \in I_{\leq 1}, |\beta'_j| > \epsilon^*] \leq n^{-\Omega(1)}.
\]
Therefore, from Corollary B.13, we have $f(A^{(\tau)}) \geq m = K\sqrt{n \log n}$ (i.e., $A^{(\tau)} \in A_2$) holds w.h.p.

for some $\tau = O(\log n)$.

**Case II: $A^{(0)} \in A_2$.** Suppose that $A^{(0)} \in A_2$ and let $j \in I_{>1}$ be the index satisfying $|\beta_j| > K\sqrt{\log n} n$.
Note that $K$ is sufficiently large. From (B1) and (50), we have $|\beta'_j| \geq (1 + 0.99\epsilon_1)|\beta_j|$. Thus, for some $\tau = O(\log n)$, we have $|\beta^{(\tau)}_i| > \epsilon^*$. Moreover, from (B5), we have $|\beta^{(\tau)}_i| \leq \epsilon^*$ for all $i \in I_{\leq 1}$.

Therefore, $A^{(\tau)} \in A_3$ holds w.h.p.

**Proof of Proposition 2.12.** Suppose that the fixed point $a^*$ satisfies the condition of Proposition 2.12. Let
\[
I_{<1} := \{ i \in [2] : |\lambda_i| < 1 \},
I_{>1} := \{ i \in [2] : |\lambda_i| > 1 \}.
\]
Note that $I_{<1} \cup I_{>1} = [2]$. Moreover, there is some constant $\epsilon > 0$ such that
\[
||\lambda_i| - 1| > 3\epsilon
\]
holds for every $i \in [2]$. For $A \subseteq V$, let $z = z(A) \in [2]$ be the least index satisfying $|\beta_z| = ||\beta||_\infty$. Consider four constants: In (50) and (51), we defined $C$ and $\epsilon$. Let $K := \frac{C}{\epsilon}$ and $\epsilon' := \frac{\epsilon}{C}$. Consider four events
\[
B_1 = \left\{ A \subseteq V : K\sqrt{\frac{\log n}{n}} < ||\beta||_\infty \leq \epsilon' \text{ and } z(A) \in I_{<1} \right\}.
\]
\[
B_2 = \left\{ A \subseteq V : \| \beta \|_\infty \leq K \sqrt{\frac{\log n}{n}} \right\},
\]
\[
B_3 = \left\{ A \subseteq V : K \sqrt{\frac{\log n}{n}} < \| \beta \|_\infty \leq \epsilon' \text{ and } z(A) \in I_{>1} \right\},
\]
\[
B_4 = \left\{ A \subseteq V : \| \beta \|_\infty > \epsilon' \text{ and } | \beta_j | \leq \epsilon' \text{ for all } j \in I_{<1} \right\}.
\]

We claim that, if \( A^{(0)} \in B_1 \), then \( A^{(\tau)} \in B_2 \) holds w.h.p. for some \( j > i \) and some \( \tau = O(\log n) \). This completes the proof of Proposition 2.12.

**Case I: \( A^{(0)} \in B_1 \).** Suppose \( A^{(0)} \in B_1 \). We claim that, if \( A^{(i)} \in B_1 \), then either \( A^{(i+1)} \in B_2 \) or \( \| \beta^{(i+1)} \|_\infty \leq (1 - \epsilon) \| \beta^{(i)} \|_\infty \) holds w.h.p. For any \( j \in I_{<1} \),

\[
| \beta_j | \leq (1 - 3\epsilon) \| \beta \|_\infty + C \| \beta \|_\infty^2 + C \sqrt{\frac{\log n}{n}}
\]

\[
\leq (1 - \epsilon) \| \beta \|_\infty - 2\epsilon \| \beta \|_\infty + C \epsilon \| \beta \|_\infty + C \sqrt{\frac{\log n}{n}}
\]

\[
= (1 - \epsilon) | \beta |_\infty - \epsilon | \beta |_\infty + C \sqrt{\frac{\log n}{n}}
\]

\[
\leq (1 - \epsilon) | \beta |_\infty
\]

holds w.h.p. from (50). If \( A^{(i+1)} \notin B_3 \), then \( \| \beta' \|_\infty = | \beta_j | \) for some \( j \in I_{<1} \); thus, we have \( \| \beta' \|_\infty \leq (1 - \epsilon) \| \beta \|_\infty \) w.h.p. Therefore, for some \( \tau = O(\log n) \), it holds w.h.p. that \( A^{(\tau)} \in B_2 \cup B_3 \).

**Case II: \( A^{(0)} \in B_2 \).** Suppose \( A^{(0)} \in B_2 \). Our strategy is to apply Corollary B.13. We will prove the following result in the last part of this subsection.

**Lemma 5.2.** For any fixed \( A \in B_2 \), the following holds w.h.p.:

(i) For every \( i \in I_{<1} \), it holds that \( | \beta_i | \leq K \sqrt{\frac{\log n}{n}} \), and

(ii) there is a constant \( h > 0 \) such that \( \mathbb{E} [ \beta_i ] \geq (1 + \epsilon) | \beta_i | \) for every \( i \in I_{>1} \) satisfying \( | \beta_i | \geq \frac{h}{\sqrt{n}} \).

Let \( m = K \sqrt{n \log n} \) and define \( f(A) := | n \cdot \max_{i \in I_{<1}} | \beta_i | \). Suppose that \( f(A^{(\tau)}) \geq K \sqrt{n \log n} \) holds w.h.p. for some \( \tau = O(\log n) \). Then, we have \( A^{(\tau)} \notin B_1 \cup B_2 \) w.h.p. since \( | \beta_i^{(\tau)} | \leq K \sqrt{\frac{\log n}{n}} \) holds w.h.p. for any \( i \in I_{<1} \). Here, we used (i) of Lemma 5.2. To show \( f(A^{(\tau)}) \geq K \sqrt{n \log n} \), we check the conditions (1') to (3') of Corollary B.13 and then apply them.

