Phase Transitions of Best-of-Two and Best-of-Three on Stochastic Block Models

Nobutaka Shimizu
The University of Tokyo, Japan
nobutaka_shimizu@mist.i.u-tokyo.ac.jp

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

Abstract

This paper 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 and discrete time step, 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). Previous studies have explored these processes on complete graphs and expander graphs, but we understand significantly less about their properties on graphs with more complicated structures.

In this paper, we study 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 density $q \le p$. We obtain two main results. First, if $p = \omega(\log n/n)$ and $r = q/p$ is a constant, we show that there is a phase transition in r with threshold r^* (specifically, $r^* = \sqrt{5} - 2$ for the Best-of-two, and $r^* = 1/7$ for the Best-of-three). If $r > r^*$, the process reaches consensus within $O(\log\log n + \log n/\log(np))$ steps for any initial opinion configuration with a bias of $\Omega(n)$. By contrast, if $r < r^*$, then there exists an initial opinion configuration with a bias of $\Omega(n)$ from which the process requires at least $2^{\Omega(n)}$ steps to reach consensus. Second, if p is a constant and $r > r^*$, we show that, for any initial opinion configuration, the process reaches consensus within $O(\log n)$ steps. To the best of our knowledge, this is the first result concerning multiple-choice voting for arbitrary initial opinion configurations on non-complete graphs.

2012 ACM Subject Classification Mathematics of computing → Stochastic processes; Mathematics of computing → Random graphs

Keywords and phrases Distributed Voting, Consensus Problem, Random Graph

Digital Object Identifier 10.4230/LIPIcs.DISC.2019.32

Related Version The full-version is [36], `https://arxiv.org/abs/1907.12212`.

Funding *Nobutaka Shimizu*: JST CREST Grant Number JPMJCR14D2 and JSPS KAKENHI Grant Number 19J12876, Japan
Takeharu Shiraga: JSPS KAKENHI Grant Number 17H07116 and 19K20214, Japan

Acknowledgements We would like to thank Colin Cooper, Nan Kang and Tomasz Radzik for helpful discussions. We also thank the anonymous reviewers for their helpful comments.

1 Introduction

This paper is concerned with *voting processes* on distributed networks. 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, e.g. social behavior, physical phenomena and biological systems [32, 30, 4]. In distributed computing, voting processes are known as a simple approach for consensus problems [20, 23].

33rd International Symposium on Distributed Computing (DISC 2019).
Editor: Jukka Suomela; Article No. 32; pp. 32:1–32:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1.1 Previous work

The synchronous *pull voting* (a.k.a. the *voter model*) is a simple and well-studied voting process [33, 25]. 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 [25] 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 exists 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[1]. Doerr et al. [19] 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[2]. 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 [8, 7, 10, 22]. For example, Ghaffari and Lengler [22] 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 of multiple-choice voting processes on non-complete graphs have considered expander graphs with an *initial bias*, i.e. a difference between the initial sizes of the largest and the second largest opinions. Cooper et al. [13] 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 λ_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 [14]. Recall that the Erdős-Rénri graph $G(n, p)$ is a graph on n vertices where each vertex pair is joined by an edge with probability p, independent of any other pairs. In [15], the authors studied the Best-of-two and the Best-of-three on regular expander graphs initially involving more than two opinions. In [3, 28], the authors studied multiple-choice voting processes on non-complete graphs with random initial configuration.

Recently, the Best-of-two on richer classes of graphs involving two opinions have been studied. Previous works proved interesting results which do not hold on complete graphs or expander graphs. Cruciani et al. [17] studied the Best-of-two on the *core periphery network*, namely 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. [18] 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 $b/a = O(n^{-0.5})$, 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 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] If the graph initially involves two possible opinions, this definition matches the rule described in Abstract.
[2] In this paper "with high probability" (w.h.p.) means probability at least $1 - n^{-c}$ for a constant $c > 0$.

1.2 Our results

This paper 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, 31], network analysis [5, 24] and machine learning [2, 1], 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, 18]. In this paper, we focus on the following model which admits two communities of equal size.

▶ **Definition 1** (Stochastic block model). *For $n \in \mathbb{N}$ and $p, q \in [0, 1]$ with $q \le p$, the stochastic block model $G(2n, p, q)$ is a graph on 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 θ, independent of any other edges, where*

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

Note that $G(2n, p, q)$ is not connected w.h.p. if $p = o(\log n/n)$ [21]. Throughout this paper, we assume $p = \omega(\log n/n)$, in which regime each community is connected w.h.p.

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

$$T_{\mathrm{cons}}(A) := \min\left\{t \ge 0 \,:\, A^{(t)} \in \{\emptyset, V\},\ A^{(0)} = A\right\}.$$

We obtain two main results, described below.

Result I: phase transition

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 [19]. 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 2** (Phase transition of the Best-of-three on $G(2n, p, q)$). *Consider the Best-of-three on $G(2n, p, q)$ such that $r := \frac{q}{p}$ is a constant.*

(i) *If $r > \frac{1}{7}$, then $G(2n, p, q)$ w.h.p. satisfies the following property: There exist two positive constants $C, C' > 0$ such that*

$$\forall A \subseteq V \text{ of } \big||A| - |V \setminus A|\big| = \Omega(n) :$$
$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C\left(\log\log n + \frac{\log n}{\log(np)}\right)\right] \ge 1 - O\left(n^{-C'}\right).$$

(ii) *If $r < \frac{1}{7}$, then $G(2n, p, q)$ w.h.p. satisfies the following property: There exist a set $A \subseteq V$ with $\big||A| - |V \setminus A|\big| = \Omega(n)$ and two positive constants $C, C' > 0$ such that*

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \ge \exp(Cn)\right] \ge 1 - O\left(n^{-C'}\right).$$

▶ **Theorem 3** (Phase transition of the Best-of-two on $G(2n,p,q)$). *Consider the Best-of-two on $G(2n,p,q)$ such that $r := \frac{q}{p}$ is a constant.*

(i) *If $r > \sqrt{5}-2$, then $G(2n,p,q)$ w.h.p. satisfies the following property: There exist two positive constants $C, C' > 0$ such that*

$$\forall A \subseteq V \text{ of } \big||A| - |V \setminus A|\big| = \Omega(n) :$$
$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C\left(\log\log n + \frac{\log n}{\log(np)}\right)\right] \ge 1 - O\left(n^{-C'}\right).$$

(ii) *If $r < \sqrt{5}-2$, then $G(2n,p,q)$ w.h.p. satisfies the following property: There exist a set $A \subseteq V$ with $\big||A| - |V \setminus A|\big| = \Omega(n)$ and two positive constants $C, C' > 0$ such that*

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \ge \exp(Cn)\right] \ge 1 - O\left(n^{-C'}\right).$$

Note that the upper bound $T_{\mathrm{cons}}(A) = O(\log\log n + \log n/\log(np))$ is tight up to a constant factor if $\log n/\log(np) \ge \log\log n$. To see this, observe that there exists an $A \subseteq V$ such that $T_{\mathrm{cons}}(A)$ 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. [21].

