The Annals of Applied Probability
2013, Vol. 23, No. 2, 637–664
DOI: 10.1214/12-AAP849

SHARP BENEFIT-TO-COST RULES FOR THE EVOLUTION OF COOPERATION ON REGULAR GRAPHS

BY YU-TING CHEN

University of British Columbia

We study two of the simple rules on finite graphs under the death–birth updating and the imitation updating discovered by Ohtsuki, Hauert, Lieberman and Nowak [*Nature* **441** (2006) 502–505]. Each rule specifies a payoff-ratio cutoff point for the magnitude of fixation probabilities of the underlying evolutionary game between cooperators and defectors. We view the Markov chains associated with the two updating mechanisms as voter model perturbations. Then we present a first-order approximation for fixation probabilities of general voter model perturbations on finite graphs subject to small perturbation in terms of the voter model fixation probabilities. In the context of regular graphs, we obtain algebraically explicit first-order approximations for the fixation probabilities of cooperators distributed as certain uniform distributions. These approximations lead to a rigorous proof that both of the rules of Ohtsuki et al. are valid and are sharp.

1. Introduction. The main objective of this paper is to investigate the simple rules for the evolution of cooperation by clever, but nonrigorous, arguments of pair-approximation on certain large graphs in Ohtsuki, Hauert, Lieberman and Nowak [18]. For convenience, we name these rules and their relatives comprehensively as *benefit-to-cost rules* (abbreviated as *b/c rules*) for reasons which will become clear later on. The work [18] takes spatial structure into consideration and gives an explanation with analytical criteria for the ubiquity of cooperative entities observed in biological systems and human societies. (See also the references in [18] for other models on structured populations.) In particular, this provides a way to overcome one of the major difficulties in theoretical biology since Darwin. (See Hamilton [8], Axelrod and Hamilton [3], Chapter 13 in Maynard Smith [14] and many others.)

We start by describing the evolutionary games defined in [18] and set some definitions. Consider a finite, connected, and simple (i.e., undirected and without loops or parallel edges) graph $G = (\mathsf{V}, \mathsf{E})$ on N vertices. (See, e.g., [4] for the standard terminology of graph theory.) Imagine the graph as a social network where a population of N individuals occupy the vertices of G and the edges denote the links

Received July 2011; revised January 2012.

MSC2010 subject classifications. Primary 60K35, 91A22; secondary 60J10, 60J28.

Key words and phrases. Evolutionary game theory, evolution of cooperation, interacting particle systems, voter model perturbations, voter model, coalescing random walks, perturbations of Markov chains.

between the individuals. The population consists of cooperators and defectors labeled by 1's and 0's, respectively. Their fitness is described through payoffs from encounters as follows. Consider a 2×2 payoff matrix

$$\mathbf{\Pi} = \begin{pmatrix} \Pi_{11} & \Pi_{10} \\ \Pi_{01} & \Pi_{00} \end{pmatrix} = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \tag{1.1}$$

Here, while positive constants are natural candidates for both benefit b and cost c, we allow arbitrary reals for their possible values unless otherwise mentioned. Each entry Π_{ij} of $\mathbf{\Pi}$ denotes the payoff that an i-player receives from a j-player. Hence, the payoff of a cooperator is $bn - ck$ if n of its k neighbors are cooperators, and the payoff of a defector is bm if m of its neighbors are cooperators. Then the fitness of an individual located at x is given by a convex combination of the *baseline fitness* with weight $1 - w$ and its payoff with weight w. The baseline fitness is normalized to 1 for convenience. The parameter w is interpreted as the *intensity of selection*. Therefore, *weak selection* means that payoff is a small proportion of fitness compared to the baseline fitness.

In contrast to game theory where strategies are decided by rational players, evolutionary game theory considers the random evolution of interacting players in which the "fitter" strategies have better chances to replicate. We study two of the updating mechanisms under weak selection in [18] for the evolution of cooperation throughout our work. Under the death–birth updating, we kill a random individual and then let its neighbors compete for the vacant vertex with success probability proportional to the fitness of its neighbors. Under the imitation updating, a random individual updates its strategy, but now it will either adhere to its original strategy or imitate one of the neighbors' strategies with success probability proportional to fitness. In this way, each updating mechanism defines a Markov chain on configurations consisting of 1's and 0's, or more specifically, a spin system in the sense of Liggett [11] where each vertex can adopt only two possible opinions, 1 and 0. Despite the simplicity of the transition rates, the readers may observe that the spin systems pose certain difficulty in terms of the classical approaches in interacting particle systems. For example, as a result of the asymmetry of payoffs, there is no symmetry between 1's and 0's in the two spin systems. In addition, it is not hard to see that in general the two spin systems are not attractive.

We are now ready to describe the b/c rules for the two evolutionary games which are surprisingly simple criterions for criticality in the asymptotic. The degree of a graph is defined in [18] to be the average number of neighbors per vertex. Put a single cooperative mutant on the vertices with a random location. Then the clever, but nonrigorous, calculations in the supplementary notes of [18], supported by several numerical simulations, lead to the following *b/c rule* for the death–birth updating under weak selection on certain large graphs of degree k; selection favors cooperation whenever $b/c > k$ and selection opposes cooperation whenever $b/c < k$. Here, selection *favors* (resp., *opposes*) cooperation if the probability that a single cooperative mutant converts the defecting population completely into a cooperative population is strictly higher (resp., lower) than the fixation probability $1/N$ of a neutral mutant. [See (2.8) for the latter probability. Equation (2.8) also shows, in particular, that if the graph is regular, then the fixation probability of a neutral mutant at an arbitrary location without further randomization is precisely $1/N$.] A similar b/c rule under the imitation mechanism is discussed in the supplementary notes of [18], with the modification that the cutoff point k should be replaced by $k+2$. We remark that the work [18] also considers the birth–death updating (in contrast to the death–birth updating) and its associated b/c rule. See [21] for a further study of these b/c rules. For more b/c rules see [8, 16] and [20], to name but a few. The monograph [17] gives an authoritative and excellent introduction to evolutionary dynamics.

Lying at the heart of the work [18] to obtain selective advantage of cooperators is the introduction of structured populations. This is manifested by the role of a fixed degree as population size becomes large. Consider instead a naive model where only fractions of players in a large population are concerned and the same payoff matrix (1.1) is in effect for evolutionary fitness. The fractions z_C and z_D of cooperators and defectors are modeled through replicator equations as

$$\begin{aligned} \dot{z}_C &= z_C(\rho_C - \bar{\rho}), \\ \dot{z}_D &= z_D(\rho_D - \bar{\rho}). \end{aligned} \tag{1.2}$$

Here, by the equality $z_C + z_D = 1$, the payoffs for cooperators and defectors are $\rho_C = bz_C - c$ and $\rho_D = bz_C$, and $\bar{\rho}$ is the average payoff given by $z_C\rho_C + z_D\rho_D$. By (1.2), the fraction of cooperators satisfies the following logistic differential equation:

$$\dot{z}_C = -cz_C(1 - z_C).$$

Hence, any proper fraction of cooperators must vanish eventually whenever cost c is positive. See, for example, Chapter 7 in [10], Chapter 4 in [17] or Section 3 in [9] for this model and more details. As discussed in more detail later on, a similar result holds in any unstructured population of finite size under the death–birth updating. Informally, a spatial structure, on one hand, promotes the formation of cliques of cooperators which collectively have a selective advantage and, on the other hand, reduces the exploitation of cooperators by defectors.

In [19], Ohtsuki and Nowak gave a rigorous proof of the b/c rules on large cycles under weak selection, in particular for the two updating mechanisms. The results in [19] exploit the fact that on cycles the fixation probabilities under each updating mechanism satisfy a system of birth–death-process type difference equations, and exact fixation probabilities can be derived accordingly. It is easy to get the exact solvability of fixation probabilities by the same approach on complete graphs, although on each *fixed* one of these graphs cooperators are always opposed under weak selection for the death–birth updating. (See [18] and Remark 1.1(3).

Note that the degree of a complete graph has the same order as the number of vertices.) It seems, however, harder to obtain fixation probabilities by extending this approach beyond cycles and complete graphs.

In this work, we will view each of the two spin systems as a *voter model perturbation* on a (finite, connected and simple) graph *of arbitrary size*. Voter model perturbations are studied in Cox and Perkins [7] and further developed in generality in Cox, Durrett and Perkins [6] on transient integer lattices $\mathbb{Z}^d$ for $d \geq 3$. On the infinite lattices considered in [6] (often sharp) conditions, based on a related reaction diffusion equation, were found to ensure the coexistence of 1's and 0's, or to ensure that one type drives the other out. In particular, a rigorous proof of the b/c rule under the death–birth updating on these infinite graphs is obtained in [6]. Here, in the context of finite graphs, the voter model perturbation associated with each of the spin systems fixates in one of the the two absorbing states of all 1's or all 0's, and we give a first-order approximation to the fixation probabilities by expansion. In spite of the apparent differences in the settings, there are interesting links between our fixation probability expansions and the reaction diffusion equation criteria in [6].

Let us now introduce voter model perturbations as a family of spin systems. Denote by x a vertex and by η a configuration. Let $c(x,\eta)$ be the flipping rate of the (nearest-neighbor) voter model at x given η. Hence, $c(x,\eta)$ is equal to the probability of drawing a neighbor of x with the opposite opinion to that of x. Then interpreted narrowly in the context considered in this paper, the rates of a voter model perturbation are given by

$$c^w(x,\eta) = c(x,\eta) + wh_{1-\eta(x)}(x,\eta) + w^2 g_w(x,\eta) \geq 0 \tag{1.3}$$

for a small perturbation rate $w > 0$. Here, h_1, h_0 and g_w for all w small are uniformly bounded. We refer the readers to Chapter V in [11] here and in the following for the classical results of voter models and to Section 1 in [6] for a general definition of voter model perturbations on transient integer lattices.

We discuss in more detail the aforementioned result in [6] which is closely related to the present work. A key result in [6] states that on the integer lattices $\mathbb{Z}^d$ for $d \geq 3$, the invariant distributions of a voter model perturbation subject to small perturbation can be determined by the reaction function through a reaction-diffusion PDE. (See Section 1.2 in [6].) Here, the reaction function takes the form

$$u \longmapsto \lim_{s\to\infty} \int D(0,\eta)\mathbb{P}_{\mu_u}(\xi_s \in d\eta), \qquad u \in [0,1], \tag{1.4}$$

where $((\xi_s), \mathbb{P}_{\mu_u})$ denotes the voter model starting at the Bernoulli product measure μ_u with density $\mu_u(\eta(x) = 1) = u$, and the difference kernel D is defined by

$$D(x,\eta) = \widehat{\eta}(x)h_1(x,\eta) - \eta(x)h_0(x,\eta) \tag{1.5}$$

with $\widehat{\eta}(x) \equiv 1 - \eta(x)$. By the duality between voter models and coalescing random walks, the reaction function defined by (1.4) can be expressed explicitly as a polynomial with coefficients consisting of coalescing probabilities of random walks.