First, we check the condition (1') of Corollary B.13. We use the same argument described in the Case I in Section 5.3. From (47), we have \( \text{Var}[\beta_i] \geq \sum_{j=1}^{2} u_{ij}^2 \text{Var}[\alpha_j] \). Moreover, for every \( i \in [2] \) there is \( j \in [2] \) such that \( u_{ij} \neq 0 \); otherwise, it contradicts to the fact that \( U \) is nonsingular. From (A2), we have \( \text{Var}[\beta_i] = \Omega(n^{-1}) \); thus, from Corollary B.11, it holds that, for any constant \( h > 0 \), there is a positive constant \( C_1 < 1 \) such that \( \Pr[f(A') \geq h \sqrt{n}] < C_1 \) holds for any \( A \subseteq V \) with \( f(A) < m \).

We check the condition (2') of Corollary B.13. For every \( i \in I_{>1} \), from Lemma 5.2, we obtain

\[
| \mathbb{E}[\beta_i^{(\tau)}|A^{(\tau)} \in B_2]| \geq (1 + \epsilon)| \beta_i^{(\tau)} |
\]

(52)
In look at (47), from the Hoeffding inequality (Lemma B.9), it holds for any set \( A^{(i)} \in B_2 \), any index \( i \in I_1 \) and any constant \( \epsilon' > 0 \) that

\[
\Pr[|\beta'_i| \leq (1 - \epsilon')|E[\beta'_i]|] \leq \exp\left(-\Omega \left( \epsilon'^2 E[|\beta'_i|^2 n] \right) \right) \leq \exp\left(-\Omega \left( \frac{\epsilon'^2 f(A)^2}{n} \right) \right).
\]

From (52) and (53), by letting \( \epsilon' = \frac{\epsilon}{2(1 + \epsilon)} \), we obtain

\[
\Pr \left[ |\beta_i^{(t+1)}| \leq \left( 1 + \frac{\epsilon}{2} \right) \cdot |\beta_i^{(t)}| \mid A^{(i)} \in B_2 \right] \leq \Pr \left[ |\beta'_i| \leq (1 - \epsilon') \cdot |E[\beta'_i|A^{(i)} \in B_2]| \right]
\leq \exp\left(-\Omega \left( \frac{f(A)^2}{n} \right) \right).
\]

In other words, for any \( A \in B_2 \) satisfying \( f(A) \geq h \sqrt{n} \) for some constant \( h > 0 \), we have

\[
\Pr \left[ f(A^{(t+1)}) < \left( 1 + \frac{\epsilon}{2} \right) f(A^{(t)}) \mid A^{(i)} = A \right] \leq \exp\left(-\Omega \left( \frac{f(A)^2}{n} \right) \right).
\]

Finally, we check the condition \( (3') \) of Corollary B.13. From Lemma 5.2, we have

\[
\Pr[ A^{(t+1)} \notin B_2 \land f(A^{(t+1)}) \leq m|A^{(t)} \in B_2] \leq \Pr \left[ \exists j \in I_{<1} : |\beta'_j| > K \sqrt{\log \frac{n}{n}} \mid A \in B_2 \right] \leq n^{-\Omega(1)}.
\]

From Corollary B.13, there is some \( \tau = O(\log n) \) such that \( f(A^{(\tau)}) \geq K \sqrt{\log \frac{n}{n}} \) and \( |\beta_j^{(\tau)}| \leq K \sqrt{\log \frac{n}{n}} \) holds w.h.p. for every \( j \in I_{<1} \). Consequently, \( A^{(\tau)} \in B_3 \cup B_4 \) holds w.h.p.

**Case III:** \( A^{(0)} \in B_3 \). Suppose that \( A^{(0)} \in B_3 \). From (50), it holds w.h.p. that

\[
|\beta'_j| \geq \hat{\lambda}_2 ||\beta|| - C ||\beta|| \infty - C \sqrt{\log \frac{n}{n}}
\geq (1 + \epsilon)|\beta| + (\epsilon|\beta| - C|\beta|) + \left( \epsilon|\beta| - C \sqrt{\log \frac{n}{n}} \right)
\geq (1 + \epsilon)|\beta|.
\]

Moreover, for any \( j \in I_{<1} \), it holds w.h.p. that

\[
|\beta'_j| \leq (1 - 3\epsilon)||\beta|| + C ||\beta|| \hat{\lambda} \infty + C \sqrt{\log \frac{n}{n}} \leq (1 - \epsilon)|\beta|.
\]

These imply that \( A^{(t+1)} \notin B_1 \cup B_2 \) holds w.h.p. whenever \( A^{(t)} \in B_3 \). Let \( \tau \) be the stopping time given by \( \tau := \min \{ t : A^{(t)} \notin B_3 \} \). Then, \( ||\beta^{(t+1)}|| \infty \geq (1 + \epsilon)||\beta^{(t)}|| \infty \) holds w.h.p. for all \( t < \tau \). Therefore, we have \( A^{(\tau)} \in B_3 \) with \( \tau = O(\log n) \), and \( |\beta_j^{(\tau)}| \leq \epsilon' \) for all \( j \in I_{<1} \).
CONCLUDING REMARK

This study considered the Best-of-two and the Best-of-three voting processes on the stochastic block model $G(2n, p, q)$. Here, we first generate $G(2n, p, q)$, then set an initial opinion configuration and observe the voting process. We presented phase transition results on $r = q/p$ for both processes. In addition, if $p \geq q > 0$ are constants, we proved that the consensus time is $O(\log n)$ for arbitrary initial opinion configurations. In the proof, we combined the theory of dynamical systems and our technical result of Theorem 2.4 which approximates the stochastic processes by the corresponding appropriate deterministic processes. To estimate the probability which the process reaches sink areas from the source area is future work to consider an application of these processes to distributed community detection algorithms. For an application to distributed community detection algorithms, it is significant to estimate the probability that the voting process reaches the sink areas (in particular, starting from the source area). This is a possible future direction of this.