We also note that the consensus time of the pull voting is $O(\mathrm{poly}(n))$ for any non-bipartite graph [25]. To the best of our knowledge, Theorem 2 and Theorem 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

The most central topic in voter processes is the *symmetry breaking*, i.e. 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 worst-case consensus time of multiple-choice voting processes deal with complete graphs [19, 7, 10, 22]. All previous work on non-complete graphs has involved some special bias setting (e.g. an initial bias [13, 14, 15], or a random initial opinion configuration [3, 18, 28]). In this paper, we present the following first worst-case analysis of non-complete graphs.

▶ **Theorem 4** (Worst-case analysis of the Best-of-three on $G(2n,p,q)$). *Consider the Best-of-three on $G(2n,p,q)$ such that p and q are positive constants. If $\frac{q}{p} > \frac{1}{7}$, then $G(2n,p,q)$ w.h.p. satisfies the following property: There exist two positive constants $C, C' > 0$ such that*

$$\forall A \subseteq V \,:\, \mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C\log n\right] \ge 1 - O\left(n^{-C'}\right).$$

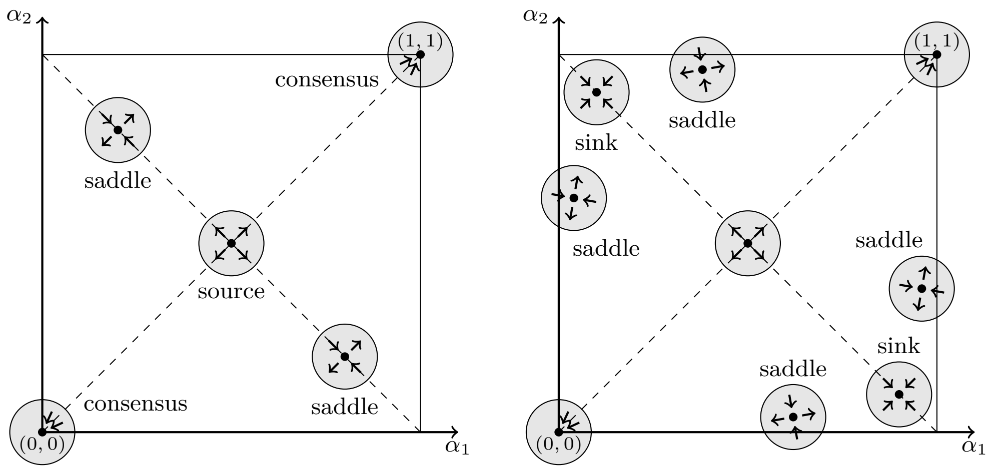
▶ **Theorem 5** (Worst-case analysis of the Best-of-two on $G(2n,p,q)$). *Consider the Best-of-two on $G(2n,p,q)$ such that p and q are positive constants. If $\frac{q}{p} > \sqrt{5}-2$, then $G(2n,p,q)$ w.h.p. satisfies the following property: There exist two positive constants $C, C' > 0$ such that*

$$\forall A \subseteq V \,:\, \mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C\log n\right] \ge 1 - O\left(n^{-C'}\right).$$

Based on these theorems, an immediate but important corollary follows.

▶ **Corollary 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 $G(n,p)$ has been extensively studied in previous works but these works put aforementioned assumptions on initial bias.



(a) $r > 1/7$ **(b)** $r < 1/7$

Figure 1 Four types of zero areas are illustrated. The sink areas do not appear if $r > 1/7$.

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, $|A'| = \sum_{v \in V} \mathbb{1}_{v \in A'}$ is the sum of independent random variables; thus, $|A'|$ concentrates on $\mathbf{E}[|A'| \mid A]$.

If the underlying graph is a complete graph, the state space can be regarded as $\{0, \dots, n\}$ (each state represents $|A|$). Therefore, $\mathbf{E}[|A'| \mid A]$ is expressed as a function of $|A|$, e.g. in the Best-of-two, $\mathbf{E}[|A'| \mid A] = f(|A|) := |A|\big(1 - (\frac{|A|}{n})^2\big) + (n - |A|)\big(\frac{|A|}{n}\big)^2 = n\big(3\big(\frac{|A|}{n}\big)^2 - 2\big(\frac{|A|}{n}\big)^3\big)$. Doerr et al. [19] exploited this idea for the Best-of-two and obtained the worst-case analysis for the consensus time on complete graphs. Somewhat interestingly, we also have $\mathbf{E}[|A'| \mid A] = f(|A|)$ in the Best-of-three.

Cooper et al. [13] extended this approach to the Best-of-two on regular expander graphs. Specifically, they proved that $\mathbf{E}[|A'| \mid A] = f(|A|) \pm O(\epsilon)$ for all $A \subseteq V$, where $\epsilon = \epsilon(n, \lambda_2) = o(n)$ is some function using the *expander mixing lemma*. This argument assumes an initial bias of size $\Omega(\epsilon)$. In another paper, Cooper et al. [14] improved this technique and proved more sophisticated results that hold for general (i.e. not necessarily regular) expander graphs.

In this paper, we consider $G(2n, p, q)$ 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 $\mathbf{E}[|A_i'| \mid A] = F_i(|A_1|, |A_2|) \pm O(\sqrt{n/p})$ for all $A \subseteq V$ in the Best-of-three, where $F_i : \mathbb{N}^2 \to \mathbb{N}$ is some function $(i = 1, 2)$. See (2) for details. We show the same result for the Best-of-two. Here, our key tool is the concentration method, specifically the Janson inequality [21] and the Kim-Vu concentration [29].

High-level proof sketch

Consider the Best-of-three on $G(2n, p, q)$, and let $A^{(0)}, A^{(1)}, \dots$ be a sequence of random vertex subsets determined by $A^{(t+1)} := (A^{(t)})'$ for each $t \geq 0$. Consider a stochastic process $\boldsymbol{\alpha}^{(t)} = (\alpha_1^{(t)}, \alpha_2^{(t)}) \in [0, 1]^2$ where $\alpha_i^{(t)} = |A^{(t)} \cap V_i|/n$ for $i = 1, 2$. Our technical result in the previous paragraph approximates the stochastic process $\boldsymbol{\alpha}^{(t)}$ by the *deterministic* process $\mathbf{a}^{(t)}$ defined as $\mathbf{a}^{(t+1)} = H(\mathbf{a}^{(t)})$ and $\boldsymbol{\alpha}^{(0)} = \mathbf{a}^{(0)}$ for some function $H : [0,1]^2 \to [0,1]^2$ (See (4) and Figure 2). The function H induces a two-dimensional dynamical system, which we call the *induced dynamical system*. Using this, we obtain two results concerning $\boldsymbol{\alpha}^{(t)}$.

First, we show that, for any initial configuration, the process reaches one of the *zero areas* (a neighbor of a fixed point of H) within a constant number of steps. To show this, in addition to the approximation result, we used the theory of *competitive dynamical systems* [26].

Second, we characterize the behavior of $\boldsymbol{\alpha}^{(t)}$ in zero areas. The zero areas depend only on $r = q/p$, and are 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 is a constant by using techniques of [19]. Intuitively speaking, in these two kinds of areas, there are drifts towards outside. To apply the techniques of [19], we show that $\mathbf{Var}[|A'_i|] = \Omega(n)$ in the area if p is constant, which leads to our worst-case analysis result. Indeed, any previous works working on expander graphs did not investigate the worst-case due to the lack of variance estimation.