The justification in [6] of the b/c rule for the death–birth updating on transient integer lattices is under a slightly different definition, for the sake of adaptation to the context of infinite graphs. Precisely, the result in [6] states that whenever $b/c > k$ (resp., $b/c < k$) and there is weak selection, given infinitely many cooperators (resp., defectors) at the beginning, any given finite set of vertices will become occupied by cooperators (resp., defectors) from some time onward almost surely. Here, k refers to the degree of each vertex in the underlying lattice and is equal to $2d$ on $\mathbb{Z}^d$. The b/c rule under the imitation updating on the same integer lattices is verified in [5], under the same definition in [6] except that, as pointed out in [18], the cutoff point k needs to be replaced by $k+2$.

We now discuss our result for voter model perturbations on (finite, connected and simple) graphs of arbitrary size. We first work with discrete-time Markov chains of the voter model perturbations. We assume, in addition, the chain starting at any arbitrary state is eventually trapped at either of the two absorbing states, the all-1 configuration $\mathbf{1}$ and the all-0 configuration $\mathbf{0}$. This is a property enjoyed by both updating mechanisms under weak selection. We proceed analytically and decompose the transition kernel P^w of a voter model perturbation with perturbation rate w as the sum of the transition kernel P of the voter model and a signed kernel K^w. We apply an elementary expansion to finite-step transitions of the P^w-chain in the spirit of Mayer's cluster expansion [13] in statistical mechanics. Then we show every linear combination of the fixation probabilities of $\mathbf{1}$ and $\mathbf{0}$ subject to small perturbation admits an infinite series expansion closely related to the voter model. A slight refinement of this expansion leads to our main result for the general voter model perturbations on which our study of the b/c rules relies.

Precisely, our main result for the voter model perturbations on finite graphs (Theorem 3.8) can be stated as follows. Regard the voter model perturbation with perturbation rate w as a continuous-time chain (ξ_s) with rates given by (1.3). Recall by our assumption that the chain starting at any state is eventually trapped at either of the absorbing states $\mathbf{1}$ and $\mathbf{0}$. Define $\tau_{\mathbf{1}}$ for the time to the absorbing state $\mathbf{1}$ and

$$\overline{H}(\xi) = \sum_{x \in \mathsf{V}} H(x, \xi)\pi(x) \tag{1.6}$$

for any $H(x, \xi)$. Here, $\pi(x)$ is the invariant distribution of the (nearest-neighbor) random walk on G given by

$$\pi(x) = \frac{d(x)}{2 \cdot \#\mathsf{E}} \tag{1.7}$$

with $d(x)$ being the degree of x, that is, the number of neighbors of x. (See, e.g., [1] and [12] for random walks on graphs.) Then as $w \longrightarrow 0+$, we have the following approximation:

$$\mathbb{P}^w(\tau_{\mathbf{1}} < \infty) = \mathbb{P}(\tau_{\mathbf{1}} < \infty) + w\int_0^\infty \mathbb{E}[\overline{D}(\xi_s)]\,ds + O(w^2). \tag{1.8}$$

Here, $\mathbb{P}^w$ and $\mathbb{P}$ (with expectation $\mathbb{E}$) denote the laws of the voter model perturbation with perturbation rate w and the voter model, respectively, both subject to the same, but arbitrary, initial distribution, and D is the difference kernel defined by (1.5). Moreover, the integral term on the right-hand side of (1.8) makes sense because $\overline{D}(\xi_s) \in L_1(d\mathbb{P} \otimes ds)$.

We apply the first-order approximation (1.8) to the two evolutionary games only on regular graphs. (A graph is a k-regular if all vertices have the same degree k and a graph is regular if it is k-regular graph for some k.) Under weak selection, the approximation (1.8) implies that we can approximate $\mathbb{P}^w(\tau_{\mathbf{1}} < \infty)$ by $\mathbb{P}(\tau_{\mathbf{1}} < \infty)$ and the 0-potential of $\overline{D}$

$$\int_0^\infty \mathbb{E}[\overline{D}(\xi_s)]\,ds, \tag{1.9}$$

all subject to the same initial distribution. Moreover, the comparison of $\mathbb{P}^w(\tau_{\mathbf{1}} < \infty)$ for small w with $\mathbb{P}(\tau_{\mathbf{1}} < \infty)$ is possible whenever the 0-potential is nonzero, with the order determined in the obvious way. For notions to be introduced later on, we take as initial distribution the uniform distribution $\mathbf{u}_n$ on the set of configurations with exactly n many 1's, where $1 \leq n \leq N-1$. Each 0-potential in (1.9) starting at $\mathbf{u}_n$ can be derived from the same 0-potentials starting at Bernoulli product measures μ_u with density $u \in [0, 1]$. Furthermore, each 0-potential with starting measure μ_u can be expressed in terms of some (expected) coalescing times of coalescing random walks. This is in contrast to the involvement of coalescing probabilities for the reaction functions in the context in [6]. By resorting to a simple identity in [1] between meeting times and hitting times of random walks, we obtain explicit forms of the coalescing times involved. Hence, by (1.8), we obtain the fixation probabilities explicit up to the first-order term, and the precise result is stated as follows.

THEOREM 1. *Let G be any (finite, connected and simple) graph on N vertices. Suppose in addition that G is k-regular, that is, every vertex of G has precisely degree k. Fix $1 \leq n \leq N-1$ and let $\mathbb{P}^w_{\mathbf{u}_n}$ denote the law of the particular evolutionary game with small intensity of selection w and initial distribution $\mathbf{u}_n$.*

(1) *Under the death–birth updating, we have*

$$\mathbb{P}^w_{\mathbf{u}_n}(\tau_{\mathbf{1}} < \infty) = \frac{n}{N} + w\left[\frac{kn(N-n)}{2N(N-1)}\right]\left[\left(\frac{b}{k} - c\right)(N-2) + b\left(\frac{2}{k} - 2\right)\right] + O(w^2)$$

as $w \longrightarrow 0+$.

(2) *Under the imitation updating*, *we have*

$$\mathbb{P}^w_{\mathbf{u}_n}(\tau_{\mathbf{1}} < \infty) = \frac{n}{N} + w\left[\frac{k(k+2)n(N-n)}{2(k+1)N(N-1)}\right] \\ \times \left[\left(\frac{b}{(k+2)} - c\right)(N-1) - \frac{(2k+1)b - ck}{k+2}\right] + O(w^2)$$

as $w \longrightarrow 0+$.

Here, *in either* (1) *or* (2), *we use Landau's notation* $O(w^2)$ *for a function* $\theta(w)$ *such that* $|\theta(w)| \leq Cw^2$ *for all small* w, *for some constant* C *depending only on the graph* G *and the particular updating mechanism.*

Before interpreting the result of Theorem 1 in terms of evolutionary games, we first introduce the following definition which is stronger than that in [18]. We say selection *strongly* favors (resp., opposes) cooperation if for *every* nontrivial n, that is, $1 \leq n \leq N-1$, the following holds: The probability that n cooperative mutants with a joint location distributed as $\mathbf{u}_n$ converts the defecting population completely into a cooperative population is strictly higher (resp., lower) than n/N. [Here, n/N is the fixation of probability of n neutral mutants again by (2.8).] Under this definition, Theorem 1 yields simple algebraic criteria for both evolutionary games stated as follows.

COROLLARY 1. *Suppose again that the underlying social network is a* k-*regular graph on* N *vertices*.

(1) *For the death–birth updating*, *if*

$$\left(\frac{b}{k} - c\right)(N-2) + b\left(\frac{2}{k} - 2\right) > 0 \qquad (\textit{resp.},\ < 0),$$

then selection strongly favors (*resp*., *opposes*) *cooperation under weak selection*.

(2) *For the imitation updating*, *if*

$$\left(\frac{b}{(k+2)} - c\right)(N-1) - \frac{(2k+1)b - ck}{k+2} > 0 \qquad (\textit{resp.},\ < 0),$$

then selection strongly favors (*resp*., *opposes*) *cooperation under weak selection*.

Applied to cycles, the algebraic criteria in Corollary 1 under the aforementioned stronger definition coincide with the algebraic criteria in [19] for the respective updating mechanism. See also equation (3) in [21] for the death–birth updating.

As an immediate consequence of Corollary 1, we have the following result.

COROLLARY 2. *Fix a degree* k.

(1) *Consider the death–birth updating. For every fixed pair* (b, c) *satisfying* $b/k > c$ (*resp.*, $b/k < c$), *there exists a positive integer* N_0 *such that on any* k-*regular graph* $G = (\mathsf{V}, \mathsf{E})$ *with* $\#\mathsf{V} \geq N_0$, *selection strongly favors* (*resp.*, *opposes*) *cooperation under weak selection.*

(2) *Consider the imitation updating. For every fixed pair* (b, c) *satisfying* $b/(k+2) > c$ [*resp.*, $b/(k+2) < c$], *there exists a positive integer* N_0 *such that on any* k-*regular graph* $G = (\mathsf{V}, \mathsf{E})$ *with* $\#\mathsf{V} \geq N_0$, *selection strongly favors* (*resp.*, *opposes*) *cooperation under weak selection.*

In this way, we rigorously prove the validity of the b/c rule in [18] under each updating mechanism. It is in fact a *universal* rule valid for any nontrivial number of cooperative mutants, and holds *uniformly* in the number of vertices, for large regular graphs with a fixed degree under weak selection.

REMARK 1.1. (1) Although we only consider payoff matrices of the special form (1.1) in our work, interests in evolutionary game theory do cover general 2×2 payoff matrices with arbitrary entries. (See, e.g., [18] and [19].) In this case, a general 2×2 matrix $\mathbf{\Pi}^* = (\Pi^*_{ij})_{i,j=1,0}$ is taken to define payoffs of players with an obvious adaptation of payoffs under $\mathbf{\Pi}$. For example, the payoff of a cooperator is $(\Pi^*_{11} - \Pi^*_{10})n + k\Pi^*_{10}$ if n of its k neighbors are cooperators. In particular, if $\mathbf{\Pi}^*$ satisfies the *equal-gains-from-switching* condition (Nowak and Sigmund [15])

$$\Pi^*_{11} - \Pi^*_{10} = \Pi^*_{01} - \Pi^*_{00}, \tag{1.10}$$

then the results in Theorem 1, Corollaries 1 and 2 still hold for $\mathbf{\Pi}^*$ by taking $\mathbf{\Pi}$ in their statements to be the “adjusted” payoff matrix

$$\mathbf{\Pi}^a := \begin{pmatrix} \Pi^*_{11} - \Pi^*_{00} & \Pi^*_{10} - \Pi^*_{00} \\ \Pi^*_{01} - \Pi^*_{00} & 0 \end{pmatrix}, \tag{1.11}$$

which is of the form in (1.1). See Remark 5.1 for this reduction.

(2) We stress that when $n = 1$ or $N - 1$ and the graphs are vertex-transitive [4] (and hence, regular) such as tori, the exact locations of mutants become irrelevant. It follows that the randomization by $\mathbf{u}_n$ is redundant in these cases.