Note that Theorem 2.4 is allowable for any polynomial function with a constant degree. For example, consider the Best-of-$(2k + 1)$ voting process for a positive constant $k$. This process is defined by $f_i(x) = f_2(x) := \Pr[\text{Bin}(2k + 1, x) \geq k + 1] = \sum_{i=k+1}^{2k+1} \binom{2k+1}{i} x^i (1-x)^{2k+1-i}$. Theorem 2.4 guarantees $||\alpha^{(n)} - \alpha^{(0)}||_{\infty} \leq C(\sqrt{1/np} + \sqrt{\log n/n})$ for this process. Moreover, using Lemma 4.1, it is not difficult to extend Theorem 2.4 to voting processes on general stochastic block models that has $c_1$ communities each of size $\Omega(n)$ and initially involving $c_2$ opinions, where $c_1$ and $c_2$ denote arbitrary positive constants. This setting yields induced dynamical systems of dimension more than two. The Jacobian matrix would be helpful to investigate several properties including the exponential time lower bound (Proposition 2.9), the fast consensus (Proposition 2.10) and escape result (Proposition 2.12) since the proofs of Sections 5.1 to 5.3 work for induced dynamical systems with higher dimensions. Unfortunately, it may not be easy to specify other properties (e.g., number of fixed points, orbit convergence). This problem is left for future work. Also, the worst-case analysis of the consensus time for sparse random graphs remains open in this study.
ACKNOWLEDGMENTS

This work is supported by JST CREST Grant Number JPMJCR14D2, and JSPS KAKENHI Grant Number 17H07116, 19J12876, and 19K20214, Japan. We would like to thank Colin Cooper, Nan Kang, and Tomasz Radzik for helpful discussions.

REFERENCES

1. E. Abbe, *Community detection and stochastic block models: Recent developments*, J. Mach. Learn. Res. **18** (2018), 1–86.
2. E. Abbe and C. Sandon, *Recovering communities in the general stochastic block model without knowing the parameters*, Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS), vol. 1, 2015, pp. 676–684.
3. M. A. Abdullah and M. Draief, *Global majority consensus by local majority polling on graphs of a given degree sequence*, Discrete Appl. Math. **121** (2002), 1–10.
4. Y. Afek, N. Alon, O. Barad, E. Hornstein, N. Barkai, and Z. Bar-Joseph, *A biological solution to a fundamental distributed computing problem*, Science **331** (2011), 183–185.
5. P. Barbilhon, S. Donnet, E. Lazega, and A. Bar-Hen, *Stochastic block models for multiplex networks: An application to a multilevel network of researchers*, J. R. Stat. Soc. A **180** (2017), 295–314.
6. L. Becchetti, A. Clementi, P. Manurangsi, E. Natale, F. Pasquale, P. Raghavendra, and L. Trevisan, *Average whenever you meet: Opportunistic protocols for community detection*, Proceedings of the 26th Annual European Symposium on Algorithms (ESA), vol. 7, 2018, pp. 1–13.
7. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, R. Silvestri, and L. Trevisan, *Simple dynamics for plurality consensus*, Distrib. Comput. **30** (2017), 293–306.
8. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan, *Stabilizing consensus with many opinions*, Proceedings of the 27th annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2016, pp. 620–635.
9. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan, *Find your place: Simple distributed algorithms for community detection*, SIAM J. Comput. **49** (2020), 821–864.
10. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan, *Fast consensus for voting on general expander graphs*, Proceedings of the 29th International Symposium on Distributed Computing (DISC), 2015, pp. 248–262.
11. C. Cooper, T. Radzik, *The power of two choices in distributed voting*, Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP), vol. 2, 2014, pp. 435–446.
12. C. Cooper, T. Radzik, N. Rivera, and T. Shiraga, *Fast plurality consensus in regular expanders*, Proceedings of the 31st International Symposium on Distributed Computing (DISC), vol. 91(13), 2017, pp. 1–16.
13. C. Cooper and N. Rivera, *The linear voting model*, Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP), vol. 55(144), 2016, 1–12.
14. E. Cruciani, E. Natale, A. Nusser, and G. Scornavacca, *Phase transition of the 2-choices dynamics on core-periphery networks*, Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), 2018, pp. 777–785.
15. E. Cruciani, E. Natale, and G. Scornavacca, *Distributed community detection via metastability of the 2-choices dynamics*, Proceedings of the 33rd AAAI conference on artificial intelligence (AAAI), 2019, pp. 6046–6053.
16. B. Daniel and T. Tassa, *Improved bounds on bell numbers and on moments of sums of random variables*, Probab. Math. Stat. **30** (2010), 185–205.
17. B. Doerr, *Probabilistic tools for the analysis of randomized optimization heuristics*, arXiv:1801.06733, 2018.
18. B. Doerr, L. A. Goldberg, L. Minder, T. Sauerwald, and C. Scheideler, *Stabilizing consensus with the power of two choices*, Proceedings of the 23rd annual ACM symposium on Parallelism in algorithms and architectures (SPAA), 2011, pp. 149–158.
19. A. L. Dontchev and R. T. Rockafellar, *Implicit functions and solution mappings: A view from variational analysis*, 2nd ed., Springer-Verlag, New York, 2014.
24. M. Fischer, N. Lynch, and M. Merritt, *Easy impossibility proofs for distributed consensus problems*, Distrib. Comput. 1 (1986), 26–39.
25. A. Frieze and M. Karoński, *Introduction to random graphs*, University Press, Cambridge, 2016.
26. M. Ghaffari and J. Lengler, *Nearly-right analysis for 2-choice and 3-majority consensus dynamics*, Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC), 2018, pp. 305–313.
27. S. Gilbert and D. Kowalski, *Distributed agreement with optimal communication complexity*, Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2010, pp. 965–977.
28. A. Goldenberg, A. X. Zheng, S. E. Fienberg, and E. M. Airoldi, *A survey of statistical network models*, Found. Trends Mach. Learn. 2 (2010), 129–233.
29. Y. Hassin and D. Peleg, *Distributed probabilistic polling and applications to proportionate agreement*, Inform. Comput. 171 (2001), 248–268.
30. M. Hirsch and H. L. Smith, *Monotone dynamical systems*, in *Handbook of Differential Equations: Ordinary Differential Equations*, Vol 2, Elsevier B. V., Amsterdam, 2005, pp. 239–357.
31. V. Kanade, F. Mallmann-Trenn, and T. Sauerwald, *On coalescence time in graphs: When is coalescing as fast as meeting?*, Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2018, pp. 305–313.
32. S. Gilbert and D. Kowalski. *Distributed agreement with optimal communication complexity*, Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2010, pp. 965–977.
33. J. H. Kim and V. H. Vu, *Concentration of multivariate polynomials and its applications*, Combinatorica 20 (2000), 417–434.
34. S. Lang, *Calculus of several variables*, Springer-Verlag, New York, 1987.
35. T. M. Liggett, *Interacting particle systems*, Springer-Verlag, New York, 1985.
36. E. Mossel, J. Neeman, and O. Tamuz, *Majority dynamics and aggregation of information in social networks*, Autonom. Agents Multi-Agent Syst. 28 (2014), 408–429.
37. E. M. Marcotte, M. Pellegrini, H.-L. Ng, D. W. Rice, T. O. Yeates, and D. Eisenberg, *Detecting protein function and protein-protein interactions from genome sequences*, Science 285 (1999), 751–753.
38. N. Kang and R. Rivera, *Best-of-three voting on dense graphs*, Proceedings of the 31st ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), 2019, pp. 115–121.
39. J. H. Kim and V. H. Vu, *Concentration of multivariate polynomials and its applications*, Combinatorica 20 (2000), 417–434.
40. R. I. Oliveira and Y. Peres, *Random walks on graphs: new bounds on hitting, meeting, coalescing and returning*, Proceedings of the 16th Workshop on Analytic Algorithmics and Combinatorics (ANALCO), 2019, pp. 119–126.
41. G. Schoenebeck and F. Yu, *Consensus of interacting particle systems on Erdős–Rényi graphs*, Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2018, pp. 1945–1964.
42. I. G. Shevtsova, *An improvement of convergence rate estimates in the Lyapunov theorem*, Doklady Math. 82 (2010), 862–864.
43. N. Shimizu and T. Shiraga, *Phase transitions of Best-of-two and Best-of-three on stochastic block models*, Proceedings of the 33rd International Symposium on Distributed Computing (DISC), 2019, pp. 32:1–32:17.