These arguments also enable us to study the Best-of-two process, which implies Theorem 3.

1.4 Related work

The consensus time of the pull voting process is investigated via its dual process, known as *coalescing random walks* [25, 12, 16]. Recently coalescing random walks have been extensively studied, including the relationship with properties of random walks such as the hitting time and the mixing time [27, 34].

Other studies have focused on voting processes with more general updating rules. Cooper and Rivera [16] 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 [35] 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 using a potential technique.

Organization

First we set notation and precise definition of the Best-of-three in Section 2. After explaining key properties of the stochastic block model in Section 3, we show some auxiliary results of the induced dynamical system in Section 4. Then we derive Theorems 2 and 4 in Section 5. Our general framework of voting processes and results of the general induced dynamical systems are given in Section 6 and Section 7, respectively. Due to the page limitation, we omit detailed proofs and the discussion of the Best-of-two. See the full paper [36] for details.

2 Best-of-three voting process

For an $\ell \in \mathbb{N}$, let $[\ell] := \{1, 2, \dots, \ell\}$. 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 $\deg_S(v) = |S \cap N(v)|$. Here, we study the Best-of-three with two possible opinions from $\{1, 2\}$.

▶ **Definition 7** (Best-of-three). *Let $G = (V, E)$ be a graph where each vertex holds an opinion from $\{1, 2\}$. Let*

$$f^{\mathrm{Bo3}}(x) := \binom{3}{3}x^3 + \binom{3}{2}x^2(1-x) = 3x^2 - 2x^3.$$

For the set A of vertices holding opinion 1, let A' denote the set of vertices that hold opinion 1 after an update. In the Best-of-three, $A' = \{v \in V : X_v = 1\}$ where $(X_v)_{v\in V}$ are independent binary random variables satisfying

$$\mathbf{Pr}[X_v = 1] = f^{\mathrm{Bo3}}\left(\frac{\deg_A(v)}{\deg(v)}\right).$$

For a given vertex subset $A^{(0)} \subseteq V$, we are interested in the behavior of the Markov chain $(A^{(t)})_{t=0}^{\infty}$, i.e. the sequence of random vertex subsets determined by $A^{(t+1)} := (A^{(t)})'$ for each $t \geq 0$. Let $A_i := V_i \cap A$ for $A \subseteq V$ and $i = 1, 2$. Since $|A_i'| = \sum_{v\in V_i} X_v$, the Hoeffding bound implies that the following holds w.h.p for $i = 1, 2$:

$$||A_i'| - \mathbf{E}[|A_i'|]| = O(\sqrt{n \log n}). \tag{1}$$

3 Concentration result for the stochastic block model

In this paper, we consider the Best-of-three on the stochastic block model $G(2n, p, q)$ (Definition 1). Then, $\mathbf{E}[|A_i'|]$ in (1) is a random variable since $G(2n, p, q)$ is a random graph. Here, our key ingredient is the following general concentration result for $G(2n, p, q)$.

▶ **Definition 8** (f-good $G(2n, p, q)$). *For a given function $f : [0, 1] \to [0, 1]$, we say $G(2n, p, q)$ is f-good if $G(2n, p, q)$ satisfies the following properties.*

(P1) *It is connected and non-bipartite.*

(P2) *A positive constant C_1 exists 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) *A positive constant C_2 exists such that, for all $A \subseteq V$, $S \in \{A, V \setminus A, V\}$ 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 9** (Main technical theorem). *Suppose that $f : [0, 1] \to [0, 1]$ is a polynomial function with constant degree, $p = \omega(\log n/n)$ and $q \geq \log n/n^2$. Then $G(2n, p, q)$ is f-good w.h.p.*

Note that the proof of 1 is not difficult since $p = \omega(\log n/n)$ and $q \geq \log n/n^2$ [21]. Proving 2 and 3, however, is more challenging: see the full version of this paper [36].

From Theorem 9, $G(2n, p, q)$ is f^{Bo3}-good w.h.p. Hence, we consider the Best-of-three on an f^{Bo3}-*good* $G(2n, p, q)$. From 2 and 3, we have

$$\mathbf{E}[|A_i'|] = \sum_{v\in V_i} f^{\mathrm{Bo3}}\left(\frac{\deg_A(v)}{\deg(v)}\right) = n f^{\mathrm{Bo3}}\left(\frac{|A_i|p + |A_{3-i}|q}{n(p+q)}\right) \begin{cases} \pm O\left(\sqrt{\frac{n}{p}}\right) \\ +O\left(|A|\sqrt{\frac{\log n}{np}}\right) \end{cases} \tag{2}$$

for all $A \subseteq V$ and $i = 1, 2$. Here, we remark that 3 is stronger than 2 if $|A|$ is sufficiently small. This property will play a key role in the proof of Proposition 14.

Idea of the proof of Theorem 9

We consider the property 2. Note that we may assume $f(x) = x^k$ for some constant k w.l.o.g. since it suffices to obtain the concentration result for each term of f. For simplicity, let us exemplify our idea on the special case of $k = 3$. It is known that $\deg(v) = n(p+q) \pm O(\sqrt{np\log n})$ holds for all $v \in V$ w.h.p. (see, e.g. [21]). This implies that $\sum_{v\in S}\left(\frac{\deg_A(v)}{\deg(v)}\right)^3 = \frac{1\pm O(\sqrt{\log n/np})}{(n(p+q))^3}\cdot\sum_{v\in S}\deg_A(v)^3$ holds for all $S, A \subseteq V$. Indeed, it is not difficult to see that the term $O(\sqrt{\log n/np})$ can be improved to $O(\sqrt{1/np})$.

The core of the proof is the concentration of $\sum_{v\in S}\deg_A(v)^3$. Note that $\sum_{v\in S}\deg_A(v) = \sum_{v\in S}\sum_{a\in A}\mathbb{1}_{\{s,a\}\in E}$ counts the number of cut edges between S and A. For fixed S and A, the Chernoff bound yields the concentration of it since each edge appears independently. Similarly, the summation $\sum_{v\in S}\deg_A(v)^3 = \sum_{v\in S}\sum_{a,b,c\in A}\mathbb{1}_{\{v,a\},\{v,b\},\{v,c\}\in E}$ counts the number of "crossing stars" between S and A. However, the Chernoff bound does not work here due to the dependency of the appearance of crossing stars. Fortunately, we can obtain a strong lower bound using the Janson inequality as follows: For $S, A, B, C \subseteq V$, let $W(S; A, B, C) := \sum_{v\in S}\deg_A(v)\deg_B(v)\deg_C(v)$. From the Janson inequality and the union bound on $S, A, B, C \subseteq V$, we can show that $W(S; A, B, C) \ge \mathbf{E}[W(S; A, B, C)] - O(n^{3.5}p^{2.5})$ holds for all $S, A, B, C \subseteq V$ w.h.p. On the other hand, it is easy to check that

$$\begin{aligned} W(S; A, B, C) = {} & W(V; V, V, V) - W(V; V, V, V\setminus C) - W(V; V, V\setminus B, C) \\ & - W(V; V\setminus A, B, C) - W(V\setminus S; A, B, C). \end{aligned}$$

The Kim-Vu concentration yields $W(V; V, V, V) \le \mathbf{E}[W(V; V, V, V)] + O(n^{3.5}p^{2.5})$ since we do not consider the union bound here. For the other terms, we apply the lower bound by the Janson inequality. Then, we have a strong concentration result that $\sum_{v\in S}\deg_A(v)^3 = W(S; A, A, A) = \mathbf{E}[W(S; A, A, A)] \pm O(n^{3.5}p^{2.5})$ holds for all $S, A \subseteq V$ w.h.p. Finally, we estimate the gap between $\mathbf{E}[W(S\cap V_i, A, A, A)]$ and $|S\cap V_i|(|A_i|p + |A_{3-i}|q)^3$. See the full version [36] for details.