(3) Let G be the complete graph on N vertices so that the spatial structure is irrelevant. Consider the death–birth updating and the “natural case” where benefit b and cost c are both positive. With the degree k set equal to $N - 1$, Theorem 1(1) gives for any $1 \leq n \leq N - 1$ the approximation

$$\mathbb{P}^w_{\mathbf{u}_n}(\tau_{\mathbf{1}} < \infty) = \frac{n}{N} + w\frac{n(N-n)}{2N}\left[-c(N-2) - \left(2 - \frac{N}{N-1}\right)b\right] + O(w^2)$$

as $w \longrightarrow 0+$. Hence, cooperators are always opposed under weak selection when $N \geq 3$.

The paper is organized as follows. In Section 2, we set up the standing assumptions of voter model perturbations considered throughout this paper and discuss their basic properties. The Markov chains associated with the two updating mechanisms in particular satisfy these standing assumptions, as stated in Propositions 2.1 and 2.2. In Section 3, we continue to work on the general voter model perturbations. We develop an expansion argument to obtain an infinite series expansion of fixation probabilities under small perturbation rates (Proposition 3.2) and then refine its argument to get the first-order approximation (1.8) (Theorem 3.8). In Section 4, we return to our study of the two evolutionary games and give the proof of Theorem 1. The vehicle for each explicit result is a simple identity between meeting times and hitting times of random walks. Finally, the proofs of Propositions 2.1 and 2.2 (that both updating mechanisms define voter model perturbations satisfying our standing assumptions) are deferred to Section 5.

2. Voter model perturbations. Recall that we consider only finite, connected and simple graphs in this paper. Fix such a graph $G = (\mathsf{V}, \mathsf{E})$ on $N = \#\mathsf{V}$ vertices. Write $x \sim y$ if x and y are neighbors to each other, that is, if there is an edge of G between x and y. We put $d(x)$ for the number of neighbors of x.

Introduce an auxiliary number $\lambda \in (0, 1]$. Take a nearest-neighbor discrete-time voter model with transition probabilities

$$\begin{aligned} P(\eta, \eta^x) &= \frac{\lambda}{N} c(x, \eta), \qquad x \in \mathsf{V}, \\ P(\eta, \eta) &= 1 - \frac{\lambda}{N} \sum_x c(x, \eta). \end{aligned} \tag{2.1}$$

Here, η^x is the configuration obtained from η by giving up the opinion of η at x for

$$\widehat{\eta}(x) := 1 - \eta(x)$$

and holding the opinions at other vertices fixed and we set

$$c(x, \eta) = \frac{\#\{y \sim x; \eta(y) = \widehat{\eta}(x)\}}{d(x)}.$$

We now define the discrete-time voter model perturbations considered throughout this paper as follows. Suppose that we are given functions h_i and g_w and a constant $w_0 \in (0, 1)$ satisfying

$$\sup_{w \in [0, w_0], x, \eta} \big(|h_1(x, \eta)| + |h_0(x, \eta)| + |g_w(x, \eta)|\big) \le C_0 < \infty. \tag{A.1}$$

$$c^w(x, \eta) := c(x, \eta) + w h_{1-\eta(x)}(x, \eta) + w^2 g_w(x, \eta) \ge 0, \tag{A.2}$$

$$c^w(x, \mathbf{1}) = c^w(x, \mathbf{0}) \equiv 0 \qquad \text{for each } x \in \mathsf{V} \tag{A.3}$$

for each $w \in [0, w_0]$. Here, $\mathbf{1}$ and $\mathbf{0}$ denote the all-1 configuration and the all-0 configuration, respectively. In (A.2), we set up a basic perturbation of voter model rates up to the second order. In terms of the voter model perturbations defined below by $c^w(x, \eta)$, we will be able to control the higher order terms in an expansion of fixation probabilities with the uniform bound imposed in (A.1). The assumption (A.3) ensures that the voter model perturbations have the same absorbing states $\mathbf{1}$ and $\mathbf{0}$ as the previously defined voter model.

Under the assumptions (A.1)–(A.3), we define for each perturbation rate $w \in [0, w_0]$ a *voter model perturbation* with transition probabilities

$$P^w(\eta, \eta^x) = \frac{\lambda}{N} c^w(x, \eta), \qquad x \in \mathsf{V}, \tag{2.2}$$
$$P^w(\eta, \eta) = 1 - \sum_x \frac{\lambda}{N} c^w(x, \eta).$$

[Here we assume without loss of generality by (A.1) that each $P^w(\eta, \cdot)$ is truly a probability measure, in part explaining the need of the auxiliary number λ.] In particular $P^0 \equiv P$.

NOTATION. We shall write $\mathbb{P}^w_\nu$ for the law of the voter model perturbation with perturbation rate w and initial distribution ν and set $\mathbb{P}_\nu := \mathbb{P}^0_\nu$. In particular we put $\mathbb{P}^w_\eta := \mathbb{P}^w_{\delta_\eta}$ and $\mathbb{P}_\eta := \mathbb{P}_{\delta_\eta}$, where δ_η is the Dirac measure at η. The discrete-time and continuous-time coordinate processes on $\{1, 0\}^{\mathsf{V}}$ are denoted by $(\xi_n; n \geq 0)$ and $(\xi_s; s \geq 0)$, respectively. Here, and in what follows, we abuse notation to read "n" and other indices for the discrete time scale and "s" for the continuous time scale whenever there is no risk of confusion.

Our last assumption, which is obviously satisfied by the P-voter model thanks to the connectivity of G, is

$$\mathbb{P}^w_\eta(\xi_n \in \{\mathbf{1}, \mathbf{0}\} \text{ for some } n) > 0 \qquad \text{for every } \eta \in \{1, 0\}^{\mathsf{V}} \tag{A.4}$$

for each $w \in (0, w_0]$. Since $\mathbf{1}$ and $\mathbf{0}$ are absorbing by the condition (A.3), it follows from the Markov property that the condition (A.4) is equivalent to the condition that the limiting state exists and can only be either of the absorbing states $\mathbf{1}$ and $\mathbf{0}$ under $\mathbb{P}^w$ for any $w \in (0, w_0]$.

PROPOSITION 2.1 ([6]). *Suppose that the graph is k-regular. Then the Markov chain associated with the death–birth updating with small intensity of selection w is a voter model perturbation with perturbation rate w satisfying* (A.1)–(A.4) *with $\lambda = 1$ and*

$$h_1 = -(b+c)kf_0 f_1 + kbf_{00} + kf_0(bf_{11} - bf_{00}), \tag{2.3}$$
$$h_0 = -h_1.$$

Here,

$$\begin{aligned} f_i(x,\eta) &= \frac{1}{k}\#\{y;\, y\sim x, \eta(y)=i\}, \\ f_{ij}(x,\eta) &= \frac{1}{k^2}\#\{(y,z);\, x\sim y\sim z, \eta(y)=i, \eta(z)=j\}. \end{aligned} \tag{2.4}$$

PROPOSITION 2.2. *Suppose that the graph is k-regular. Then the Markov chain associated with the imitation updating with small intensity of selection w is a voter model perturbation with perturbation rate w satisfying* (A.1)–(A.4) *with* $\lambda = \frac{k}{k+1}$ *and*

$$\begin{aligned} h_1 &= k[(b-c)f_{11} - cf_{10}] - \frac{k^2}{k+1} f_1[(b-c)f_{11} - cf_{10} + bf_{01}] \\ &\quad - \frac{k}{k+1} bf_1^2, \\ h_0 &= kbf_{01} - \frac{k^2}{k+1} f_0[(b-c)f_{11} - cf_{10} + bf_{01}] \\ &\quad - \frac{k}{k+1} f_0[(b-c)f_1 - cf_0], \end{aligned} \tag{2.5}$$

where f_i and f_{ij} are as in (2.4).

The proofs of Propositions 2.1 and 2.2 are deferred to Section 5.

The assumptions (A.1)–(A.4) *are in force from now on.*

Let us consider some basic properties of the previously defined discrete-time chains. First, as has been observed, we know that

$$1 = \mathbb{P}^w(\tau_{\mathbf{1}} \wedge \tau_{\mathbf{0}} < \infty) = \mathbb{P}^w(\tau_{\mathbf{1}} < \infty) + \mathbb{P}^w(\tau_{\mathbf{0}} < \infty),$$

where we write τ_η for the first hitting time of η. Observe that $\mathbb{P}^w(\tau_{\mathbf{1}} < \infty)$ is independent of the auxiliary number $\lambda > 0$. Indeed, the holding time of each configuration $\eta \neq \mathbf{1}, \mathbf{0}$ is finite and the probability of transition from η to η^x at the end of the holding time is given by

$$\frac{c^w(x,\eta)}{\sum_{y\in \mathsf{V}} c^w(y,\eta)},$$

which is independent of $\lambda > 0$.

We can estimate the equilibrium probability $\mathbb{P}^w(\tau_{\mathbf{1}} < \infty)$ by a "harmonic sampler" of the voter model from finite time. Let $p_1(\xi)$ be the weighted average of 1's in the vertex set

$$p_1(\eta) = \sum_x \eta(x)\pi(x), \tag{2.6}$$

where $\pi(x)$ is the invariant distribution of the (nearest-neighbor) random walk on G and is given by (1.7). Since $p_1(\mathbf{1}) = 1 - p_1(\mathbf{0}) = 1$ and the chain is eventually trapped at $\mathbf{1}$ or $\mathbf{0}$, it follows from dominated convergence that

$$\lim_{n\to\infty} \mathbb{E}^w[p_1(\xi_n)] = \mathbb{P}^w(\tau_{\mathbf{1}} < \infty). \tag{2.7}$$

On the other hand, the function p_1 is harmonic for the voter model

$$\begin{aligned}\mathbb{E}_\eta[p_1(\xi_1)] = p_1(\eta) + &\frac{\lambda}{N \cdot 2\#\mathsf{E}} \\ &\times \Bigl(\sum_{\eta(x)=0} \#\{y \sim x; \eta(y) = 1\} - \sum_{\eta(x)=1} \#\{y \sim x; \eta(y) = 0\}\Bigr) \\ = p_1(\eta).&\end{aligned}$$

In particular, (2.7) applied to $w = 0$ entails

$$\mathbb{P}_\eta(\tau_{\mathbf{1}} < \infty) = p_1(\eta) = \frac{\sum_{\eta(x)=1} d(x)}{2 \cdot \#\mathsf{E}}, \tag{2.8}$$

where the last equality follows from the explicit form (1.7) of π.

REMARK 2.3. Since every harmonic function f for the voter model satisfies

$$f(\eta) \equiv \mathbb{E}_\eta[f(\xi_{\tau_{\mathbf{1}} \wedge \tau_{\mathbf{0}}})],$$

(2.8) implies that the vector space of harmonic functions is explicitly characterized as the span of the constant function 1 and p_1. Recall also the foregoing display gives a construction of any harmonic function with preassigned values at $\mathbf{1}$ and $\mathbf{0}$. (See, e.g., Chapter 2 in [1].)

3. Expansion. We continue to study the discrete-time voter model perturbations defined in Section 2. For each $w \in [0, w_0]$, consider the signed kernel

$$K^w = P^w - P,$$

which measures the magnitude of perturbations of transition probabilities. We also define a nonnegative kernel $|K^w|$ by $|K^w|(\eta, \widetilde{\eta}) = |K^w(\eta, \widetilde{\eta})|$.