**How to cite this article:** Shimizu N, Shiraga T. Phase transitions of Best-of-two and Best-of-three on stochastic block models. *Random Struct Alg* 2021;59:96–140. https://doi.org/10.1002/rsa.20992

**APPENDIX A**

**VOTING PROCESSES ON G(2n,p,q) WITH A NON-CONSTANT r=r(n)**

Here, we check that the proof of Theorem 1.2 works even if \( r = r(n) \) is a function. This study is motivated by the understanding of the Best-of-three on \( G(2n,p,q) \) with \( q/p = o(1) \). The discussion given in this section also works for the Best-of-two (i.e., the proof of Theorem 1.3). Consider a graph \( G(2n, p, q) \) with \( q/p = r = r(n) \). Suppose that the limit \( \tilde{r} := \lim_{n \to \infty} r(n) \in [0, 1] \) exists. Let \( u = u(n) := \frac{1 - r(n)}{1 + r(n)} \) and \( \tilde{u} := \lim_{n \to \infty} u(n) = \frac{1 - \tilde{r}}{1 + \tilde{r}} \), respectively.
A.1  |  Orbit convergence

Let \( f : \mathbb{R}^\ell \rightarrow \mathbb{R} \) be a function having continuous partial derivatives up to order 1. Then, from the Taylor’s formula (see e.g., Theorem 5.2 in [35]), for any \((x_i)_{i=1}^\ell, (y_i)_{i=1}^\ell \in [-1, 1]^\ell\), we have

\[
[f(x_1, \ldots, x_\ell) - f(y_1, \ldots, y_\ell)] \leq K_1(f) \sum_{i=1}^\ell |x_i - y_i|,  \tag{A1}
\]

where \( K_1(f) := \max_{i \in [\ell]} \max_{c_1, \ldots, c_\ell} \left| \frac{\partial f}{\partial x_i}(c_1, \ldots, c_\ell) \right| \). For each \( i = 1, 2, \) consider \( T_i(d_1, d_2, u) = T_i(d_1, d_2) \) defined in \( (12) \). Then there is a positive constant \( C_1 \) such that \( \max_{i \in \{1, 2\}} K_1(T_i) \leq C_1 \) since \( T_1 \) and \( T_2 \) are polynomials.