4 Induced dynamical system

Let $\alpha_i := \frac{|A_i|}{n}$, $\alpha'_i := \frac{|A'_i|}{n}$ and $r := \frac{q}{p}$. Suppose that r is a constant. Then, for an f^{Bo3}-good $G(2n, p, q)$, it holds w.h.p. that

$$\left|\alpha'_i - f^{\mathrm{Bo3}}\left(\frac{\alpha_i + r\alpha_{3-i}}{1+r}\right)\right| = O\left(\sqrt{\frac{1}{np}} + \sqrt{\frac{\log n}{n}}\right) \tag{3}$$

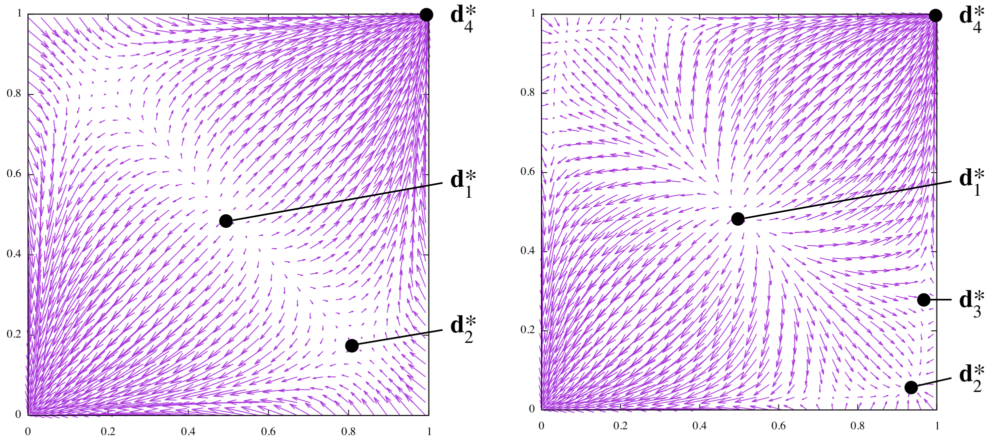
for all $A \subseteq V$ and $i = 1, 2$ since (1) and (2) hold.

Throughout this paper, we use $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ and $\boldsymbol{\alpha}' = (\alpha'_1, \alpha'_2)$ as vector-valued random variables. Equation (3) leads us to the dynamical system H, where we define $H : \mathbb{R}^2 \to \mathbb{R}^2$ as

$$H : \mathbf{a} \mapsto (H_1(\mathbf{a}), H_2(\mathbf{a})), \tag{4}$$

and $H_i(a_1, a_2) := f^{\mathrm{Bo3}}\left(\frac{a_i + ra_{3-i}}{1+r}\right)$.

By combining (3) with the Lipschitz condition, it is not difficult to show the following result; see Section 6 for the proof.



(a) $r = 1/6$ **(b)** $r = 1/9$

Figure 2 The induced dynamical system H of (4). The points $\mathbf{d}_i^*$ are the fixed points given in (7). Here, the horizontal and vertical axes correspond to α_1 and α_2, respectively. We can observe two sink points in (b), but none in (a).

▶ **Theorem 10.** *Consider the Best-of-three on an f^{Bo3}-good $G(2n, p, q)$, starting with the vertex set $A^{(0)} \subseteq V$ holding opinion 1. Let $(\boldsymbol{\alpha}^{(t)})_{t=0}^{\infty}$ be a stochastic process given by $\boldsymbol{\alpha}^{(t)} = (\alpha_1^{(t)}, \alpha_2^{(t)})$ and $\alpha_i^{(t)} = |A^{(t)} \cap V_i|/n$. Let H be the mapping (4) and define $(\mathbf{a}^{(t)})_{t=0}^{\infty}$ as*

$$\begin{cases} \mathbf{a}^{(0)} = \boldsymbol{\alpha}^{(0)}, \\ \mathbf{a}^{(t+1)} = H(\mathbf{a}^{(t)}). \end{cases} \tag{5}$$

Then there exists a positive constant $C > 0$ such that

$$\forall 0 \le t \le n^{o(1)}, \forall A^{(0)} \subseteq V :$$
$$\mathbf{Pr}\left[\|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_\infty \le C^t \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right)\right] \ge 1 - n^{-\Omega(1)}.$$

Broadly speaking, Theorem 10 approximates the behavior of $\boldsymbol{\alpha}^{(t)}$ by the orbit $\mathbf{a}^{(t)}$ of the corresponding dynamical system H. We call the mapping H the *induced dynamical system*. Indeed, the same results as (2) hold for the Best-of-two voting. Therefore, analogous results of Theorem 10 hold, which enable us to analyze the Best-of-two on $G(2n, p, q)$ via its induced dynamical system. The dynamical system H of (4) is illustrated in Figure 2.

To make the calculations more convenient, we change the coordinate of H by

$$\boldsymbol{\delta} = (\delta_1, \delta_2) := (\alpha_1 - \alpha_2, \alpha_1 + \alpha_2 - 1).$$

Note that δ_1 and δ_2 axes are corresponding to the dotted lines of Figure 1. Let $u := \frac{1-r}{1+r}$. Then we have

$$\mathbf{E}[\delta_i' \mid A] = T_i(\delta_1, \delta_2) + O\left(\frac{1}{\sqrt{np}}\right),$$

where

$$T_1(d_1, d_2) := \frac{ud_1}{2}\left(3 - (ud_1)^2 - 3d_2^2\right), \; T_2(d_1, d_2) := \frac{d_2}{2}\left(3 - 3(ud_1)^2 - d_2^2\right).$$

This suggests another dynamical system $T(\mathbf{d}) = (T_1(\mathbf{d}), T_2(\mathbf{d}))$. Here, we use $\mathbf{d} = (d_1, d_2)$ as a specific point and $\boldsymbol{\delta} = (\delta_1, \delta_2)$ as a vector-valued random variable. Consider $\boldsymbol{\delta}^{(t)} = (\delta_1^{(t)}, \delta_2^{(t)})$ and $(\mathbf{d}^{(t)})_{t=0}^{\infty}$, where $\mathbf{d}^{(0)} = \boldsymbol{\delta}^{(0)}$ and $\mathbf{d}^{(t+1)} = T(\mathbf{d}^{(t)})$ for each $t \geq 0$. From Theorem 10, it holds w.h.p. that

$$\|\boldsymbol{\delta}^{(t)} - \mathbf{d}^{(t)}\|_\infty \leq C^t \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} \right) \tag{6}$$

for sufficiently large constant $C > 0$, any $0 \leq t \leq n^{o(1)}$ and any initial configuration $A^{(0)} \subseteq V$. For notational convenience, we use $\boldsymbol{\delta}' := \boldsymbol{\delta}^{(t+1)}$ for $\boldsymbol{\delta} = \boldsymbol{\delta}^{(t)}$. Similarly, we refer $\mathbf{d}'$ to $T(\mathbf{d})$.