LEMMA 3.1. *For any $w \in [0, w_0]$ and any $f : \{1, 0\}^{\mathsf{V}} \longrightarrow \mathbb{R}$, we have*

$$K^w f(\eta) = \frac{\lambda}{N} \sum_x [w h_{1-\eta(x)}(x, \eta) + w^2 g_w(x, \eta)][f(\eta^x) - f(\eta)] \tag{3.1}$$

and

$$\||K^w| f\|_\infty \le 4C_0 w \|f\|_\infty, \tag{3.2}$$

where C_0 is the constant in (A1).

PROOF. We notice that for any η and any x,

$$K^w(\eta,\eta^x)=\frac{\lambda}{N}\big[wh_{1-\eta(x)}(x,\eta)+w^2g_w(x,\eta)\big],$$
$$K^w(\eta,\eta)=-\sum_x\frac{\lambda}{N}\big[wh_{1-\eta(x)}(x,\eta)+w^2g_w(x,\eta)\big],$$

by the definitions of c^w and P^w. Our assertions (3.1) and (3.2) then follow at one stroke. $\square$

Using the signed kernel K^w, we can rewrite every T-step transition probability P^w_T of the voter model perturbation as

$$\begin{aligned}P^w_T(\eta_0,\eta_T)&=\sum_{\eta_1,\dots,\eta_{T-1}}P^w(\eta_0,\eta_1)P^w(\eta_1,\eta_2)\cdots P^w(\eta_{T-1},\eta_T)\\&=\sum_{\eta_1,\dots,\eta_{T-1}}(P+K^w)(\eta_0,\eta_1)\cdots(P+K^w)(\eta_{T-1},\eta_T)\\&=P_T(\eta_0,\eta_T)+\sum_{n=1}^{T}\sum_{\mathbf{j}\in\mathcal{I}_T(n)}\sum_{\eta_1,\dots,\eta_{T-1}}\Delta^{w,\mathbf{j}}_T(\eta_0,\dots,\eta_T).\end{aligned}\tag{3.3}$$

Here, $\mathcal{I}_T(n)$ is the set of strictly increasing n-tuples with entries in $\{1,\dots,T\}$, and for $\mathbf{j}=(j_1,\dots,j_n)\in\mathcal{I}_T(n)$

$$\Delta^{w,\mathbf{j}}_T(\eta_0,\dots,\eta_T)\tag{3.4}$$

is the signed measure of the path $(\eta_0,\eta_1,\dots,\eta_T)$ such that the transition from η_r to η_{r+1} is determined by $K^w(\eta_r,\eta_{r+1})$ if $r+1$ is one of the (integer-valued) indices in $\mathbf{j}$ and is determined by $P(\eta_r,\eta_{r+1})$ otherwise. For convenience, we set for each $\mathbf{j}\in\mathcal{I}_T(n)$

$$Q^{w,\mathbf{j}}_T(\eta_0,\eta_T)=\sum_{\eta_1,\dots,\eta_{T-1}}\Delta^{w,\mathbf{j}}_T(\eta_0,\dots,\eta_T)$$

as the T-step transition signed kernel, and we say $Q^{w,\mathbf{j}}_T$ has *n faults* (up to time T) and $\mathbf{j}$ is its *fault sequence*. Then by (3.3), we can write for any $f:\{1,0\}^{\mathsf{V}}\longrightarrow\mathbb{R}$,

$$\begin{aligned}\mathbb{E}^w_{\eta_0}[f(\xi_T)]&=\mathbb{E}_{\eta_0}[f(\xi_T)]+\sum_{n=1}^{T}\sum_{\mathbf{j}\in\mathcal{I}_T(n)}\sum_{\eta_1,\dots,\eta_T}\Delta^{w,\mathbf{j}}_T(\eta_0,\dots,\eta_T)f(\eta_T)\\&=\mathbb{E}_{\eta_0}[f(\xi_T)]+\sum_{n=1}^{T}\sum_{\mathbf{j}\in\mathcal{I}_T(n)}Q^{w,\mathbf{j}}_Tf(\eta_0).\end{aligned}\tag{3.5}$$

Write $\mathcal{I}(n)\equiv\mathcal{I}_\infty(n)$ for the set of strictly increasing n-tuples with entries in $\mathbb{N}$. We now state the key result which in particular offers an expansion of fixation probabilities.

PROPOSITION 3.2. *Recall the parameter $w_0 > 0$ in the definition of the voter model perturbations. There exists $w_1 \in (0, w_0]$ such that for any harmonic function f for the voter model*,

$$f(\mathbf{1})\mathbb{P}^w_\eta(\tau_{\mathbf{1}} < \infty) + f(\mathbf{0})\mathbb{P}^w_\eta(\tau_{\mathbf{0}} < \infty) = f(\eta) + \sum_{n=1}^{\infty} \sum_{\mathbf{j}\in\mathcal{I}(n)} Q^{w,\mathbf{j}}_{j_n} f(\eta), \tag{3.6}$$

where the series converges absolutely and uniformly in $w \in [0, w_1]$ and in $\eta \in \{1, 0\}^{\mathsf{V}}$.

REMARK 3.3. (i) There are alternative perspectives to state the conclusion of Proposition 3.2. Thanks to Remark 2.3 and the fact $Q^{w,\mathbf{j}}_T \mathbb{1} \equiv 0$, it is equivalent to the validity of the same expansion for p_1 defined in (2.6) (for any small w). By Remark 2.3 again, it is also equivalent to an analogous expansion of any linear combination of the two fixation probabilities under $\mathbb{P}^w$.

(ii) The series expansion (3.6) has the flavor of a Taylor series expansion in w, as hinted by Lemma 3.6.

The proof of Proposition 3.2 is obtained by passing T to infinity for both sides of (3.5). This immediately gives the left-hand side of (3.6) thanks to our assumption (A.4). There are, however, two technical issues when we handle the right-hand sides of (3.5). The first one is minor and is the dependence on T of the summands $Q^{w,\mathbf{j}}_T f(\eta_0)$. For this, the harmonicity of f implies that such dependence does not exist, as asserted in Lemma 3.4. As a result, the remaining problem is the absolute convergence of the series on the right-hand side of (3.6) for any small parameter $w > 0$. This is resolved by a series of estimates in Lemmas 3.5, 3.6 and finally Lemma 3.7.

LEMMA 3.4. *For any harmonic function f for the voter model, any $T \geq 1$, and any $\mathbf{j} \in \mathcal{I}_T(n)$,*

$$Q^{w,\mathbf{j}}_T f(\eta_0) \equiv Q^{w,\mathbf{j}}_{j_n} f(\eta_0),$$

where we identify $\mathbf{j} \in \mathcal{I}_{j_n}(n)$ in the natural way.

PROOF. This follows immediately from the martingale property of a harmonic function f for the voter model and the definition of the signed measures $\Delta^{w,\mathbf{j}}_T$ in (3.4). $\square$

LEMMA 3.5. *There exist $C_1 = C_1(G) \geq 1$ and $\delta = \delta(G) \in (0, 1)$ such that*

$$\sup_{\eta\neq\mathbf{1},\mathbf{0}} \mathbb{P}_\eta(\xi_n \neq \mathbf{1}, \mathbf{0}) \leq C_1\delta^n \qquad \text{for any } n \geq 1. \tag{3.7}$$

PROOF. Recall that the voter model starting at any arbitrary state is eventually trapped at either $\mathbf{1}$ or $\mathbf{0}$. By identifying $\mathbf{1}$ and $\mathbf{0}$, we deduce (3.7) from some standard results of nonnegative matrices, for suitable constants $C_1 = C_1(G) \geq 1$ and $\delta \in (0, 1)$. (See, e.g., [2], Lemma I.6.1 and Proposition I.6.3.) $\square$

LEMMA 3.6. *Let $C_1 = C_1(G)$ and $\delta = \delta(G)$ be the constants in Lemma 3.5, and set $C = C(G, C_0) = \max(4C_0, C_1)$. Then for any $\mathbf{j} \in \mathcal{I}(n)$, any $w \in [0, w_0]$ and any harmonic function f for the voter model*,

$$\|Q_{j_n}^{w,\mathbf{j}} f\|_\infty \leq \|f\|_\infty w^n C^{2n} \delta^{j_n - n}. \tag{3.8}$$

PROOF. Without loss of generality, we may assume $\|f\|_\infty = 1$. By definition,

$$Q_{j_n}^{w,\mathbf{j}} f(\eta_0) = \sum_{\eta_1,\ldots,\eta_{j_n}} \Delta_{j_n}^{w,\mathbf{j}}(\eta_0, \ldots, \eta_{j_n}) f(\eta_{j_n}). \tag{3.9}$$

If $\Delta_{j_n}^{w,\mathbf{j}}(\eta_0, \ldots, \eta_{j_n})$ is nonzero, then none of the $\eta_0, \ldots, \eta_{j_n-1}$ is $\mathbf{1}$ or $\mathbf{0}$. Indeed, if some of the η_i, $0 \leq i \leq j_n - 1$, is $\mathbf{1}$, then $\eta_{i+1}, \ldots, \eta_{j_n}$ can only be $\mathbf{1}$ by (A.3) and therefore $K^w f(\eta_{j_n-1}) = 0$. This is a contradiction. Similarly, we cannot have $\eta_i = \mathbf{0}$ for some $0 \leq i \leq j_n - 1$. Hence, the nonvanishing summands of the right-hand side of (3.9) range over $\Delta_{j_n}^{w,\mathbf{j}}(\eta_1, \ldots, \eta_{j_n}) f(\eta_{j_n})$ for which none of the $\eta_{j_1}, \ldots, \eta_{j_n-1}$ is $\mathbf{1}$ or $\mathbf{0}$. With η_0 fixed, write $\Delta_{U,\eta_0}^{w,\mathbf{j}'}$ for the signed measure $\Delta_U^{w,\mathbf{j}'}$ restricted to paths starting at η_0. Thus we get from (3.9) that

$$Q_{j_n}^{w,\mathbf{j}} f(\eta_0) = \Delta_{j_n,\eta_0}^{w,\mathbf{j}}\big[f(\xi_{j_n})\mathbb{1}_{[\xi_1,\ldots,\xi_{j_n-1} \neq \mathbf{1},\mathbf{0}]}\big].$$