For a given initial configuration \( \mathbf{d}^{(0)} \in [-1, 1]^2 \), consider two sequences \((\mathbf{d}^{(i)})_{i=0}^\infty \) and \((\tilde{\mathbf{d}}^{(i)})_{i=0}^\infty \) defined as follows. The initial point is \( \tilde{\mathbf{d}}^{(0)} := \mathbf{d}^{(0)} \), and for each \( t \geq 0 \),

\[
\mathbf{d}^{(t+1)} := \left( T_1(d_1^{(t)}, d_2^{(t)}, u), T_2(d_1^{(t)}, d_2^{(t)}, u) \right), \quad \tilde{\mathbf{d}}^{(t+1)} := \left( T_1(\tilde{d}_1^{(t)}, \tilde{d}_2^{(t)}, \tilde{u}), T_2(\tilde{d}_1^{(t)}, \tilde{d}_2^{(t)}, \tilde{u}) \right). \tag{A2}
\]

Then, for any positive constant \( t \),

\[
\|\mathbf{d}^{(t)} - \tilde{\mathbf{d}}^{(t)}\|_{\infty} = \left\| \begin{pmatrix} T_1(d_1^{(t)}, d_2^{(t)}, u) \\ T_2(d_1^{(t)}, d_2^{(t)}, u) \end{pmatrix} - \begin{pmatrix} T_1(\tilde{d}_1^{(t)}, \tilde{d}_2^{(t)}, \tilde{u}) \\ T_2(\tilde{d}_1^{(t)}, \tilde{d}_2^{(t)}, \tilde{u}) \end{pmatrix} \right\|_{\infty}
\leq C_1 \left( \sum_{i=1}^{2} |d_i^{(t-1)} - \tilde{d}_i^{(t-1)}| + |u - \tilde{u}| \right)
\leq 2C_1 \|\mathbf{d}^{(t-1)} - \tilde{\mathbf{d}}^{(t-1)}\|_{\infty} + C_1 |u - \tilde{u}|
\leq \cdots \leq (2C_1)^t \|\mathbf{d}^{(0)} - \tilde{\mathbf{d}}^{(0)}\|_{\infty} + |u - \tilde{u}| \sum_{i=0}^{t-1} C_1 \leq C_2 |u - \tilde{u}|,
\]

where \( C_1 \) and \( C_2 \) are sufficiently large constants. Thus, from \((9)\), the following holds w.h.p. for any positive constant \( t \) and a sufficiently large positive constant \( C \):

\[
\|\mathbf{d}^{(t)} - \tilde{\mathbf{d}}^{(t)}\|_{\infty} \leq C \left( \frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} + |u - \tilde{u}| \right).  \tag{A3}
\]

Note that \( |u - \tilde{u}| = o_{n \to \infty}(1) \) from the definition. Furthermore, \( \|\mathbf{d}^{(i)}_i - \tilde{\mathbf{d}}^{(i)}_i\|_{\infty} = O(|u - \tilde{u}|) = o_{n \to \infty}(1) \) for each \( i \in \{4\} \). Thus, Proposition 3.2 implies that \( \langle \mathbf{d}^{(i)}_i \rangle_{\ast} \in B(\mathbf{d}^{(i)}_i, \varepsilon) \) within a constant number of steps w.h.p. for some \( i \in \{4\} \).

A.2  |  Local dynamics around fixed points

Let \( f : \mathbb{R}^\ell \rightarrow \mathbb{R} \) be a function having continuous partial derivatives up to order 2. Applying Taylor’s formula (see e.g., Theorem 5.2 in [35]), for any \((x_i)_{i=1}^\ell, (y_i)_{i=1}^\ell \in [-1, 1]^\ell\),

\[
[f(x_1, \ldots, x_\ell) - f(y_1, \ldots, y_\ell)] - \sum_{i=1}^\ell \frac{\partial f}{\partial x_i}(y_1, \ldots, y_\ell)(x_i - y_i) \leq \frac{K_2(f)}{2} \left( \sum_{i=1}^\ell |x_i - y_i| \right)^2.  \tag{A4}
\]


where $K_2(f) := \max_{i,j \in [\ell], (c_1, \ldots, c_\ell) \in [-1,1]^\ell} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(c_1, \ldots, c_\ell) \right|$.

1. **Sink points.** Suppose $r \leq \frac{1}{7} - \varepsilon$ for a positive constant $\varepsilon$. Here, $u \geq \frac{3}{4} + \Theta(\varepsilon)$ holds. It is sufficient to check that the largest singular value $\sigma_{\text{max}}$ of the Jacobian matrix at the sink point is at most $1 - \Omega(\varepsilon)$; indeed, the proof of Proposition 2.9 works if this condition holds. In the Best-of-three, the corresponding Jacobian matrix is $3\left( \begin{array}{cc} 1 - u & 0 \\ 0 & (1/u) - 1 \end{array} \right)$ (see (15)); thus, we have $\sigma_{\text{max}} \leq 1 - \Theta(\varepsilon)$.

2. **Consensus points.** Consider $H_i(\alpha_1, \alpha_2, r) = f^{Bo3}\left( \frac{\alpha_1 + \alpha_2}{1+r} \right)$ for $i = 1, 2$. Then, it is not difficult to see that there is a positive constant $C$ satisfying $\max_{i \in [2]} K_2(H_i) \leq C$. Thus using (A2) at $(\alpha_1, \alpha_2) = (0, 0)$,

$$H_1(\alpha_1, \alpha_2, r) + H_2(\alpha_1, \alpha_2, r) \leq C \left( \alpha_1 + \alpha_2 \right)^2 + C \left( \alpha_1 + \alpha_2 \right)^2 = C\alpha^2.$$  

Proposition 5.1 can be applied by combining (A3) and (P3).