Note that $\boldsymbol{\delta}$ satisfies $|\delta_1| + |\delta_2| \leq 1$. In addition, the dynamical system T is symmetric: Precisely, $T_1(\pm d_1, \mp d_2) = \pm T_1(d_1, d_2)$ and $T_2(\pm d_1, \mp d_2) = \mp T_2(d_1, d_2)$ hold. In Lemma 11, we assert that the sequence $(\mathbf{d}^{(t)})_{t=0}^{\infty}$ is closed in

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

From now on, we focus on S and consider the behavior of $\boldsymbol{\delta}$ around fixed points. A straightforward calculation shows that $\mathbf{d}' = \mathbf{d} \in S$ if and only if $\mathbf{d} \in \{\mathbf{d}_1^*, \mathbf{d}_2^*, \mathbf{d}_3^*, \mathbf{d}_4^*\}$, where

$$\mathbf{d}_i^* := \begin{cases} (0,0) & \text{if } i = 1, \\ \left(\sqrt{\frac{3u-2}{u^3}}, 0\right) & \text{if } i = 2 \text{ and } u \geq \frac{2}{3}, \\ \left(\sqrt{\frac{1}{4u^3}}, \sqrt{\frac{4u-3}{4u}}\right) & \text{if } i = 3 \text{ and } u \geq \frac{3}{4}, \\ (0,1) & \text{if } i = 4. \end{cases} \tag{7}$$

Here, we provide auxiliary results needed for the proofs of Theorems 2 and 4. Section 7 contains generalized form of these results.

For $\mathbf{x} \in \mathbb{R}^2$ and $\epsilon > 0$, let $B(\mathbf{x}, \epsilon) = \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{y}\|_\infty < \epsilon\}$ be the open ball. For $\mathbf{d} = (d_1, d_2) \in \mathbb{R}^2$, let $\langle \mathbf{d} \rangle_+ := (|d_1|, |d_2|) \in \mathbb{R}^2$.

▶ **Lemma 11** (S is closed). *For any $\mathbf{d} \in S$, it holds that $\mathbf{d}' \in S$.*

▶ **Proposition 12** (Orbit convergence). *For any sequence $(\mathbf{d}^{(t)})_{t=0}^{\infty}$, $\lim_{t\to\infty} \langle \mathbf{d}^{(t)} \rangle_+ = \mathbf{d}_i^*$ for some $i \in \{1,2,3,4\}$. In addition, if $u < \frac{3}{4}$ and a positive constant $\kappa > 0$ exists such that the initial point $\mathbf{d}^{(0)} = (d_1^{(0)}, d_2^{(0)}) \in S$ satisfies $|d_2^{(0)}| > \kappa$, then $\lim_{t\to\infty} \langle \mathbf{d}^{(t)} \rangle_+ = \mathbf{d}_4^*$.*

▶ **Proposition 13** (Dynamics around $\mathbf{d}_2^*$). *Consider the Best-of-three on an f^{Bo3}-good $G(2n, p, q)$ such that $r = q/p < 1/7$ is a constant. Then there exists a positive constant $\epsilon = \epsilon(r)$ satisfying*

$$\mathbf{Pr}\left[\langle \boldsymbol{\delta}' \rangle_+ \notin B(\mathbf{d}_2^*, \epsilon) \,\middle|\, \langle \boldsymbol{\delta} \rangle_+ \in B(\mathbf{d}_2^*, \epsilon)\right] \leq \exp(-\Omega(n)).$$

In particular, $T_{\mathrm{cons}}(A) = \exp(\Omega(n))$ w.h.p. for any A satisfying $\langle \boldsymbol{\delta} \rangle_+ \in B(\mathbf{d}_2^*, \epsilon)$.

▶ **Proposition 14** (Towards consensus). *Consider the Best-of-three on an f^{Bo3}-good $G(2n, p, q)$ such that $r = q/p$ is a constant. Then, there exists a universal constant $\epsilon = \epsilon(r) > 0$ satisfying the following: $T_{\mathrm{cons}}(A) \leq O(\log\log n + \log n/\log(np))$ holds w.h.p. for all $A \subseteq V$ with $\min\{|A|, 2n - |A|\} \leq \epsilon n$.*

▶ **Proposition 15** (Escape from fixed points). *Consider the Best-of-three on an f^{Bo3}-good $G(2n, p, q)$ such that p and q are constants. If $q/p > 1/7$ and $|\delta_2^{(0)}| = o(1)$, then it holds w.h.p. that $|\delta_2^{(\tau)}| > \kappa$ for some $\tau = O(\log n)$ and some constant $\kappa > 0$.*

Intuitive explanations for Propositions 13 to 15

In Propositions 13 to 15, we consider the behavior of $\boldsymbol{\alpha}^{(t)}$ around the fixed points (7). Let H be the induced dynamical system and let J be the Jacobian matrix of H at a fixed point $\mathbf{a}^*$ with two eigenvalues λ_1, λ_2. If the eigenvectors are linearly independent, we can rewrite J as $J = U^{-1}\Lambda U$, where $\Lambda := \mathrm{diag}(\lambda_1, \lambda_2)$ and U is some nonsingular matrix. Let $\boldsymbol{\beta} := U(\boldsymbol{\alpha} - \mathbf{a}^*)$. Roughly speaking, if $\boldsymbol{\alpha}$ is closed to $\mathbf{a}^*$, the Taylor expansion at $\mathbf{a}^*$ (i.e. $H(\boldsymbol{\alpha}) \approx \mathbf{a}^* + J(\boldsymbol{\alpha} - \mathbf{a}^*)$) yields

$$\mathbf{E}[\boldsymbol{\beta}' \mid A] = U(\mathbf{E}[\boldsymbol{\alpha}' \mid A] - \mathbf{a}^*) \approx U(H(\boldsymbol{\alpha}) - \mathbf{a}^*) \approx \Lambda \boldsymbol{\beta}.$$

In other words, $\beta_i' \approx \lambda_i \beta_i$. If $\max\{|\lambda_1|, |\lambda_2|\} < 1 - c$ for some constant $c > 0$, we might expect that $\|\boldsymbol{\beta}\| = \Theta(\|\boldsymbol{\alpha} - \mathbf{a}^*\|)$ is likely to keep being small. Here, we do not restrict this argument on the Best-of-three. We will prove Proposition 19, which is a generalized version of Proposition 13. If $\max\{|\lambda_1|, |\lambda_2|\} > 1 + c$ for some constant $c > 0$, the norm $\|\boldsymbol{\beta}\|$ seems to become large in a small number of steps. We will exploit this insight and prove Proposition 25, which immediately implies Proposition 15. Indeed, for consensus areas (i.e. $\mathbf{a}^* \in \{(0,0),(1,1)\}$), the induced dynamical systems of the Best-of-three and the Best-of-two satisfy $\lambda_1 = \lambda_2 = 0$. Then, the Taylor expansion yields $\|\boldsymbol{\alpha}' - \mathbf{a}^*\| \approx O(\|\boldsymbol{\alpha} - \mathbf{a}^*\|^2)$. This observation and the property 3 lead to the proof of Proposition 20 as well as Proposition 14.