Here, our usage of the compact notation on the right-hand side is analogous to the convention in the modern theory of stochastic processes. Recall that $|K^w|$ stands for the kernel $|K^w|(\eta, \tilde{\eta}) = |K^w(\eta, \tilde{\eta})|$, and put $|\Delta|_{\eta_0,U}^{w,\mathbf{j}'}$ for the measure on paths $(\eta_0, \ldots, \eta_U)$ obtained by replacing all the K^w in $\Delta_{U,\eta_0}^{w,\mathbf{j}'}$ by $|K^w|$. Since $\|f\|_\infty = 1$, the foregoing display implies

$$\begin{aligned}
|Q_{j_n}^{w,\mathbf{j}} f(\eta_0)| &\leq |\Delta|_{j_n,\eta_0}^{w,\mathbf{j}}(\xi_1, \ldots, \xi_{j_n-1} \neq \mathbf{1}, \mathbf{0}) \\
&\leq |\Delta|_{j_{n-1},\eta_0}^{w,(j_1,\ldots,j_{n-1})}(\xi_1, \ldots, \xi_{j_{n-1}-1} \neq \mathbf{1}, \mathbf{0}) \\
&\quad \times \Big(\sup_{\eta \neq \mathbf{1},\mathbf{0}} \mathbb{P}_\eta(\xi_{j_n - j_{n-1} - 1} \neq \mathbf{1}, \mathbf{0})\Big) \||K^w|\mathbb{1}\|_\infty
\end{aligned}$$

with $j_0 = 0$, where $\sup_{\eta \neq \mathbf{1},\mathbf{0}} \mathbb{P}_\eta(\xi_{j_n - j_{n-1} - 1} \neq \mathbf{1}, \mathbf{0})$ bounds the measure of the yet "active" paths from j_{n-1} to $j_n - 1$ and $\||K^w|\mathbb{1}\|_\infty$ bounds the measure of the transition from $j_n - 1$ to j_n. Iterating the last inequality, we get

$$
\begin{aligned}
|Q_{j_n}^{w,\mathbf{j}} f(\eta_0)| &\le \| |K^w| \mathbb{1} \|_\infty^n \prod_{r=1}^n \Big(\sup_{\eta \neq \mathbf{1},\mathbf{0}} \mathbb{P}_\eta(\xi_{j_r - j_{r-1}-1} \neq \mathbf{1},\mathbf{0}) \Big) \\
&\le (4C_0)^n w^n \prod_{r=1}^n \Big(\sup_{\eta \neq \mathbf{1},\mathbf{0}} \mathbb{P}_\eta(\xi_{j_r - j_{r-1}-1} \neq \mathbf{1},\mathbf{0}) \Big),
\end{aligned} \tag{3.10}
$$

where the last inequality follows from Lemma 3.1. Since $\sum_{r=1}^n (j_r - j_{r-1} - 1) = j_n - n$ and $C = \max(4C_0, C_1)$, Lemma 3.5 applied to the right-hand side of (3.10) gives

$$
|Q_{j_n}^{w,\mathbf{j}} f(\eta_0)| \le w^n (C^2)^n \delta^{j_n - n}. \tag{3.11}
$$

The proof of (3.8) is complete. $\square$

LEMMA 3.7. *Recall the constants $C = C(G, C_0)$ and $\delta = \delta(G)$ in Lemmas 3.6 and 3.5, respectively. There exists $w_1 \in (0, w_0]$ such that*

$$
\sum_{n=1}^\infty \sum_{\mathbf{j} \in \mathcal{I}(n)} w_1^n (C^2)^n \delta^{j_n - n} < \infty. \tag{3.12}
$$

PROOF. Observe that every index in $\bigcup_{n=1}^\infty \mathcal{I}(n)$ can be identified uniquely by the time of the last fault and the fault sequence before the time of the last fault. Hence, letting S denote the time of the last fault and m the number of faults within $\{1, \dots, S-1\}$, we can write for any $w > 0$

$$
\sum_{n=1}^\infty \sum_{\mathbf{j} \in \mathcal{I}(n)} w^n (C^2)^n \delta^{j_n - n} = \sum_{S=1}^\infty \sum_{m=0}^{S-1} \binom{S-1}{m} w^{m+1} (C^2)^{m+1} \delta^{S-m-1}. \tag{3.13}
$$

For each S, write

$$
\sum_{m=0}^{S-1} \binom{S-1}{m} w^{m+1} (C^2)^{m+1} \delta^{S-m-1} = w C^2 \delta^{S-1} (1 + w C^2 \delta^{-1})^{S-1}. \tag{3.14}
$$

With $\delta \in (0, 1)$ fixed, we can choose $w_1 \in (0, w_0]$ small such that

$$
C^2 (1 + w_1 C^2 \delta^{-1})^{S-1} \le \left(\frac{1}{\sqrt{\delta}} \right)^{S-1}
$$

for any large S. Apply the foregoing inequality for large S to the right-hand side of (3.14) with w replaced by w_1. This gives

$$
\sum_{m=0}^{S-1} \binom{S-1}{m} w_1^{m+1} (C^2)^{m+1} \delta^{S-m-1} \le w_1 \big(\sqrt{\delta}\big)^{S-1},
$$

where the right-hand side converges exponentially fast to 0 as $\delta < 1$. By (3.13), the asserted convergence of the series in (3.12) now follows. □

PROOF OF PROPOSITION 3.2. We pick w_1 as in the statement of Lemma 3.7. By Lemma 3.6 and the choice of w_1, the series in (3.6) converges absolutely and uniformly in $w \in [0, w_1]$ and $\eta \in \{1, 0\}^{\mathsf{V}}$.

By (3.5) and dominated convergence, it remains to show that

$$\lim_{T\to\infty}\sum_{n=1}^{T}\sum_{\mathbf{j}\in\mathcal{I}_T(n)} Q_T^{w,\mathbf{j}} f(\eta_0) = \sum_{n=1}^{\infty}\sum_{\mathbf{j}\in\mathcal{I}(n)} Q_{j_n}^{w,\mathbf{j}} f(\eta_0).$$

To see this, note that by Lemma 3.4, we can write

$$\sum_{n=1}^{T}\sum_{\mathbf{j}\in\mathcal{I}_T(n)} Q_T^{w,\mathbf{j}} f(\eta_0) = \sum_{n=1}^{T}\sum_{\substack{\mathbf{j}\in\mathcal{I}(n)\\ j_n\leq T}} Q_{j_n}^{w,\mathbf{j}} f(\eta_0),$$

where the right-hand side is a partial sum of the infinite series in (3.6). The validity of (3.6) now follows from the absolute convergence of the series in the same display. The proof is complete. □

For the convenience of subsequent applications, we consider from now on the continuous-time Markov chain (ξ_s) with rates given by (A.2). We can define this chain (ξ_s) from the discrete-time Markov chain (ξ_n) by

$$\xi_s = \xi_{M_s},$$

where (M_s) is an independent Poisson process with $\mathbb{E}[M_s] = \frac{sN}{\lambda}$. (Recall our time scale convention: "n" for the discrete time scale and "s" for the continuous time scale.) Under this setup, the potential measure of (ξ_s) and the potential measure of (ξ_n) are linked by

$$\mathbb{E}\left[\int_0^\infty f(\xi_s)\,ds\right] = \frac{\lambda}{N}\mathbb{E}\left[\sum_{n=0}^{\infty} f(\xi_n)\right] \tag{3.15}$$

for any nonnegative f. In addition, the fixation probability to $\mathbf{1}$ for this continuous-time Markov chain (ξ_s) is the same as that for the discrete-time chain (ξ_n). (See the discussion after Proposition 2.2.)

We now state a first-order approximation of $\mathbb{P}^w(\tau_{\mathbf{1}} < \infty)$ by the voter model. Recall the difference kernel D defined by (1.5) with h_i as in (A.1) and the π-expectation $\overline{D}$ defined by (1.6).

THEOREM 3.8. *Let ν be an arbitrary distribution on $\{1, 0\}^{\mathsf{V}}$. Then as $w \longrightarrow 0+$, we have*

$$\mathbb{P}_\nu^w(\tau_{\mathbf{1}} < \infty) = \mathbb{P}_\nu(\tau_{\mathbf{1}} < \infty) + w\int_0^\infty \mathbb{E}_\nu[\overline{D}(\xi_s)]\,ds + O(w^2). \tag{3.16}$$

Here, the convention for the function $O(w^2)$ is as in Theorem 1. Moreover, $\overline{D}(\xi_s) \in L_1(d\mathbb{P}_\nu \otimes ds)$.

PROOF. It suffices to prove the theorem for $\nu = \delta_{\eta_0}$ for any $\eta_0 \in \{1,0\}^{\mathsf{V}}$. Recall that the function p_1 defined by (2.6) is harmonic for the voter model, and hence, the expansion (3.6) applies. By (3.8) and Lemma 3.7, it is plain that

$$\sum_{n=2}^{\infty} \sum_{\mathbf{j} \in \mathcal{I}(n)} Q_{j_n}^{w,\mathbf{j}} p_1(\eta_0) = O(w^2). \tag{3.17}$$

We identify each $\mathbf{j} \in \mathcal{I}(1)$ as $\mathbf{j} = (j) = j$ and look at the summands $Q_j^{w,j} p_1$. Write $\mathbb{E}^{w,j}$ for the expectation of the time-inhomogeneous Markov chain where the transition of each step is governed by P except that the transition from $j-1$ to j is governed by P^w. Then

$$\begin{aligned} Q_j^{w,j} p_1(\eta_0) &= \mathbb{E}_{\eta_0}^{w,j}[p_1(\xi_j)] - \mathbb{E}_{\eta_0}[p_1(\xi_j)] \\ &= \mathbb{E}_{\eta_0}\big[\mathbb{E}_{\xi_{j-1}}^{w}[p_1(\xi_1)] - \mathbb{E}_{\xi_{j-1}}[p_1(\xi_1)]\big] \\ &= \mathbb{E}_{\eta_0}[K^w p_1(\xi_{j-1})] \\ &= \mathbb{E}_{\eta_0}[K^w p_1(\xi_{j-1}); \tau_{\mathbf{1}} \wedge \tau_{\mathbf{0}} \geq j], \end{aligned} \tag{3.18}$$

where the last equality follows from the definition of K^w and the fact that $\mathbf{1}$ and $\mathbf{0}$ are both absorbing. Moreover, we deduce from Lemma 3.1 that

$$K^w p_1(\eta) = \frac{\lambda}{N} w \overline{D}(\eta) + \frac{\lambda}{N} w^2 \overline{G}_w(\eta), \tag{3.19}$$

where

$$G_w(x,\eta) = g_w(x,\eta)\big(1 - 2\eta(x)\big).$$

Note that $\mathbb{E}_{\eta_0}[\tau_{\mathbf{1}} \wedge \tau_{\mathbf{0}}] < \infty$ by Lemma 3.5. Hence, by (3.18) and (3.19), we deduce that

$$\begin{aligned} \sum_{j=1}^{\infty} Q_j^{w,j} p_1(\eta_0) &= \frac{\lambda w}{N} \sum_{j=1}^{\infty} \mathbb{E}_{\eta_0}[\overline{D}(\xi_{j-1})] + O(w^2) \\ &= w \mathbb{E}_{\eta_0}\left[\int_0^{\infty} \overline{D}(\xi_s)\, ds\right] + O(w^2), \end{aligned} \tag{3.20}$$

where the last equality follows from (3.15). Moreover, $\overline{D}(\xi_s) \in L_1(d\mathbb{P}_{\eta_0} \otimes ds)$. The approximation (3.16) for each $w \leq w_2$ for some small $w_2 \in (0, w_1]$ now follows from (3.17) and (3.20) applied to the expansion (3.6) for p_1. The proof is complete. $\square$

4. First-order approximations. In this section, we give the proof of Theorem 1. We consider only regular graphs throughout this section. (Recall that a graph is regular if all vertices have the same number of neighbors.)