3. **Source and saddle points.** It is sufficient to check (50) for general $u = u(n)$. It is obvious since $T_1$ and $T_2$ are polynomials. To state it more formally, there is a positive constant $C$ satisfying $\max_{i \in [2]} K_2(T_i) \leq C$. Thus, for any $d, d' \in [-1,1]^2$ and $u \in [0, 1]$,

$$\left\| \begin{pmatrix} T_1(d_1, d_2, u) \\ T_2(d_1, d_2, u) \end{pmatrix} - \begin{pmatrix} T_1(d'_1, d'_2, u) \\ T_2(d'_1, d'_2, u) \end{pmatrix} - J(d'_1, d'_2) \begin{pmatrix} d_1 - d'_1 \\ d_2 - d'_2 \end{pmatrix} \right\|_{\infty} \leq C \left( \sum_{i=1}^{2} |d_i - d'_i| \right)^2 \leq 4C||d - d'||^2_{\infty}$$

holds from (A2). Here, $J(d'_1, d'_2)$ denotes the Jacobian matrix at $(d'_1, d'_2)$ (see (15)).

**APPENDIX B**

**B | TOOLS**

**B.1 | Linear algebra**

**Definition B.1** (singular value). For a real matrix $A \in \mathbb{R}^{m \times n}$, singular values $\sigma_1, \ldots, \sigma_m$ of $A$ are nonnegative square roots of eigenvalues of $AA^\top$. We write $\sigma_i(A)$, when we specify $A$. In particular, the maximum singular value, denoted by $\sigma_{\text{max}}$, is the largest value among all singular values.

**Proposition B.2.** For a real matrix $A \in \mathbb{R}^{m \times n}$, it holds that

$$\sigma_{\text{max}} = \max_{v \in \mathbb{R}^n : \|v\|_2 = 1} \|Av\|_2,$$

where the norm $\| \cdot \|_2$ is the $\ell^2$ norm.

In particular, it holds that

$$\|Av\|_2 \leq \sigma_{\text{max}} \|v\|_2$$

for any vector $v \in \mathbb{R}^n$. 
B.2 | Real analysis

The inverse function theorem. The inverse function theorem is a fundamental result in real analysis and can be found in many textbooks [23,34].

**Theorem B.3** (The inverse function theorem (see e.g., Theorem 12.17 in [34] and Theorem 1A.1 in [23])). Let \( f \) be a continuously differentiable function from an open set \( U \subseteq \mathbb{R}^k \) into \( \mathbb{R}^k \). Suppose that the Jacobian matrix \( J \) at \( p \in U \) is invertible. Then, there is a neighborhood \( V \) of \( p \) such that the restriction of \( f \) to \( V \) is invertible. Moreover, the Jacobian matrix of \( f^{-1} \) at \( p \) is given by \( J^{-1} \).

Lipschitz functions.

**Definition B.4.** Consider a function \( H : S \rightarrow T \), where \( S \subseteq \mathbb{R}^m \) and \( T \subseteq \mathbb{R}^n \) are closed sets. The function \( H \) satisfies the Lipschitz condition if there is a constant \( C > 0 \) such that

\[
\|H(x) - H(y)\|_{\infty} \leq C \|x - y\|_{\infty}
\]

holds for any \( x, y \in S \).

It should be noted that the definition of the Lipschitz condition does not depend on the norm \( \| \cdot \| \) on \( \mathbb{R}^n \). The following is a well-known result in real analysis.

**Proposition B.5** (Exercise 1D.3 in [23]). Let \( O \subseteq \mathbb{R}^k \) be an open set and \( S \subseteq O \) be a compact convex subset of \( O \). Suppose that \( H : O \rightarrow \mathbb{R}^k \) is continuously differentiable on an open set \( O \). Then \( H \) is Lipschitz continuous on \( C \) and

\[
\|H(x) - H(y)\|_{\infty} \leq \max_{p \in S} \sigma_{\max}(J_p) \|x - y\|_{\infty},
\]

where \( J_p \) is the Jacobian matrix at \( p \).

**Corollary B.6.** Let \( H : \mathbb{R}^m \rightarrow \mathbb{R}^n \) be a function given by

\[
H(x) = (H_1(x), \ldots, H_n(x)),
\]

where \( H_i(x) = H_i(x_1, \ldots, x_m) \) is a polynomial on \( x_1, \ldots, x_m \) for all \( i \in [n] \). Then, \( H \) satisfies the Lipschitz condition on \([0, 1]^m\).

B.3 | Probabilistic tools

**Lemma B.7** (The Chernoff bound (see e.g., Theorem 10.1, Corollary 10.4 and Theorem 10.5 in [21])). Let \( X_1, X_2, \ldots, X_n \) be independent random variables taking values in \([0, 1]\). Let \( X = \sum_{i=1}^n X_i \). Then, the following hold:

(i) For any \( \delta \geq 0 \),

\[
\Pr[X \geq (1 + \delta)E[X]] \leq \exp\left(-\frac{\min\{\delta, \delta^2\}E[X]}{3}\right).
\]
(ii) for any $\delta \in [0, 1]$,

$$\Pr[X \leq (1 - \delta)\mathbb{E}[X]] \leq \exp\left(-\frac{\delta^2\mathbb{E}[X]}{2}\right).$$

(iii) for any $k \geq 2e\mathbb{E}[X]$,

$$\Pr[X \geq k] \leq 2^{-k}.$$

Here, $e$ denotes the Napier constant.