5 Derive Theorems 2 and 4

Here, we prove Theorems 2 and 4 using Propositions 12 to 15.

Proof of Theorem 2. If $r > \frac{1}{7}$ and $A^{(0)} \subseteq V$ satisfies $\big||A^{(0)}| - n\big| = \Omega(n)$, then we have $|d_2^{(0)}| = |\delta_2^{(0)}| > \kappa$ for some constant $\kappa > 0$. Next, for any constant $\epsilon > 0$, Proposition 12 implies $\langle \mathbf{d}^{(l)} \rangle_+ \in B(\mathbf{d}_4^*, \epsilon)$ for some constant $l = l(\epsilon)$. From (6), we have $\langle \boldsymbol{\delta}^{(l)} \rangle_+ \in B(\mathbf{d}_4^*, \epsilon)$ for sufficiently large n. Set ϵ be the constant mentioned in Proposition 14. Then, from Proposition 14, it holds w.h.p. that $T_{\mathrm{cons}}(A^{(0)}) \leq l + T_{\mathrm{cons}}(A^{(l)}) \leq O(\log\log n + \log n / \log(np))$.

If $r < \frac{1}{7}$, Proposition 13 yields $T_{\mathrm{cons}}(A^{(0)}) \geq \exp(\Omega(n))$ w.h.p. for any $A^{(0)} \subseteq V$ with $\boldsymbol{\delta}^{(0)} \in B(\mathbf{d}_2^*, \epsilon)$, where $\epsilon > 0$ is the constant from Proposition 13, which completes the proof of ii. ◀

Proof of Theorem 4. If $|\delta^{(0)}| = o(1)$, then Proposition 15 yields that $|\delta^{(\tau)}| > \kappa$ for some constant $\kappa > 0$ and some $\tau = O(\log n)$. Then, from Theorem 2, we have $T_{\mathrm{cons}}(A^{(\tau)}) \leq O(\log\log n + \log n / \log(np))$. Thus, $T_{\mathrm{cons}}(A^{(0)}) \leq \tau + T_{\mathrm{cons}}(A^{(\tau)}) \leq O(\log n)$. ◀

6 Polynomial voting processes

Using Theorem 9, we can prove the same results as Theorem 10 for various models including the Best-of-two. Hence, in this paper, we do not restrict our interest to the Best-of-three: Instead, we prove general results that hold for *polynomial voting process* on $G(2n, p, q)$.

▶ **Definition 16** ((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*

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

In other words, for $i = 1, 2$,

$$\mathbf{Pr}[v \in A' \mid A, v \text{ has opinion } i] = f_i\left(\frac{\deg_A(v)}{\deg(v)}\right).$$

Polynomial voting process includes several known voting models including the Best-of-two, the Best-of-three, and so on. For example, $f_1(x) = f_2(x) = f^{\mathrm{Bo3}}(x) = 3x^2 - 2x^3$ for the Best-of-three. For the Best-of-two, $f_1(x) = 2x(1-x)$ and $f_2(x) = x^2$. We can define induced dynamical system for any polynomial voting process on $G(2n, p, q)$ via the following result:

▶ **Theorem 17** (Theorem 10 for polynomial voting processes). *Let f_1 and f_2 be polynomials with constant degree. Consider an (f_1, f_2)-polynomial voting process, on an f_1-good and f_2-good $G(2n, p, q)$ starting with 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 $(\boldsymbol{\alpha}^{(t)})_{t=0}^{\infty}$, where $\boldsymbol{\alpha}^{(t)} = (\alpha_1^{(t)}, \alpha_2^{(t)})$ and $\alpha_i^{(t)} = |A^{(t)} \cap V_i|/n$. Define a mapping $H = (H_1, H_2)$ as*

$$H_i(a_1, a_2) = a_i f_1\left(\frac{a_i + ra_{3-i}}{1+r}\right) + (1 - a_i) f_2\left(\frac{a_i + ra_{3-i}}{1+r}\right) \quad \text{for } i = 1, 2.$$

Define $(\mathbf{a}^{(t)})_{t=0}^{\infty}$ as $\mathbf{a}^{(0)} = \boldsymbol{\alpha}^{(0)}$ and $\mathbf{a}^{(t+1)} = H(\mathbf{a}^{(t)})$ for each $t \geq 0$. Then, there exists a constant $C > 0$ such that

$$\forall 0 \leq t \leq n^{o(1)}, \forall A^{(0)} \subseteq V:$$
$$\mathbf{Pr}\left[\|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_\infty \leq C^t\left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right)\right] \geq 1 - n^{-\Omega(1)}.$$

Remark that the mapping H of Theorem 17 is the induced dynamical system.

Proof. For any polynomial voting process, the cardinality $|A_i'| = \sum_{v \in V_i} X_v$ is the sum of independent random variables. Thus, if we fix $A \subseteq V$, the Hoeffding bound implies that (1) holds w.h.p. Since

$$\mathbf{E}[|A_i'|] = \sum_{v \in V_i} \mathbf{E}[X_v] = \sum_{v \in A_i} f_1\left(\frac{\deg_A(v)}{\deg(v)}\right) + \sum_{v \in V \setminus A_i} f_2\left(\frac{\deg_A(v)}{\deg(v)}\right),$$

the property 2 and (1) lead to

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

for some constant $C_1 > 0$.

Note that the function H satisfies the Lipschitz condition. Hence, a positive constant C_2 exists such that

$$\|H(\mathbf{x}) - H(\mathbf{y})\|_\infty \leq C_2\|\mathbf{x} - \mathbf{y}\|_\infty$$

holds for any $\mathbf{x}, \mathbf{y} \in [0,1]^2$. Let $\boldsymbol{\alpha}^{(t)} = (\alpha_1^{(t)}, \alpha_2^{(t)})$ be the vector-valued stochastic process and $\mathbf{a}^{(t)} = (\mathbf{a}_1^{(t)}, \mathbf{a}_2^{(t)})$ be the vector sequence given in (5). Then, we have

$$\begin{aligned}
\|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_\infty &= \|\boldsymbol{\alpha}^{(t)} - H(\boldsymbol{\alpha}^{(t-1)}) + H(\boldsymbol{\alpha}^{(t-1)}) - H(\mathbf{a}^{(t-1)})\|_\infty \\
&\leq \|\boldsymbol{\alpha}^{(t)} - H(\boldsymbol{\alpha}^{(t-1)})\|_\infty + C_2\|\boldsymbol{\alpha}^{(t-1)} - \mathbf{a}^{(t-1)}\|_\infty \\
&\leq C_2\|\boldsymbol{\alpha}^{(t-1)} - \mathbf{a}^{(t-1)}\|_\infty + C_1\left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right) \\
&\leq C^t\left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right),
\end{aligned}$$

where C is sufficiently large constant. ◀

7 Results of general induced dynamical systems

Now let us focus on the orbit $(\boldsymbol{\alpha}^{(t)})_{t=1}^{\infty}$ such that $H(\boldsymbol{\alpha}^{(0)}) = \boldsymbol{\alpha}^{(0)}$ holds, where H is the induced dynamical system. In this case, Theorem 17 does not provide enough information about the dynamics. In dynamical system theory, a natural approach for the local behavior around fixed points is to consider the *Jacobian matrix*. Recall that, the Jacobian matrix J of a function $H : \mathbf{x} \mapsto (H_1(\mathbf{x}), H_2(\mathbf{x}))$ at $\mathbf{a} \in \mathbb{R}^2$ is a 2×2 matrix given by

$$J = \left(\frac{\partial H_i}{\partial x_j}(\mathbf{a}) \right)_{i,j \in [2]}.$$

In the following subsections, we will investigate the local dynamics from the viewpoint of the maximum singular value and eigenvalue of the Jacobian matrix.