As a preliminary, let us introduce the convenient notion of Bernoulli transforms and discuss its properties. For each $u \in [0,1]$, let μ_u be the Bernoulli product measure on $\{1,0\}^{\mathsf{V}}$ with density $\mu_u(\xi(x)=1)=u$. For any function $f:\{1,0\}^{\mathsf{V}} \longrightarrow \mathbb{R}$, define the *Bernoulli transform* of f by

$$\mathscr{B}f(u) := \int f\,d\mu_u = \sum_{n=0}^{N}\Biggl[\sum_{\eta:\#\{x;\eta(x)=1\}=n} f(\eta)\Biggr]u^n(1-u)^{N-n}, \qquad u \in [0,1]. \tag{4.1}$$

The Bernoulli transform of f uniquely determines the coefficients

$$A_f(n) := \sum_{\eta:\#\{x;\eta(x)=1\}=n} f(\eta), \qquad 0 \le n \le N.$$

Indeed, $\mathscr{B}f(0) = f(\mathbf{0}) = A_f(0)$ and for each $1 \le n \le N$,

$$A_f(n) = \lim_{u\downarrow 0+}\frac{1}{u^n}\Biggl(\mathscr{B}I(u) - \sum_{i=0}^{n-1} u^i(1-u)^{N-i}A_f(i)\Biggr).$$

The Bernoulli transform $\mathscr{B}f(u)$ is a polynomial $\sum_{i=0}^{N}\alpha_i u^i$ of order at most N. Let us invert the coefficients $A_f(n)$ from α_i by basic combinatorics. By the binomial theorem,

$$u^i = \sum_{n=0}^{N-i}\binom{N-i}{n}u^{i+n}(1-u)^{N-i-n}.$$

Hence, summing over $i+n$, we have

$$\sum_{i=0}^{N}\alpha_i u^i = \sum_{n=0}^{N}\Biggl[\sum_{i=0}^{n}\alpha_i\binom{N-i}{n-i}\Biggr]u^n(1-u)^{N-n},$$

and the uniqueness of the coefficients A_f implies

$$A_f(n) = \sum_{i=0}^{n}\alpha_i\binom{N-i}{n-i}, \qquad 0 \le n \le N. \tag{4.2}$$

As a corollary, we obtain

$$\int f\,d\mathbf{u}_n = \frac{1}{\binom{N}{n}}A_f(n) = \frac{1}{\binom{N}{n}}\sum_{i=0}^{n}\alpha_i\binom{N-i}{n-i}, \qquad 1 \le n \le N-1, \tag{4.3}$$

if we regard $\mathbf{u}_n$, the uniform distribution on the set of configurations with precisely n many 1's, as a measure on $\{1,0\}^{\mathsf{V}}$ in the natural way.

We will specialize the application of Bernoulli transforms to the function

$$I(\eta) := \int_0^\infty \mathbb{E}_\eta[\overline{D}(\xi_s)]\,ds.$$

To obtain the explicit approximations (up to the first order) asserted in Theorem 1, we need to compute by Theorem 3.8 the 0-potentials $\int_0^\infty \mathbb{E}_{\mathbf{u}_n}[\overline{D}(\xi_s)]\,ds$ for $1 \le n \le N-1$ under each updating mechanism. On the other hand, we will see that the Bernoulli transform of each 0-potential I is analytically tractable and

$$\mathscr{B}I(u) = \Gamma u(1-u) \tag{4.4}$$

for some explicit constant Γ. Note that we have $A_I(N) = A_I(0) = 0$ for the updating mechanisms under consideration. Hence, the formula (4.3) entails

$$\int_0^\infty \mathbb{E}_{\mathbf{u}_n}[\overline{D}(\xi_s)]\,ds = \frac{\Gamma n(N-n)}{N(N-1)}, \qquad 1 \le n \le N-1, \tag{4.5}$$

since $\binom{N-1}{n-1} - \binom{N-2}{n-2} = \binom{N-2}{n-1}$ for $n \ge 2$ and $\binom{N-1}{0} = \binom{N-2}{0} = 1$.

4.1. *Proof of Theorem* 1(1). Assume that the graph G is k-regular. Recall that the death–birth updating defines a Markov chain of voter model perturbation satisfying (A.1)–(A.4) by Proposition 2.1 under weak selection. The functions h_i in (A.1) for this updating are given by (2.3). Hence, the difference kernel D is given by $D(x,\xi) = h_1(x,\xi)$, and for any $x \in \mathsf{V}$,

$$\begin{aligned}
\frac{1}{k}\mathbb{E}_{\mu_u}[D(x,\xi_s)] &= -(b+c)\mathbb{E}_{\mu_u}[f_0 f_1(x,\xi_s)] + b\mathbb{E}_{\mu_u}[f_{00}(x,\xi_s)] \\
&\quad + b\mathbb{E}_{\mu_u}[f_0 f_{11}(x,\xi_s)] - b\mathbb{E}_{\mu_u}[f_0 f_{00}(x,\xi_s)] \\
&= -(b+c)\mathbb{E}_{\mu_u}[f_0 f_1(x,\xi_s)] + b\mathbb{E}_{\mu_u}[f_0 f_{11}(x,\xi_s)] \\
&\quad + b\mathbb{E}_{\mu_u}[f_1 f_{00}(x,\xi_s)].
\end{aligned} \tag{4.6}$$

In analogy to the computations in [6] for coalescing probabilities, we resort to the duality between voter models and coalescing random walks for the right-hand side of (4.6). Let $\{B^x; x \in \mathsf{V}\}$ be the rate-1 coalescing random walks on G, where B^x starts at x. The random walks move independently of each other until they meet another and move together afterward. The duality between the voter model and the coalescing random walks is given by

$$\mathbb{P}_\eta\big(\xi_s(x) = i_x, x \in Q\big) = \mathbb{P}\big(\eta(B_s^x) = i_x, x \in Q\big)$$

for any $Q \subset \mathsf{V}$ and $(i_x; x \in Q) \in \{1,0\}^Q$. (See Chapter V in [11].) Introduce two independent discrete-time random walks $(X_n; n \ge 0)$ and $(Y_n; n \ge 0)$ starting at the same vertex, both independent of $\{B^x; x \in \mathsf{V}\}$. Fix x and assume that the chains

(X_n) and (Y_n) both start at x. Recall that we write $\widehat{\eta} \equiv 1 - \eta$. Then by duality, we deduce from (4.6) that

$$
\begin{aligned}
\frac{1}{k}\mathbb{E}_{\mu_u}[D(x,\xi_s)] &= -(b+c)\int \mu_u(d\eta)\mathbb{E}[\widehat{\eta}(B_s^{X_1})\eta(B_s^{Y_1})] \\
&\quad + b\int \mu_u(d\eta)\mathbb{E}[\widehat{\eta}(B_s^{X_1})\eta(B_s^{Y_1})\eta(B_s^{Y_2})] \\
&\quad + b\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{X_1})\widehat{\eta}(B_s^{Y_1})\widehat{\eta}(B_s^{Y_2})] \\
&= -c\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{Y_1})] + c\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{X_1})\eta(B_s^{Y_1})] \\
&\quad + b\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{Y_1})\eta(B_s^{Y_2})] \\
&\quad - b\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{X_1})\eta(B_s^{Y_2})] \\
&\quad + b\int \mu_u(d\eta)\mathbb{E}[\eta(B_s^{X_1}) - \eta(B_s^{Y_1})].
\end{aligned}
\tag{4.7}
$$

For clarity, let us write from now on $\mathbf{P}_\rho$ and $\mathbf{E}_\rho$ for the probability measure and the expectation, respectively, under which the common initial position of (X_n) and (Y_n) is distributed as ρ. Recall that $\overline{D}(\xi)$ is the π-expectation of $x \longmapsto D(x,\xi)$ defined by (1.6). Write $M_{x,y} = \inf\{t \in \mathbb{R}_+; B_t^x = B_t^y\}$ for the first meeting time of the random walks B^x and B^y, so B^x and B^y coincide after $M_{x,y}$. Then from (4.7), the spatial homogeneity of the Bernoulli product measures implies that

$$
\begin{aligned}
\frac{1}{k}\mathbb{E}_{\mu_u}[\overline{D}(\xi_s)] &= -cu + c[u\mathbf{P}_\pi(M_{X_1,Y_1} \le s) + u^2\mathbf{P}_\pi(M_{X_1,Y_1} > s)] \\
&\quad + b[u\mathbf{P}_\pi(M_{Y_1,Y_2} \le s) + u^2\mathbf{P}_\pi(M_{Y_1,Y_2} > s)] \\
&\quad - b[u\mathbf{P}_\pi(M_{X_1,Y_2} \le s) + u^2\mathbf{P}_\pi(M_{X_1,Y_2} > s)] \\
&= -cu(1-u)\mathbf{P}_\pi(M_{X_1,Y_1} > s) - bu(1-u)\mathbf{P}_\pi(M_{Y_1,Y_2} > s) \\
&\quad + bu(1-u)\mathbf{P}_\pi(M_{X_1,Y_2} > s).
\end{aligned}
$$

To obtain $\mathscr{B}I$, we integrate both sides of the foregoing equality with respect to s over $\mathbb{R}_+$. This gives

$$
\mathscr{B}I(u) = ku(1-u)(-c\mathbf{E}_\pi[M_{X_1,Y_1}] - b\mathbf{E}_\pi[M_{Y_1,Y_2}] + b\mathbf{E}_\pi[M_{X_1,Y_2}]).
\tag{4.8}
$$

We now turn to a simple identity between first meeting times and first hitting times. Let $T_y = \inf\{n \ge 0; X_n = y\}$, the first hitting time of y by (X_n). Observe that for the random walks on (connected) regular graphs, the invariant distribution is uniform and $\mathbf{E}_x[T_y] = \mathbf{E}_y[T_x]$ for any x, y. Hence, the proof of Proposition 14.5 in [1] implies

$$
\mathbf{E}[M_{x,y}] = \tfrac{1}{2}\mathbf{E}_x[T_y], \qquad x, y \in \mathsf{V}.
\tag{4.9}
$$

Write

$$f(x,y) := \mathbf{E}_x[T_y] = \mathbf{E}_y[T_x], \qquad x, y \in \mathsf{V},$$

where $\mathbf{E}_x = \mathbf{E}_{\delta_x}$.

LEMMA 4.1. *For any* $z \in \mathsf{V}$,

$$\mathbf{E}_z[f(X_0, X_1)] = \mathbf{E}_z[f(Y_1, Y_2)] = N - 1, \tag{4.10}$$

$$\mathbf{E}_z[f(X_1, Y_1)] = \mathbf{E}_z[f(Y_0, Y_2)] = N - 2, \tag{4.11}$$

$$\mathbf{E}_z[f(X_1, Y_2)] = \left(1 + \frac{1}{k}\right)(N-1) + \frac{1}{k} - 2. \tag{4.12}$$

PROOF. The proof of the equality $\mathbf{E}_z[f(X_0, X_1)] = N - 1$ can be found in Chapter 3 of [1] or [12]. We restate its short proof here for the convenience of readers. Let $T_x^+ = \inf\{n \geq 1; X_n = x\}$ denote the first return time to x. A standard result of Markov chains says $\mathbf{E}_x[T_x^+] = \pi(x)^{-1} = N$ for any x. The equalities in (4.10) now follow from the Markov property.