Lemma B.8 (The additive Chernoff bound (see e.g., Theorems 10.10 and 10.11 in [21])). Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values in $[0, 1]$. Let $X = \sum_{i=1}^{n} X_i$. Then, for any $\delta \geq 0$,

$$\Pr[X \geq \mathbb{E}[X] + \delta] \leq \exp\left(-\frac{1}{3} \min\left\{ \frac{\delta^2}{\mathbb{E}[X]}, \delta \right\}\right),$$

$$\Pr[X \leq \mathbb{E}[X] - \delta] \leq \exp\left(-\frac{\delta^2}{2\mathbb{E}[X]}\right).$$

Lemma B.9 (The Hoeffding bound (see e.g., Theorem 10.9 in [21])). Let $X_1, X_2, \ldots, X_n$ be independent random variables. Assume that each $X_i$ takes values in a real interval $[a_i, b_i]$ of length $c_i := b_i - a_i$. Let $X = \sum_{i=1}^{n} X_i$. Then, for any $\delta > 0$,

$$\Pr[X \geq \mathbb{E}[X] + \delta] \leq \exp\left(-\frac{2\delta^2}{\sum_{i=1}^{n} c_i^2}\right),$$

$$\Pr[X \leq \mathbb{E}[X] - \delta] \leq \exp\left(-\frac{2\delta^2}{\sum_{i=1}^{n} c_i^2}\right).$$

Lemma B.10 (The Berry-Esseen theorem (see e.g., [42])). Let $X_1, X_2, \ldots, X_n$ be independent random variables such that $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_i^2] > 0$, $\mathbb{E}[|X_i|^3] < \infty$ for all $i \in [n]$ and $\sum_{i=1}^{n} \mathbb{E}[X_i^2] = 1$. Let $X = \sum_{i=1}^{n} X_i$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$ (the cumulative distribution function of the standard normal distribution). Then

$$\sup_{x \in \mathbb{R}} \left| \Pr[X \leq x] - \Phi(x) \right| \leq 5.6 \sum_{i=1}^{n} \mathbb{E}[|X_i|^3].$$

Corollary B.11. Let $X_1, X_2, \ldots, X_n$ be $n$ independent random variables such that $\text{Var}[X] \neq 0$ and $|X_i - \mathbb{E}[X_i]| \leq C < \infty$ for all $i \in [n]$ where $X = \sum_{i=1}^{n} X_i$. Then, for any $x \in \mathbb{R}$,

$$\left| \Pr\left[\frac{X - \mathbb{E}[X]}{\sqrt{\text{Var}[X]}} \leq x \right] - \Phi(x) \right| \leq \frac{5.6C}{\sqrt{\text{Var}[X]}}.$$

Proof. For all $i \in [n]$, let

$$Z_i := \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\text{Var}[X]}}.$$
Note that $E[Z_i^2] = 0$ $\iff$ $\sum_i Z_i = 0$ $\iff$ $P(Z_i = 0) = 1$. Then, for all $i \in \{j \in [n] : E[Z_j^2] > 0\}$, it is easy to check that $E[Z_i] = 0$, $E[Z_i^2] > 0$, and $E[|Z_i|^3] \leq \frac{C^3}{\text{Var}[X]^{3/2}} < \infty$. Furthermore,

\[
\sum_{i \in [n]: E[Z_i^2] > 0} E[Z_i^2] = \sum_{i \in [n]} E[Z_i^2] = \frac{\sum_{i \in [n]} E[(X_i - E[X_i])^2]}{\text{Var}[X]} = 1.
\]

Thus, the following can be obtained by applying Lemma B.10 to $Z$:

\[
\left| \Pr \left[ \frac{X - E[X]}{\sqrt{\text{Var}[X]}} \leq x \right] - \Phi(x) \right| = \left| \Pr \left[ \sum_{i=1}^n Z_i \leq x \right] - \Phi(x) \right|
\]

\[
= \left| \Pr [Z \leq x] - \Phi(x) \right|
\]

\[
\leq 5.6 \sum_{i \in [n]: E[Z_i^2] > 0} E[|Z_i|^3]
\]

\[
\leq \frac{5.6C}{\sqrt{\text{Var}[X]}} \sum_{i=1}^n E[Z_i^2] = \frac{5.6C}{\sqrt{\text{Var}[X]}}.
\]

\[\blacksquare\]

**Lemma B.12** (Lemma 4.5 of [12]). Consider a Markov chain $(X_i)_{i=1}^\infty$ with finite state space $\Omega$ and a function $f : \Omega \to \{0, \ldots, n\}$. Let $C_3$ be arbitrary constant and $m = C_3 \sqrt{n \log n}$. Suppose that $\Omega, f$ and $m$ satisfy the following conditions:

1. For any positive constant $h$, there is a positive constant $C_1 < 1$ such that

\[
\Pr \left[ f(X_{i+1}) < h \sqrt{n} \left| f(X_i) \leq m \right. \right] < C_1.
\]

2. There are three positive constants $\epsilon, C_2$ and $h$ such that, for any $x \in \Omega$ satisfying $h \sqrt{n} \leq f(x) < m$,

\[
\Pr \left[ f(X_{i+1}) < (1 + \epsilon)f(X_i) \left| X_i = x \right. \right] < \exp \left( -C_2 \frac{f(x)^2}{n} \right).
\]

Then $f(X_{\tau}) \geq m$ holds for some $\tau = O(\log n)$.

**Corollary B.13.** Consider a Markov chain $(X_i)_{i=1}^\infty$ with a finite state space $\Omega$ and a function $f : \Omega \to \{0, \ldots, n\}$. Let $C_3$ be arbitrary constant and $m = C_3 \sqrt{n \log n}$. Consider a set $B \subseteq \Omega$ such that

\[
B \subseteq \{x \in \Omega : f(x) < m\}.
\]

Suppose that $\Omega, f, m$ and $B$ satisfy the following conditions:

1. For any positive constant $h$, there is a positive constant $C_1 < 1$ such that

\[
\Pr \left[ f(X_{i+1}) < h \sqrt{n} \left| f(X_i) \leq m, X_i \in B \right. \right] < C_1.
\]
(2') There are three positive constants $\epsilon, C_2, \text{ and } h$ such that, for any $x \in B$ satisfying $h\sqrt{n} \leq f(x) < m,$

$$\Pr\left[ f(X_{t+1}) < (1 + \epsilon)f(X_t) \mid X_t = x \right] < \exp \left( -C_2 \frac{f(x)^2}{n} \right).$$