In contrast to the local dynamics, it is quite difficult to predicate the orbit of general dynamical systems since some of them exhibits so-called chaos phenomenon. Therefore, the proof of the orbit convergence (e.g. Proposition 12) is not trivial. Fortunately, the induced dynamical system of the Best-of-three on $G(2n, p, q)$ is *competitive*, a well-known nice property for predicting the future orbit [26]. We can show Proposition 12 using known results of competitive dynamical systems. The same argument leads to the orbit convergence for the Best-of-two. Details are presented in the full version [36].

7.1 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 $A^\top A$.

▶ **Definition 18** (sink point). *For a dynamical system H, a fixed point $\mathbf{a}^* \in \mathbb{R}^2$ is* sink *if the Jacobian matrix J at $\mathbf{a}^*$ satisfies $\sigma_{\max} < 1$, where $\sigma_{\max}$ is the largest singular value of J.*

▶ **Proposition 19.** *Consider an (f_1, f_2)-polynomial voting process on an f_1-good and f_2-good $G(2n, p, q)$ such that $r = \frac{q}{p}$ is a constant. Let H be the induced dynamical system. Then, for any sink point $\mathbf{a}^*$ and any sufficiently small $\epsilon = \omega(\sqrt{1/np})$,*

$$\mathbf{Pr}\left[\boldsymbol{\alpha}' \notin B(\mathbf{a}^*, \epsilon) \,|\, \boldsymbol{\alpha} \in B(\mathbf{a}^*, \epsilon)\right] \le \exp(-\Omega(\epsilon^2 n))$$

holds. In particular, let

$$\tau := \inf\left\{t \in \mathbb{N} : \boldsymbol{\alpha}^{(t)} \notin B(\mathbf{a}^*, \epsilon)\right\}$$

be a stopping time. Then, $\tau \ge \exp(\Omega(\epsilon^2 n))$ holds w.h.p. conditioned on $\boldsymbol{\alpha}^{(0)} \in B(\mathbf{a}^*, \epsilon)$ for any ϵ satisfying $\epsilon = \omega(\max\{1/\sqrt{np}, \sqrt{\log n/n}\})$.

7.2 Fast consensus

Suppose that the Jacobian matrix at the consensus point (i.e. $\alpha \in \{(0,0),(1,1)\}$ is the all-zero matrix. Then, we claim that the polynomial voting process reaches consensus within a small number of iterations if the initial set $A^{(0)}$ has small size.

▶ **Proposition 20.** *Consider an (f_1, f_2)-polynomial voting process on an f_1-good and f_2-good $G(2n, p, q)$ such that $\frac{p}{q}$ is a constant. Suppose that the Jacobian matrix at the point $\boldsymbol{\alpha} = (0,0)$ is the all-zero matrix. Then, there exists a constant $C_1, C_2, \delta > 0$ such that*

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C_1 \left(\log\log n + \frac{\log n}{\log np}\right)\right] \ge 1 - n^{-C_2}$$

for all $A \subseteq V$ satisfying $|A| \le \delta n$.

To show Proposition 20, we prove the following result which might be an independent interest:

▶ **Proposition 21.** *Consider a polynomial voting process on a graph G of n vertices. Suppose that there exist absolute constants $C, \delta > 0$ and a function $\epsilon = \epsilon(n) = o(1)$ such that*

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

holds for all $A \subseteq V$ satisfying $|A| \le \delta n$. Then, positive constants δ', C', C'' exist such that

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C'\left(\log\log n + \frac{\log n}{\log \epsilon^{-1}}\right)\right] \ge 1 - n^{-C''}$$

holds for all $A \subseteq V$ satisfying $|A| \le \delta' n$.

It should be noted that in Proposition 21, we do not restrict the underlying graph G to be random graphs.

7.3 Escape from a fixed point

Consider an (f_1, f_2)-polynomial voting process on an f_1-good and f_2-good $G(2n, p, q)$ such that p and q are constants. Let $\mathbf{a}^* \in \mathbb{R}^2$ be a fixed point of the induced dynamical system H. Let J be the Jacobian matrix of H at $\mathbf{a}^*$ and λ_1, λ_2 be its eigenvalues. Let $\mathbf{u}_i$ be the eigenvector of J corresponding to λ_i. Suppose that $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent. Then, we can rewrite J as $J = U^{-1}\Lambda U$, where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$ and $U = (\mathbf{u}_1\, \mathbf{u}_2)^{-1}$. For a fixed point $\mathbf{a}^* \in \mathbb{R}^2$, let $\boldsymbol{\beta} = (\beta_1, \beta_2)$ be a vector-valued random variable defined as

$$\boldsymbol{\beta} = U(\boldsymbol{\alpha} - \mathbf{a}^*). \tag{8}$$

Roughly speaking, from the Taylor expansion of H at $\mathbf{a}^*$, we have

$$\mathbf{E}[\boldsymbol{\beta}'] \approx \Lambda\boldsymbol{\beta}$$

if $\|\boldsymbol{\beta}\|_\infty$ is sufficiently small. Thus, $|\beta_i'| \approx |\lambda_i||\beta_i|$.

Recall that $B(\mathbf{x}, R)$ is the open ball of radius R centered at $\mathbf{x}$. If $|\lambda_i| > 1$ for some $i \in [2]$, one may expect that $\boldsymbol{\alpha}^{(\tau)} \notin B(\mathbf{a}^*, \epsilon_0)$ holds for any $A^{(0)} \subseteq V$ and for some constant $\epsilon_0 > 0$. We aim to prove this under some assumptions.

▶ **Asspmption 22** (Basic assumptions). *We consider an (f_1, f_2)-polynomial voting process on an f_1-good and f_2-good $G(2n, p, q)$ for constants $p \ge q \ge 0$. Let $\mathbf{a}^*$ be a fixed point and J be the corresponding Jacobian matrix satisfying*

(A1) *The eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent.*

(A2) *A positive constant ϵ_0 exists such that $\mathbf{Var}[\alpha_i' \mid A] \ge \Omega(n^{-1})$ for all $i \in \{1, 2\}$ and all $A \subseteq V$ of $\boldsymbol{\alpha} \in B(\mathbf{a}^*, \epsilon_0)$.*

(A3) *The matrix J contains an eigenvalue λ satisfying $|\lambda| > 1$.*

Under Assumption 22, we can define the random variable $\boldsymbol{\beta}$ of (8). Further, we put the following.