Next, we prove (4.11). By (4.10) and the symmetry of f, we have

$$\begin{aligned} N - 1 = \mathbf{E}_z[f(X_0, X_1)] &= \sum_{x\sim z} \frac{1}{k}\mathbf{E}_z[T_x] \\ &= \sum_{x\sim z}\sum_{y\sim z} \frac{1}{k^2}(\mathbf{E}_y[T_x] + 1) = \mathbf{E}_z[f(Y_1, X_1)] + 1, \end{aligned}$$

so $\mathbf{E}_z[f(X_1, Y_1)] = N - 2$. Here, our summation notation $\sum_{x\sim z}$ means summing over indices x with z fixed, and the same convention holds in the proof of (4.12) and Section 5 below. A similar application of the Markov property to the coordinate Y_1 in $\mathbf{E}_z[f(Y_0, Y_1)]$ gives $\mathbf{E}_z[f(Y_0, Y_2)] = N - 2$. This proves (4.11).

Finally, we need to prove (4.12). We use (4.10) and (4.11) to get

$$\begin{aligned} \mathbf{E}_z[f(X_0, X_1)] &= 1 + \mathbf{E}_z[f(X_1, Y_1)] \\ &= 1 + \sum_{x\sim z}\sum_{\substack{y\sim z \\ y\neq x}} \frac{1}{k^2}\mathbf{E}_x[T_y] \\ &= 1 + \sum_{x\sim z}\sum_{\substack{y\sim z \\ y\neq x}} \frac{1}{k^2}\left(\sum_{w\sim y}\frac{1}{k}\mathbf{E}_x[T_w] + 1\right) \qquad (4.13) \\ &= 1 + \sum_{x\sim z}\sum_{\substack{y\sim z \\ y\neq x}} \frac{1}{k^2} + \sum_{x\sim z}\sum_{y\sim z}\sum_{w\sim y}\frac{1}{k^3}\mathbf{E}_x[T_w] - \sum_{x\sim z}\sum_{w\sim x}\frac{1}{k^3}\mathbf{E}_x[T_w] \\ &= 2 - \frac{1}{k} + \mathbf{E}_z[f(X_1, Y_2)] - \frac{1}{k}\mathbf{E}_z[f(X_1, X_0)]. \qquad (4.14) \end{aligned}$$

Here, in (4.13) we use the symmetry of f, and the last equality follows from (4.10). A rearrangement of both sides of (4.14) and an application of (4.10) then lead to (4.12), and the proof is complete. $\square$

Apply Lemma 4.1 and (4.9) to (4.8), and we obtain the following result.

PROPOSITION 4.2. *For any* $u \in [0,1]$,

$$\mathscr{B}I(u) = \frac{ku(1-u)}{2}\left[\left(\frac{b}{k} - c\right)(N-2) + b\left(\frac{2}{k} - 2\right)\right]. \tag{4.15}$$

Finally, since $\mathscr{B}I(u)$ takes the form (4.4), we may apply (4.5) and Proposition 4.2 to obtain the explicit formula for the coefficient of w in (3.16), subject to each initial distribution $\mathbf{u}_n$. This proves our assertion in Theorem 1(1).

4.2. *Proof of Theorem* 1(2). The proof of Theorem 1(2) follows from almost the same argument for Theorem 1(1) except for more complicated arithmetic. For this reason, we will only point out the main steps, leaving the detailed arithmetic to the interested readers. In the following, we continue to use the notation for the random walks in the proof of Theorem 1(1).

Fix $x \in \mathsf{V}$ and assume the chains (X_n) and (Y_n) both start at x. By Proposition 2.2, we have

$$\begin{aligned}
&\frac{1}{k}\mathbb{E}_{\mu_u}[D(x,\xi_s)] \\
&\quad = \mathbb{E}_{\mu_u}[(b-c)\widehat{\xi}_s(x) f_{11}(x,\xi_s) - c\widehat{\xi}_s(x) f_{10}(x,\xi_s) - b\xi_s(x) f_{01}(x,\xi_s)] \\
&\qquad - \frac{k}{k+1}\mathbb{E}_{\mu_u}\big[\big((b-c) f_{11}(x,\xi_s) - c f_{10}(x,\xi_s) + b f_{01}(x,\xi_s)\big) \\
&\qquad\qquad \times \big(\widehat{\xi}_s(x) f_1(x,\xi_s) - \xi_s(x) f_0(x,\xi_s)\big)\big] \\
&\qquad - \frac{1}{k+1}\mathbb{E}_{\mu_u}\big[b\widehat{\xi}_s(x) f_1^2(x,\xi_s) \\
&\qquad\qquad - \xi_s(x) f_0(x,\xi_s)\big((b-c) f_1(x,\xi_s) - c f_0(x,\xi_s)\big)\big] \\
&\quad = \int \mu_u(d\eta)\Big(\mathbb{E}[(b-c)\widehat{\eta}(B_s^x)\eta(B_s^{Y_1})\eta(B_s^{Y_2}) - c\widehat{\eta}(B_s^x)\eta(B_s^{Y_1})\widehat{\eta}(B_s^{Y_2}) \\
&\qquad\qquad - b\eta(B_s^x)\widehat{\eta}(B_s^{Y_1})\eta(B_s^{Y_2})] \\
&\qquad\qquad - \frac{k}{k+1}\mathbb{E}\big[\big((b-c)\eta(B_s^{Y_1})\eta(B_s^{Y_2}) - c\eta(B_s^{Y_1})\widehat{\eta}(B_s^{Y_2}) \\
&\qquad\qquad\qquad + b\widehat{\eta}(B_s^{Y_1})\eta(B_s^{Y_2})\big) \\
&\qquad\qquad\qquad \times \big(\eta(B_s^{X_1}) - \eta(B_s^x)\big)\big]
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{k+1}\mathbb{E}[b\widehat{\eta}(B_s^x)\eta(B_s^{X_1})\eta(B_s^{Y_1}) \\
&\qquad - (b-c)\eta(B_s^x)\widehat{\eta}(B_s^{X_1})\eta(B_s^{Y_1}) \\
&\qquad\qquad + c\eta(B_s^x)\widehat{\eta}(B_s^{X_1})\widehat{\eta}(B_s^{Y_1})]\Big),
\end{aligned}
$$

where the last equality follows again from duality. The last equality gives

$$
\begin{aligned}
&\frac{1}{k}\mathbb{E}_{\mu_u}[D(x,\xi_s)] \\
&\quad = \int \mu_u(d\eta)\mathbb{E}\Big[b\eta(B_s^{Y_1})\eta(B_s^{Y_2}) + \frac{c+b}{k+1}\eta(B_s^x)\eta(B_s^{Y_1}) \\
&\qquad - c\eta(B_s^{Y_1}) - \frac{b}{k+1}\eta(B_s^x)\eta(B_s^{Y_2}) + \frac{kc-b}{k+1}\eta(B_s^{X_1})\eta(B_s^{Y_1}) \\
&\qquad - \frac{kb}{k+1}\eta(B_s^{X_1})\eta(B_s^{Y_2}) + \frac{c}{k+1}\eta(B_s^x)\eta(B_s^{X_1}) \\
&\qquad - \frac{c}{k+1}\eta(B_s^x)\Big].
\end{aligned}
$$

Recall that $X_1 \overset{(\mathrm{d})}{=} Y_1$. Hence, by the definition of $\overline{D}$ and $\mathbf{P}_\pi$, the foregoing implies that

$$
\begin{aligned}
\frac{1}{k}\mathbb{E}_{\mu_u}[\overline{D}(\xi_s)] = {}& -bu(1-u)\mathbf{P}_\pi(M_{Y_1,Y_2} > s) \\
& - \frac{2c+b}{k+1}u(1-u)\mathbf{P}_\pi(M_{X_0,X_1} > s) \\
& + \frac{b}{k+1}u(1-u)\mathbf{P}_\pi(M_{Y_0,Y_2} > s) \\
& - \frac{kc-b}{k+1}u(1-u)\mathbf{P}_\pi(M_{X_1,Y_1} > s) \\
& + \frac{kb}{k+1}u(1-u)\mathbf{P}_\pi(M_{X_1,Y_2} > s).
\end{aligned}
$$

Again, we integrate both sides of the foregoing display with respect to s and then apply (4.9) and Lemma 4.1 for the explicit form of $\mathscr{B}I$. The result is given by the following proposition.

PROPOSITION 4.3. *For any* $u \in [0,1]$,

$$
\mathscr{B}I(u) = \frac{k(k+2)u(1-u)}{2(k+1)}\left[\left(\frac{b}{(k+2)} - c\right)(N-1) - \frac{(2k+1)b - ck}{k+2}\right].
$$

Our assertion for Theorem 1(2) now follows from an application of Proposition 4.3 similar to that of Proposition 4.2 for Theorem 1(1). The proof is now complete.

5. Proofs of Propositions 2.1 and 2.2.

5.1. *Proof of Proposition* 2.1. Suppose that $\xi \in \{1,0\}^{\mathsf{V}}$ is the present configuration on the graph. Let $n_i(x) = n_i(x,\xi)$ be the number of neighboring i players for an individual located at vertex x for $i = 1, 0$. Let $w \in [0,1]$ denote the intensity of selection. By definition, the fitness $\rho_i(x) = \rho_i(x,\xi)$ of an i-player located at x is given by

$$\rho_i(x) = (1-w) + w\,[\,\Pi_{i1} \quad \Pi_{i0}\,]\begin{bmatrix} n_1(x) \\ n_0(x) \end{bmatrix} = (1-w) + w\boldsymbol{\Pi}_i\mathbf{n}(x). \tag{5.1}$$

Here, $\boldsymbol{\Pi}_i$ is the payoff row of an i-player of the matrix $\boldsymbol{\Pi}$ and $\mathbf{n}(x)$ is the column vector $[n_1(x) \;\; n_0(x)]^\top$. Hence, there exists $w_0 > 0$ depending only on k and $\boldsymbol{\Pi}$ such that $\rho_i > 0$ for every $w \in [0, w_0]$ (see [6]).

The game with the death–birth updating under weak selection defines a Markov chain with transition probabilities P^w taking the form (2.2) and

$$c^w(x,\xi) = r_{1-\xi(x)}(x,\xi) \geq 0, \tag{5.2}$$

$$r_i(x,\xi) = \frac{\sum_{y\sim x}\rho_i(y)\mathbb{1}_{\xi(y)=i}}{\sum_{y\sim x}[\rho_1(y)\xi(y) + \rho_0(y)\widehat{\xi}(y)]}. \tag{5.3}$$

It has been shown in Section 1.4 of [6] that the rates c^w define voter model perturbations satisfying (A.1) and (A.2). Moreover, $\lambda = 1$ and the functions h_i in the expansion (A.2) are given by (2.3). Plainly, $c^w(x,\mathbf{1}) \equiv r_0(x,\mathbf{1}) \equiv 0$ and $c^w(x,\mathbf{0}) \equiv r_1(x,\mathbf{0}) \equiv 0$. Hence, (A.3) is also satisfied.