(3') For some constant $C_4 > 0,$

$$\Pr\left[ X_{t+1} \not\in B \text{ and } f(X_{t+1}) < m \mid X_t \in B \right] \leq O(n^{-C_4}).$$

Then,

$$\Pr\left[ f(X_t) \geq m \mid X_0 \in B \right] \geq 1 - n^{-\Omega(1)}$$

holds for some $\tau = O(\log n).$

**Proof.** Let $\Omega' = B \cup \{a, b\}$ be the state space with two special states $a$ and $b.$ We consider a Markov chain $(X'_t)_{t+1}^{\infty}$ on $\Omega'$ by

$$\Pr[X'_{t+1} = x \mid X'_t = y] = \begin{cases} 
\Pr[X_{t+1} = x \mid X_t = y] & \text{if } x, y \in B, \\
\Pr[X_{t+1} \not\in B \land f(X_{t+1}) < m \mid X_t = y] & \text{if } x = a \text{ and } y \in B, \\
\Pr[f(X_{t+1}) \geq m \mid X_t = y] & \text{if } x = b \text{ and } y \in B, \\
1 & \text{if } x = y \in \{a, b\}.
\end{cases}$$

In other words, the special state $a$ corresponds to the event “$f(x) < m$ and $x \not\in B$,” and $b$ does “$f(x) \geq m.$”

Suppose that $X'_0 \in B$ and let $\tau' = \min\{t : X'_t \not\in B\} > 0$ be the stopping time. Then, the above definition of $X'_t$ naturally yields a coupling $(X'_t, X'_t)_{t<\tau'}$ satisfying $X_t = X'_t$ for $t < \tau'.$

Let $f' : \Omega' \to \{0, \ldots, n\}$ be a function given by

$$f'(x) = \begin{cases} f(x) & \text{if } x \in B, \\
n & \text{if } x \in \{a, b\}. \end{cases}$$

Then, the Markov chain $(X'_t)$ on $\Omega'$ and the function $f'$ satisfies the conditions (1) and (2) of Lemma B.12. Hence, for some $\tau = O(\log n),$ it holds that $X'_\tau \in \{a, b\}.$ We insist that $X'_\tau = b,$ that is, $f(X_\tau) \geq m.$ Indeed, from the condition (3'), we have $\Pr[X'_\tau = a \mid X'_0 \in B] \leq \tau \cdot O(n^{-C_4}) \leq n^{-\Omega(1)}.$

We say a function $f : \{0, 1\}^M \to \mathbb{R}$ is monotone increasing if $f(x) \leq f(y)$ whenever $x = (x_1, \ldots, x_M), y = (y_1, \ldots, y_M) \in \{0, 1\}^M$ satisfies $x_i \leq y_i$ for every $i = 1, \ldots, M.$

**Lemma B.14** (The FKG inequality, Theorem 21.5 in [25]). For a given set $[M] = \{1, 2, \ldots, M\},$ let $I_1, I_2, \ldots, I_M$ be independent binary random variables. Then for any two monotone increase functions $f, g : \{0, 1\}^M \to \mathbb{R},$ it holds that

$$E[f(I)g(I)] \geq E[f(I)]E[g(I)],$$

where $I = (I_1, I_2, \ldots, I_M) \in \{0, 1\}^M.$
Lemma B.15 (Janson’s inequality (Theorem 21.12 in [25])). For a given set \([M] = \{1, 2, \ldots, M\}\), let \(I_1, I_2, \ldots, I_M\) be independent binary random variables. Let \((F_1, F_2, \ldots, F_N)\) be a family of \(N\) subsets of \([M]\) \((F_i \subseteq [M]\) for every \(i \in [N]\)) and

\[
Y := \sum_{i \in [M]} \prod_{e \in F_i} I_e.
\]

Then, it holds for any \(t \leq E[Y]\) that

\[
\Pr \left[ Y \leq E[Y] - t \right] \leq \exp \left( -\frac{t^2}{2\nabla} \right),
\]

where

\[
\nabla := \sum_{i \in N, j \in N: F_i \cap F_j \neq \emptyset} \mathbb{E} \left[ \left( \prod_{e \in F_i} I_e \right) \left( \prod_{e' \in F_j} I_{e'} \right) \right].
\]

Lemma B.16 (The Kim-Vu concentration (main Theorem in [33])). For a given set \([M] = \{1, 2, \ldots, M\}\), let \(I_1, I_2, \ldots, I_M\) be independent binary random variables. Let \(\mathcal{E} \subseteq 2^{[M]}\) be a collection of subsets of \([M]\) \((F \subseteq [M]\) for all \(F \in \mathcal{E}\)) and

\[
Y = \sum_{F \in \mathcal{E}} w(F) \prod_{e \in F} I_e,
\]

where \(w(F)\) are positive coefficients. For a subset \(A \subseteq [M]\), define \(Y_A\) as

\[
Y_A = \sum_{F \in \mathcal{E}: F \supseteq A} w(F) \prod_{e \in F \setminus A} I_e.
\]

If the polynomial \(Y\) has a degree of at most \(k\) (i.e., \(\max_{F \in \mathcal{E}} |F| \leq k\)), then the following holds for any positive \(\lambda > 1\):

\[
\Pr \left[ |Y - E[Y]| \geq \sqrt{k! \max_{A \subseteq [M]} E[Y_A]} \max_{A \subseteq [M]: A \neq \emptyset} E[Y_A](8\lambda)^k \right] \leq 2 \exp(2 + (k - 1) \log M - \lambda).
\]