▶ **Asspmption 23.** *In addition to Assumption 22, we assume that there exists a positive constant ϵ^* satisfying the followings:*

(A4) *There exist two positive constants* ϵ_1, C *such that*

$$|\mathbf{E}[\beta_i' \mid A]| \geq (1+\epsilon_1)|\beta_i| - \frac{C}{\sqrt{n}}$$

holds for any $A \subseteq V$ *of* $\|\boldsymbol{\beta}\| \leq \epsilon^*$ *and any* $i \in [2]$ *of* $|\lambda_i| > 1$.

(A5) *For any* $i \in [2]$ *of* $|\lambda_i| \leq 1$,

$$\mathbf{Pr}[|\beta_i'| \leq \epsilon^* \mid |\beta_i| \leq \epsilon^*] \geq 1 - n^{-\Omega(1)}.$$

Sometimes, it might be not easy to check the conditions of Assumption 23. In this paper, we provide the following alternative condition which is easy to check:

▶ **Asspmption 24.** *In addition to Assumption 22, we assume the following:*

(A6) *The eigenvalues* λ_1, λ_2 *of* J *satisfies* $|\lambda_i| \neq 1$ *for all* $i \in [2]$.

Based on the assumptions, we prove the following result:

▶ **Proposition 25** (Escape from source and sink areas)**.** *Let* $\mathbf{a}^*$ *be a fixed point satisfying either Assumption 23 or 24. Then, there exist* $\tau = O(\log n)$ *and a constant* $\epsilon' > 0$ *such that the followings hold w.h.p.:*

(i) $\|\beta^{(\tau)}\|_\infty > \epsilon'$, *and*

(ii) $|\beta_j^{(\tau)}| \leq \epsilon'$ *for any* $j \in [2]$ *of* $|\lambda_j| \leq 1$.

References

1. E. Abbe. Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.
2. E. Abbe and C. Sandon. Recovering communities in the general stochastic block model without knowing the parameters. *In Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS)*, 1:676–684, 2015.
3. M. A. Abdullah and M. Draief. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Applied Mathematics*, 1(10):1–10, 2015.
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(6014):183–185, 2011.
5. P. Barbillon, S. Donnet, E. Lazega, and A. Bar-Hen. Stochastic block models for multiplex networks: an application to a multilevel network of researchers. *Journal of the Royal Statistical Society Series A*, 180(1):295–314, 2017.
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. *In Proceedings of the 26th Annual European Symposium on Algorithms (ESA)*, 7:1–13, 2018.
7. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, R. Silvestri, and L. Trevisan. Simple dynamics for plurality consensus. *Distributed Computing*, 30(4):293–306, 2017.
8. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Stabilizing consensus with many opinions. *In Proceedings of the 27th annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 620–635, 2016.
9. L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Find your place: Simple distributed algorithms for community detection. *In Proceedings of the 28th annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 940–959, 2017.
10. P. Berenbrink, A. Clementi, R. Elsässer, P. Kling, F. Mallmann-Trenn, and E. Natale. Ignore or comply? On breaking symmetry in consensus. *In Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, pages 335–344, 2017.

11 J. Chen and B. Yuan. Detecting functional modules in the yeast protein-protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.
12 C. Cooper, R. Elsässer, H. Ono, and T. Radzik. Coalescing random walks and voting on connected graphs. *SIAM Journal on Discrete Mathematics*, 27(4):1748–1758, 2013.
13 C. Cooper, R. Elsässer, and T. Radzik. The power of two choices in distributed voting. *In Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, 2:435–446, 2014.
14 C. Cooper, R. Elsässer, T. Radzik, N. Rivera, and T. Shiraga. Fast consensus for voting on general expander graphs. *In Proceedings of the 29th International Symposium on Distributed Computing (DISC)*, pages 248–262, 2015.
15 C. Cooper, T. Radzik, N. Rivera, and T. Shiraga. Fast plurality consensus in regular expanders. *In Proceedings of the 31st International Symposium on Distributed Computing (DISC)*, 91(13):1–16, 2017.
16 C. Cooper and N. Rivera. The linear voting model. *In Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, 55(144):1–12, 2016.
17 E. Cruciani, E. Natale, A. Nusser, and G. Scornavacca. Phase transition of the 2-choices dynamics on core-periphery networks. *In Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 777–785, 2018.
18 E. Cruciani, E. Natale, and G. Scornavacca. Distributed community detection via metastability of the 2-choices dynamics. *In Proceedings of the 33rd AAAI conference on artificial intelligence (AAAI)*, pages 6046–6053, 2019.
19 B. Doerr, L. A. Goldberg, L. Minder, T. Sauerwald, and C. Scheideler. Stabilizing consensus with the power of two choices. *In Proceedings of the 23rd annual ACM symposium on Parallelism in algorithms and architectures (SPAA)*, pages 149–158, 2011.
20 M. Fischer, N. Lynch, and M. Merritt. Easy impossibility proofs for distributed consensus problems. *Distributed Computing*, 1(1):26–39, 1986.
21 A. Frieze and M. Karoński. *Introduction to random graphs.* Campridge University Press, 2016.
22 M. Ghaffari and J. Lengler. Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. *In Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, pages 305–313, 2018.
23 S. Gilbert and D. Kowalski. Distributed agreement with optimal communication complexity. *In Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 965–977, 2010.
24 A. Goldenberg, A. X. Zheng, S. E. Fienberg, and E. M. Airoldi. A survey of statistical network models. *Foundations and Trends in Machine Learning*, 2(2):129–233, 2010.
25 Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001.
26 M. Hirsch and H. L. Smith. Monotone dynamical systems. *In Handbook of Differential Equations: Ordinary Differential Equations*, 2(4):239–357, 2005.
27 V. Kanade, F. Mallmann-Trenn, and T. Sauerwald. On coalescence time in graphs: When is coalescing as fast as meeting? *In Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 956–965, 2019.
28 N. Kang and R. Rivera. Best-of-Three voting on dense graphs. *In Proceedings of the 31st ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 115–121, 2019.
29 J. H. Kim and V. H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20:417–434, 2000.
30 T. M. Liggett. *Interacting particle systems.* Springer-Verlag, 1985.
31 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(5428):751–753, 1999.
32 E. Mossel, J. Neeman, and O. Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, 28(3):408–429, 2014.

33 T. Nakata, H. Imahayashi, and M. Yamashita. Probabilistic local majority voting for the agreement problem on finite graph. *In Proceedings of the 5th Annual International Computing and Combinatorics Conference (COCOON)*, pages 330–338, 1999.

34 R. I. Oliveira and Y. Peres. Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. *In Proceedings of the 16th Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 119–126, 2019.

35 G. Schoenebeck and F. Yu. Consensus of interacting particle systems on Erdős-Rényi graphs. *In Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1945–1964, 2018.

36 N. Shimizu and T. Shiraga. Phase Transitions of Best-of-Two and Best-of-Three on Stochastic Block Models. *arXiv*, 2019. `arXiv:1907.12212`.