It remains to check that (A.4) is satisfied. Since (A.4) is satisfied when $w = 0$, it is enough to show that

$$P(\xi,\xi^x) > 0 \quad \Longleftrightarrow \quad P^w(\xi,\xi^x) > 0 \tag{5.4}$$

for any $\xi \neq \mathbf{1}, \mathbf{0}$ and any x. However, this is immediate from (5.3) if we notice that $\rho_i(\cdot)$ and the constant function 1, both regarded as measures on V in the natural way, are equivalent. Our proof of Proposition 2.1 is complete.

REMARK 5.1. Suppose now that payoff is given by a general 2×2 payoff matrix $\boldsymbol{\Pi}^* = (\Pi^*_{ij})_{i,j=1,0}$ subject only to the "equal-gains-from-switching" condition (1.10). Let us explain how to reduce the games with payoff matrix $\boldsymbol{\Pi}^*$ to the games with payoff matrix $\boldsymbol{\Pi}^a$ under weak selection, where $\boldsymbol{\Pi}^a$ is defined by (1.11).

In this case, payoffs of players are as described in Remark 1.1, and fitness is given by

$$\rho_i^{\boldsymbol{\Pi}^*}(x) = (1-w) + w\boldsymbol{\Pi}^*_i\mathbf{n}(x), \qquad x \in \mathsf{V}. \tag{5.5}$$

Here again, $\mathbf{\Pi}_i^*$ is the payoff row of an i-player. We put the superscript $\mathbf{\Pi}^*$ (only in this remark) to emphasize the dependence on the underlying payoff matrix $\mathbf{\Pi}^*$, so, in particular, the previously defined fitness ρ in (5.1) is equal to $\rho^{\mathbf{\Pi}}$.

Suppose that the graph is k-regular. The transition probabilities under the death–birth updating with payoff matrix $\mathbf{\Pi}^*$ are defined in the same way as before through (5.2) and (5.3) with ρ replaced by $\rho^{\mathbf{\Pi}^*}$. Note that $n_1(x) + n_0(x) \equiv k$. Then for all small w

$$\begin{aligned}\frac{1}{1-(1-k\Pi_{00}^*)w}\rho_i^{\mathbf{\Pi}^*}(x) &= 1 + \frac{w}{1-(1-k\Pi_{00}^*)w}\mathbf{\Pi}_i^a\mathbf{n}(x)\\ &= 1 + \frac{w^a}{1-w^a}\mathbf{\Pi}_i^a\mathbf{n}(x)\\ &= \frac{1}{1-w^a}\rho_i^{\mathbf{\Pi}^a}(x)\end{aligned}$$

for some w^a. Here, w and w^a are defined continuously in terms of each other by

$$w^a = \frac{w}{1+k\Pi_{00}^* w} \quad \text{and} \quad w = \frac{w^a}{1-k\Pi_{00}^* w^a},$$

so $\lim_{w^a\to 0} w = \lim_{w\to 0} w^a = 0$. Consequently, by (5.2) and (5.3), the foregoing display implies that the death–birth updating with payoff matrix $\mathbf{\Pi}^*$ and intensity of selection w is "equivalent" to the death–birth updating with payoff matrix $\mathbf{\Pi}^a$ and intensity of selection w^a, whenever w^a or w is small. Here, "equivalent" means equality of transition probabilities.

A similar reduction applies to the imitation updating by using its formal definition described in the next subsection, and we omit the details.

5.2. *Proof of Proposition* 2.2. Under the imitation updating, the Markov chain of configurations has transition probabilities given by

$$\begin{aligned}P^w(\xi,\xi^x) &= \frac{1}{N}d^w(x,\xi),\\ P^w(\xi,\xi) &= 1 - \frac{1}{N}\sum_x d^w(x,\xi),\end{aligned} \tag{5.6}$$

where

$$d^w(x,\xi) = s_{1-\xi(x)}(x,\xi), \tag{5.7}$$

$$s_i(x,\xi) = \frac{\sum_{y\sim x}\rho_i(y)\mathbb{1}_{\xi(y)=i}}{\sum_{y\sim x}[\rho_1(y)\xi(y)+\rho_0(y)\widehat{\xi}(y)] + \rho_{1-i}(x)} \tag{5.8}$$

and the fitness ρ_i are defined as before by (5.1). We assume again that the intensity of selection w is small such that $\rho_i > 0$. To simplify notation, let us set the column vectors

$$\mathbf{f}(x) = [\, f_1(x) \quad f_0(x)\,]^\top,$$
$$\mathbf{f}_{i\bullet}(x) = [\, f_{i1}(x) \quad f_{i0}(x)\,]^\top,$$
$$\mathbf{n}_{i\bullet}(x) = [\, n_{i1}(x) \quad n_{i0}(x)\,]^\top,$$

where the functions f_i and f_{ij} are defined by (2.4). By (5.1) and (5.8), we have

$$\begin{aligned} s_i(x,\xi) &= \frac{(1-w)n_i(x) + w\boldsymbol{\Pi}_i\mathbf{n}_{i\bullet}(x)}{(1-w)(k+1) + w\sum_{j=0}^{1}\boldsymbol{\Pi}_j\mathbf{n}_{j\bullet}(x) + w\boldsymbol{\Pi}_{1-i}\mathbf{n}(x)} \\ &= \frac{(1-w)k/(k+1)f_i(x) + k^2/(k+1)w\boldsymbol{\Pi}_i\mathbf{f}_{i\bullet}(x)}{(1-w) + wk^2/(k+1)\sum_{j=0}^{1}\boldsymbol{\Pi}_j\mathbf{f}_{j\bullet}(x) + wk/(k+1)\boldsymbol{\Pi}_{1-i}\mathbf{f}(x)} \\ &= \frac{k/(k+1)f_i(x) + w(k^2/(k+1)\boldsymbol{\Pi}_i\mathbf{f}_{i\bullet}(x) - k/(k+1)f_i(x))}{1 + w(k^2/(k+1)\sum_{j=0}^{1}\boldsymbol{\Pi}_j\mathbf{f}_{j\bullet}(x) + k/(k+1)\boldsymbol{\Pi}_{1-i}\mathbf{f}(x) - 1)}. \end{aligned}$$

Note that the functions f_i and f_{ij} are uniformly bounded. Apply Taylor's expansion in w at 0 to the right-hand side of the foregoing display. We deduce from (5.7) that the transition probabilities (5.6) takes the form (2.2) with $\lambda = \frac{k}{k+1}$ and the rates c^w satisfying (A.1) and (A.2) for some small w_0. Moreover, the functions h_i are given by

$$\begin{aligned} h_i &= (k\boldsymbol{\Pi}_i\mathbf{f}_{i\bullet} - f_i) - f_i\left(\frac{k^2}{k+1}\sum_{j=0}^{1}\boldsymbol{\Pi}_j\mathbf{f}_{j\bullet} + \frac{k}{k+1}\boldsymbol{\Pi}_{1-i}\mathbf{f} - 1\right) \\ &= k\boldsymbol{\Pi}_i\mathbf{f}_{i\bullet} - \frac{k^2}{k+1}f_i\left(\sum_{j=0}^{1}\boldsymbol{\Pi}_j\mathbf{f}_{j\bullet}\right) - \frac{k}{k+1}f_i\boldsymbol{\Pi}_{1-i}\mathbf{f}. \end{aligned}$$

By the definition of $\boldsymbol{\Pi}$ in (1.1), we get (2.5). The verifications of (A.3) and (A.4) follow from similar arguments for those of (A.3) and (A.4) under the death–birth updating, respectively. This completes the proof of Proposition 2.2.

Acknowledgments. I am grateful to my advisor, Professor Ed Perkins, for inspiring discussions and to Christoph Hauert and the anonymous referees for valuable comments.

REFERENCES

[1] ALDOUS, D. and FILL, J. A. (1994). Reversible Markov chains and random walks on graphs. Unpublished manuscript. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. Springer, New York. MR1978607

[3] AXELROD, R. and HAMILTON, W. D. (1981). The evolution of cooperation. *Science* **211** 1390–1396. MR0686747

[4] BOLLOBÁS, B. (1998). *Modern Graph Theory*. *Graduate Texts in Mathematics* **184**. Springer, New York. MR1633290

[5] CHEN, Y. T. (2010). The b/c rule under the imitation updating on $\mathbb{Z}^d$ for $d \geq 3$. Unpublished manuscript.

[6] COX, J. T., DURRETT, R. and PERKINS, E. A. (2011). Voter model perturbations and reaction diffusion equations. Available at http://arxiv.org/abs/1103.1676.

[7] COX, J. T. and PERKINS, E. A. (2005). Rescaled Lotka–Volterra models converge to super-Brownian motion. *Ann*. *Probab*. **33** 904–947. MR2135308

[8] HAMILTON, W. D. (1964). The genetical evolution of social behaviour. I. *J*. *Theoret*. *Biol*. **7** 1–16.

[9] HAUERT, C. (2008). Evolutionary dynamics. In *Evolution from Cellular to Social Scales*, *NATO Science for Peace and Security Series* 11–44. Springer, Dordrecht.

[10] HOFBAUER, J. and SIGMUND, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge Univ. Press, Cambridge. MR1635735

[11] LIGGETT, T. M. (2005). *Interacting Particle Systems*. Springer, Berlin. MR2108619

[12] LOVÁSZ, L. (1996). Random walks on graphs: A survey. In *Combinatorics*, *Paul Erdős Is Eighty*, *Vol*. 2 (*Keszthely*, 1993). *Bolyai Society Mathematical Studies* **2** 353–397. János Bolyai Math. Soc., Budapest. MR1395866

[13] MAYER, J. E. and MONTROLL, E. (1941). Molecular distribution. *J*. *Chem*. *Phys*. **9** 2–16.

[14] MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*. Cambridge Univ. Press, Cambridge.

[15] NOWAK, M. and SIGMUND, K. (1990). The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl*. *Math*. **20** 247–265. MR1081589

[16] NOWAK, M. A. (2006). Five rules for the evolution of cooperation. *Nature* **314** 1560–1563.

[17] NOWAK, M. A. (2006). *Evolutionary Dynamics*: *Exploring the Equations of Life*. The Belknap Press of Harvard Univ. Press, Cambridge, MA. MR2252879

[18] OHTSUKI, H., HAUERT, C., LIEBERMAN, E. and NOWAK, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441** 502–505.

[19] OHTSUKI, H. and NOWAK, M. A. (2006). Evolutionary games on cycles. *Proc*. *Biol*. *Sci*. **273** 2249–2256.

[20] SIGMUND, K. (2010). *The Calculus of Selfishness*. Princeton Univ. Press, Princeton, NJ. MR2590059

[21] TAYLOR, P. D., DAY, T. and WILD, G. (2007). Evolution of cooperation in a finite homogeneous graph. *Nature* **447** 469–472.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
1984 MATHEMATICS RD.
VANCOUVER, BRITISH COLUMBIA
CANADA V6T 1Z2
E-MAIL: ytchen@math.ubc.ca